THE CENTENNIAL EVOLUTION OF GEOMAGNETIC ACTIVITY AND ITS DRIVING MECHANISMS

LEONIE JOHANNA LISA PICK, M.SC.

Kumulative Dissertation
zur Erlangung des akademischen Grades
„doctor rerum naturalium“
(Dr. rer. nat.)
in der Wissenschaftsdisziplin „Geophysik“

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
Institut für Geowissenschaften
der Universität Potsdam

Disputation: Potsdam, 19.06.2020

Hauptbetreuerin: Prof. Dr. Claudia Stolle

Zweitbetreuerin: Dr. Monika Korte

Gutachter*innen: Prof. Dr. Claudia Stolle
Dr. Monika Korte
Dr. Timo Asikainen

Published online in the
Institutional Repository of the University of Potsdam:
https://doi.org/10.25932/publishup-47275
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-472754

"The construction of meaning is a fundamentally individual, subjective, creative enterprise, and an intimidating responsibility."

– Sean Carroll, The Big Picture: On the Origins of Life, Meaning, and the Universe Itself, 2017

ACKNOWLEDGMENTS

First and foremost I would like to thank my supervisors Prof. Dr. Claudia Stolle, who provided valuable strategical guidance in the broader scientific context, and Dr. Monika Korte, who supported me continuously and thoughtfully on a day-to-day basis.
My sincere gratitude also goes to my mentor Prof. Dr. Matthias Hort from the Institute of Geophysics at the University of Hamburg for his very helpful advice in both professional and personal matters.
I am grateful for the eager and productive help of my former student assistant Yannik Thomas. Furthermore, I profited greatly from the competent input of Prof. Dr. Christopher Finlay and Prof. Dr. Nils Olsen, who generously welcomed me at DTU Space in Copenhagen and took time for in-depth discussions. The same applies to Dr. Natalie Krivova and Dr. Chi-Ju Wu from the Max Planck Institute for Solar System Research in Göttingen with whom I found it very enjoyable and fruitful to collaborate.
I am thankful for the professional interest, solidarity and optimism I was able to experience as a member of DFG's priority program 1788 "Dynamic Earth". In the framework of the program, Dr. Mirjam Langhans and Dr. Stefanie Weege did an excellent job of organizing the annual progress meetings.
I am very fortunate to have shared the office with Dr. Julien Baerenzung, Dr. Sanja Panovska, and Dr. Foteini Vervelidou, whose way of working impressed and inspired me - I hope some of it rubbed off on me! In addition to being another great office mate, Irina Zhelavskaya supported me with her expertise and positive attitude, together with Dr. Frederic Effenberger. In general, the integrative and friendly working atmosphere in GFZ's sections 2.3 "Geomagnetism" and 2.8 "Magnetospheric physics" was an important factor for the completion of this work. Particularly, I would like to acknowledge the exceptional commitment of Alexander Jordan, Martina Krüger, and Antje Timmermann.
Working at GFZ in Potsdam has definitely been a personal gain for me. I am happy to have shared this positive and instructive experience with Dr. Tarique Siddiqui, Robin Senftleben, and Juan Rodríguez-Zuluaga as well as many other fellow Ph.D. students from the whole of GFZ.
I am especially thankful for the huge support of Prof. Dr. Joachim Vogt from the Jacobs University in Bremen during the last phase of my Ph.D. time. The conversations with him and Dr. Adrian Blagau from ISS Bucharest-Magurele were extremely valuable and have encouraged me to finalize this work.
Lastly, I acknowledge the work of André Miede and Ivo Pletikosić, whose LaTeX-template I used to compile the thesis.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS & ABBREVIATIONS

ACE	Advanced Composition Explorer
AMC	Annual Magnetospheric Currents Index
BGS	British Geological Survey
CIR	Co-rotating Interaction Region
DISCOVR	Deep Space Climate Observatory
DPS	Dessler-Parker-Sckopke
Dst	Disturbed Storm Time Index
EEJ	Equatorial Electrojet
F10.7	Solar Radio Flux
FAC	Field-aligned Current
GEO	Geographic Coordinate System
GSM	Geocentric Solar Magnetic Coordinate System
HCS	Heliospheric Current Sheet
HMC	Hourly Magnetospheric Currents Index
HSS	High-speed Stream
IAGA	International Association For Geomagnetism And Aeronomy
ICME	Interplanetary Coronal Mass Ejection
IGRF	International Geomagnetic Reference Field
IMF	Interplanetary Magnetic Field
LEO	Low Earth Orbit
LT	Local Time
MAG	Geomagnetic Coordinate System

MLT	Magnetic Local Time
OSF	Open Solar Magnetic Flux
PRC	Partial Ring Current
SC	Solar Cycle
SH	Spherical Harmonics
SIR	Stream Interaction Region
SM	Solar Magnetic Coordinate System
Sq	Solar-quiet
SR	Solar Rotation Period
SRC	Symmetric Ring Current
SV	Secular Variation
UT	Universal Time

LIST OF CONSTANTS

Sign	**Description**	**SI unit**	**Value**
$R_\odot$	Sun's volumetric mean radius	m	$6.957{\cdot}10^{8}$
ϵ_0	Permittivity of free space	Fm^{-1}	$8.854{\cdot}10^{-12}$
μ_0	Permeability of free space	Hm^{-1}	$1.257{\cdot}10^{-6}$
a	Earth's volumetric mean radius; $R_E = a$	m	$6.371{\cdot}10^{6}$
c	Speed of light, $c = (\mu_0\epsilon_0)^{-\frac{1}{2}}$	ms^{-1}	$2.998{\cdot}10^{8}$
r_{L1}	Radial distance of Earth's first Lagrangian point (geocentric)	m	$1.488{\cdot}10^{9}$
AU	Astronomical Unit	m	$1.496{\cdot}10^{11}$
eV	Electronvolt	J	$1.602{\cdot}10^{-19}$

Sign	Description	SI unit
M	Magnetic dipole moment	Am^2
N	Maximum degree of spherical harmonic expansion	1
P	Plasma pressure	Pa
V	Scalar magnetic potential	Vsm^{-1}
Φ	Magnetic flux	Wb
$\boldsymbol{B}$	Magnetic flux density	T
$\boldsymbol{E}$	Electric field	Vm^{-1}
$\boldsymbol{R_c}$	Local radius of curvature	m
$\boldsymbol{j}$	Current density	Am^{-2}
$\boldsymbol{v_B}$	Bulk (average) velocity	ms^{-1}
$\boldsymbol{v_D}$	Magnetic guiding center drift velocity	ms^{-1}
$\boldsymbol{v_E}$	E-cross-B drift velocity	ms^{-1}
$\boldsymbol{v_{G\perp}}$	Transverse bulk velocity (guiding center system)	ms^{-1}
$\boldsymbol{v}$	Instantaneous velocity	ms^{-1}
ϵ_n^m	External Gauss coefficient of degree n and order m Internal: $\imath_n^m$	T
$\hat{\boldsymbol{m}}$	Unit vector along the centered dipole axis (northward)	1
ω_C	Cyclotron frequency, $f_C = \frac{\omega_C}{2\pi}$	s^{-1}
ϕ	Longitude	rad
ρ_N	Number density	m^{-3}
ρ	Electric charge density	Cm^{-3}
τ_C	Cyclotron period	s
θ	Colatitude; latitude $\theta' = \frac{\pi}{2} - \theta$	rad
h	Geodetic height	m
m	Mass	kg
q	Electric charge	C
r_M	Radius of coronal magn. flux measurement sphere (heliocentric)	m
r_S	Radius of coronal magn. source flux sphere (heliocentric)	m
r	Radial distance	m

Part I

INTRODUCTION

1

FRAMEWORK

1.1 RESEARCH QUESTION

The framework of this cumulative thesis is set by the following overarching question:

How is the evolution of electric currents in the inner magnetosphere imprinted in the ground-level geomagnetic field over the last century, and what is its origin?

The Earth's magnetic field is felt not only at its surface but also in a part of its space environment, the magnetosphere, where it influences the dynamics of charged particles. These dynamics are externally driven by the solar wind whose highly variable interaction with the Earth's magnetic field at the magnetopause allows the transfer of energy, momentum, and solar wind plasma into the magnetosphere. Charge separation in the magnetosphere then leads to the formation of large-scale, time-varying electric current systems which induce an "external" disturbance magnetic field that superposes the Earth's "internal" magnetic field.
The foundations of our current understanding of this chain of effects were laid at the beginning of the 20th century by the analysis of measurements from sparsely distributed geomagnetic observatories. However, a comprehensive view of solar-terrestrial interactions is rendered possible only in combination with spacecraft missions dedicated to the measurement of magnetospheric and solar wind state variables (first in 1958/1959) as well as Earth's magnetic field (first in 1965).
Taking into account the insights gained from the geospace exploration missions, I tackle the research question by reanalyzing the signatures of magnetospheric currents contained in the global geomagnetic observatory record since 1900. I focus on the magnetospheric ring current because it produces the most intense, large-scale geomagnetic disturbance field at ground level. Global averages of the latter constitute the class of ring current targeting geomagnetic activity indices. To date, several such indices have been developed none of which provides a consistent, reasonable absolute level and solar cycle (SC) variability over a substantial part of the observatory era. This void causes problems regarding both continuous geomagnetic field modeling as well as multi-event studies of particularly disturbed periods, i.e., geomagnetic storms.
I take the concept of geomagnetic indices as a vantage point (publication 1) from the basis of which I explore to what extent the long-term ring current variability can be traced back to the Sun (publication 2), and what additional information on the solar wind driving is contained in the spatially varying disturbance field (publication 3).

1.2 BACKGROUND

This chapter provides the background needed to grasp the complexity of the research question. Owing to the thesis' interdisciplinary nature, combining the fields of geomagnetism and near-Earth space physics, only the most relevant facts and phenomena are highlighted with limited theoretical derivations. In compliance with the logic of the approach, the information is presented bottom-up, i.e., starting with the observatory measurements and the basics of geomagnetic field modeling (section 1.2.1), followed by the solar wind forcing of magnetospheric current dynamics (section 1.2.2). Both these topics are prerequisites for understanding the capabilities and limitations of ring current indices, which are treated last (section 1.2.3).

1.2.1 *Observation and modeling of the geomagnetic field*

The analyses conducted in the underlying publications are based on continuous vector magnetic field measurements from the global network of geomagnetic observatories since 1900. Despite their sparseness, especially in the first half of the 20th century, they reflect well the dipolar structure of the geomagnetic field. Satellite measurements contribute implicitly through the used geomagnetic field models, which rely on the improved spatial sampling offered by satellite missions to properly separate internal from external field sources.

1.2.1.1 *Measurements of the geomagnetic field*

The magnetic field vector ($\boldsymbol{B}$), measured at geomagnetic observatories, is commonly expressed in a local Cartesian coordinate system $(X, Y, Z)^T$ tied to the geodetic position (latitude θ'_d, longitude ϕ, height h) and the universal time (UT) of the measurement (Figure 1.1a). Alternatively, $\boldsymbol{B}$ is given in local cylindrical or spherical systems, including the horizontal (H) and total ($F = |\boldsymbol{B}|$) field intensities together with the angles declination (D) and inclination (I):

$$\begin{aligned} &\text{cylindrical:} \quad H = \sqrt{X^2 + Y^2}, \quad D = \tan^{-1}(Y/X), \quad Z \\ &\text{spherical:} \quad D, \qquad\qquad\qquad I = \tan^{-1}(Z/H), \quad F = \sqrt{H^2 + Z^2} \end{aligned} \tag{1.1}$$

According to the state-of-the-art measurement routine (see Matzka et al., 2010; Rasson et al., 2011), the variations of the magnetic field are continuously recorded by a digital three-component Fluxgate magnetometer, which suffers from instrument and environmental drifts caused by, e.g., temperature variations. These are calibrated regularly, e.g., once a week, using an absolute measurement of F by a proton, Overhauser, or optically pumped magnetometer in combination with a measurement of the directions (D, I) by a Fluxgate theodolite ("DI-flux").
The global network of observatories (Figure 1.2a) is coordinated by the International Association for Geomagnetism and Aeronomy (IAGA), Division V-OBS. Since 2007, the corresponding geomagnetic data master catalog (Figure 1.2b), containing digital minute and hourly mean observatory data, is held by the World Data Center for Geomagnetism, Edinburgh, which is hosted by the British Geological Survey (BGS). Before the availability of digital minute means (first in ∼1970), hourly means were

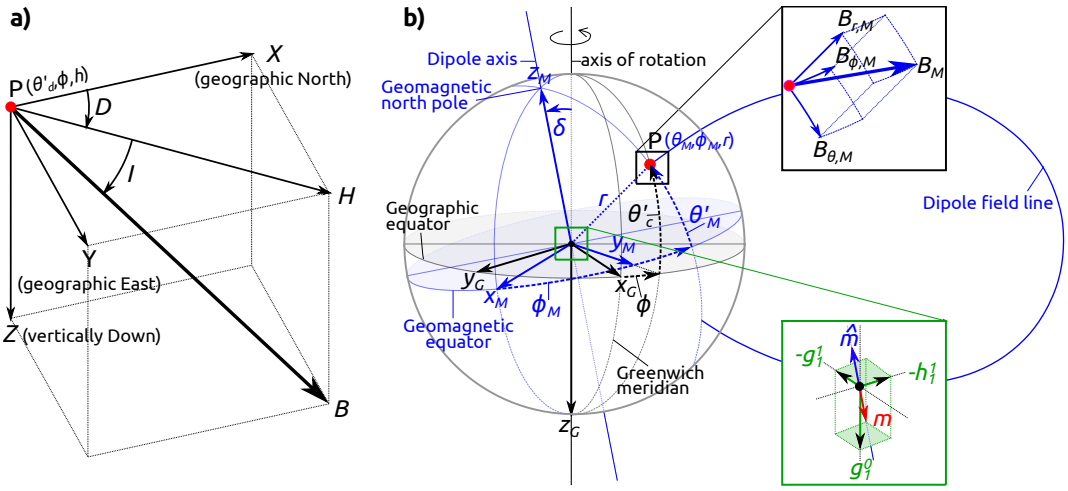

Figure 1.1: Observatory position (P) and measured magnetic field elements in **a)** the local Cartesian coordinate system; reproduced from Merrill et al., (1998), Fig. 2.1 and **b)** the geocentric (GEO, black) and geomagnetic (MAG, blue) coordinate systems; modification from Campbell, (1997), Fig. 1.8. The axis of the centered dipole field is aligned with $\hat{\boldsymbol{m}}$ (green box) and slightly tilted (angle δ) with respect to the rotation axis.

determined by visual judgment from photographic paper records and centered on half hours. The annual means, centered on half decimal years, are calculated from the hourly means and stored separately by the BGS. Consequently, whenever there are annual means, corresponding hourly means must exist. The fact that this is not reflected in Figure 1.2b) means that a considerable fraction of hourly mean values are not yet digitized from the yearbooks they were historically documented in.
The important point communicated by Figure 1.2 is the irregularity of the geomagnetic observatory record regarding both the spatial distribution (top) and the growth rate throughout time (bottom). Unfortunately, these network parameters are not guided by the changing morphology of the geomagnetic field but rather by the ability and incentive of policymakers to invest in research. As indicated by a notable upward spike in Figure 1.2b), the perhaps most effective incentive to date was the International Geophysical Year from 1957–1958. The problem of inadequate measurement coverage is largely overcome by dedicated magnetic field satellite missions at Low Earth Orbit (LEO), the most recent one being Swarm, which persistently supplement the observatory measurements since 1999 (gray bars in Figure 1.2b, section 1.2.1.2). However, due to the space-time ambiguity of satellite measurements, observatories are still needed to provide the ground truth for modern geomagnetism.

1.2.1.2 *A global field representation*

To a first approximation, Earth's magnetic field in space behaves as a magnetic dipole because the higher-order field components decline more quickly with radial distance. The concept was derived already in the early 19th century around when the observatory network came into existence. It was observed that the electric field of two point charges (q [C]) of opposite sign, separated by a short distance, is of the

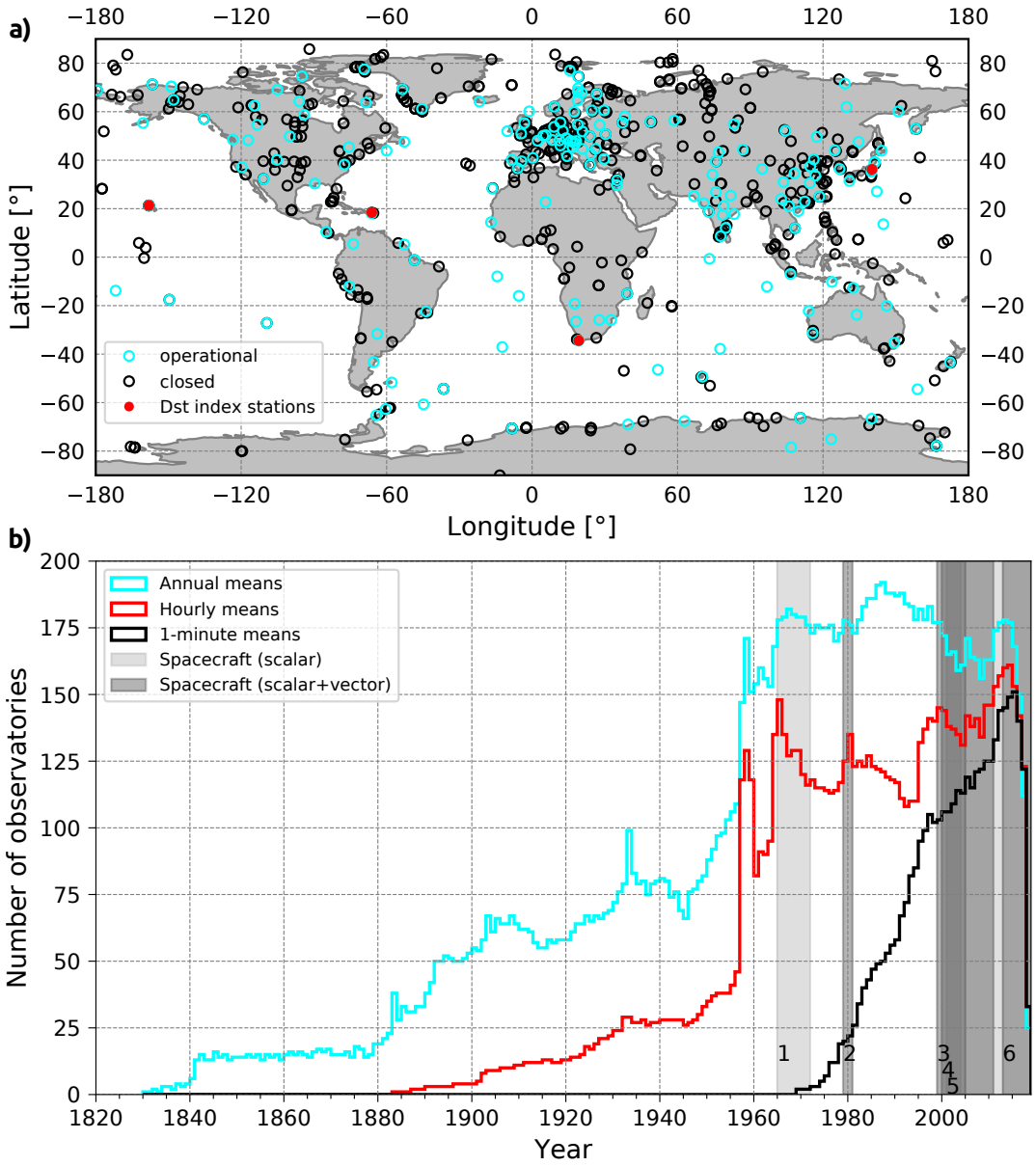

Figure 1.2: Observatory locations and available data as of September 2019. **a)** Locations of operational (cyan) and historical (black) geomagnetic observatories as given by the IAGA, Division V-OBS (`https://www.bgs.ac.uk/iaga/vobs/home.html`). The Dst index observatories (Honululu, San Juan, Hermanus, and Kakioka) are highlighted in red. **b)** Number of observatories per year delivering annual, hourly and one-minute means according to the catalogs held by the BGS (`http://www.wdc.bgs.ac.uk/`); updated version of Matzka et al., (2010), Fig. 6. Time periods highlighted in gray indicate LEO magnetic field satellite missions: (1) POGO series, (2) Magsat, (3) Ørsted (since 1999; vector data up to 2004), (4) CHAMP (2000–2010), (5) SAC-C (2001–2004), (6) Swarm.

same form as the magnetic field produced by a current (I [A]) that flows in a wire loop and encloses the area A [m^2] with $IA =: M$. M is called the magnetic dipole

moment. At the Earth's surface (radial distance $r = a = 6371$ km), the dipole field takes the form

$$\begin{aligned} B_r &= -Z = \frac{\mu_0 M \cos(\theta)}{2\pi a^3} = Z_0 \cos(\theta) \\ B_\theta &= -H = \frac{\mu_0 M \sin(\theta)}{4\pi a^3} = H_0 \sin(\theta) \\ B_\phi &= 0, \end{aligned} \tag{1.2}$$

where μ_0 is the permeability of free space.
How is M determined from the measurements of a globally distributed observatory ensemble? First, a transformation from the local Cartesian frame into the spherical geographic coordinate system (GEO) is needed (Figure 1.1b, black), which requires knowledge of the reference ellipsoid according to the World Geodetic System 1984 as well as a transformation between Cartesian and spherical coordinate systems (see Laundal and Richmond, 2017, section 2). Next, Maxwell's equations (cf. Roederer and Zhang, 2014, p. 174 in SI units) come into play,

Solenoidality of $\boldsymbol{B}$ $$\nabla \cdot \boldsymbol{B} = 0 \tag{1.3}$$

Ampère's law $$\nabla \times \boldsymbol{B} = \mu_0 \left(\boldsymbol{j} + \epsilon_0 \frac{\partial \boldsymbol{E}}{\partial t} \right) \tag{1.4}$$

Gauss's law $$\nabla \cdot \boldsymbol{E} = \frac{\rho}{\epsilon_0} \tag{1.5}$$

Maxwell-Faraday equation $$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \tag{1.6}$$

which include the electric field ($\boldsymbol{E}$), the space charge and current densities (ρ, $\boldsymbol{j}$), and the permittivity of free space (ϵ_0; $\mu_0 \epsilon_0 = c^{-2}$).
A justifiable simplification arises from the treatment of the geospace plasma as a quasi-neutral mixture of charged particles of species s, $\rho = \sum_s \rho_{N_s} q_s \simeq 0$, so that $\boldsymbol{j} = \sum_s \rho_{N_s} q_s \boldsymbol{v}_{\boldsymbol{B}s}$ is nonzero only if the bulk velocities of the plasma species, $\boldsymbol{v}_{\boldsymbol{B}s}$, differ (see section 1.2.2). Charge neutrality is consistent with the negligence of the "displacement current" in the non-relativistic limit ($c^{-2} \frac{\partial \boldsymbol{E}}{\partial t} \simeq 0$ in Equation 1.4), which implies the closure of currents ($\nabla \cdot \boldsymbol{j} = 0$).
Coming back to the determination of the Earth's dipole moment, the solenoidality of $\boldsymbol{B}$ states that the magnetic flux $\Phi = \int \boldsymbol{B} d\boldsymbol{S}$ through a *closed* surface S vanishes. The Earth's surface can be considered a current-free region ($\boldsymbol{j} = 0$) where the magnetic field is curl-free (Ampère's law) and can thus be represented as the gradient of a scalar potential, $\boldsymbol{B} = -\nabla V$. The insertion of this into Equation 1.3 yields Laplace's equation, $\nabla^2 V = 0$. One example of a potential satisfying Laplace's equation is the dipole potential, which represents the dipole magnetic field (Equation 1.2):

$$V_d = \frac{\mu_0 M \cos(\theta)}{4\pi r^2} \tag{1.7}$$

In 1838, Carl Friedrich Gauss presented a general solution of Laplace's equation in the form of an expansion into spherical harmonics (SH):

$$V(r, \theta, \phi) = \Re \left\{ a \sum_{n=1}^{\infty} \sum_{m=0}^{n} \left[\underbrace{\left(\frac{a}{r}\right)^{n+1} \iota_n^m}_{\text{internal}} + \underbrace{\left(\frac{r}{a}\right)^{n} \epsilon_n^m}_{\text{external}} \right] P_n^m \left(\cos(\theta)\right) e^{im\phi} \right\} \tag{1.8}$$

Above, P_n^m are the Schmidt quasi-normalized associated Legendre polynomials of degree n and order m (see Winch et al., 2005) and $\iota_n^m = g_n^m - ih_n^m$, $\epsilon_n^m = q_n^m - is_n^m$ are the complex internal and external Gauss coefficients. The differentiation of the contributions to V from internal and external sources is discussed in section 1.2.1.3. In practice, the expansion has to be cut at some maximum degree N. For the internal dipole field ($N = 1$) the corresponding magnetic potential in terms of SHs is:

$$V_d = \Re \left\{ \frac{a^3}{r^2} \left[\iota_1^0 \cos(\theta) + \iota_1^1 \sin(\theta) e^{i\phi} \right] \right\} \tag{1.9}$$

This expression can be simplified by changing the reference system from the geographic (subscript G) to the geomagnetic coordinate system (MAG; Laundal and Richmond, 2017, chapter 3; Figure 1.1b, blue):

$$\hat{\boldsymbol{Z}}_M = \hat{\boldsymbol{m}}, \quad \hat{\boldsymbol{Y}}_M = \frac{\hat{\boldsymbol{Z}}_G \times \hat{\boldsymbol{Z}}_M}{|\hat{\boldsymbol{Z}}_G \times \hat{\boldsymbol{Z}}_M|}, \quad \hat{\boldsymbol{X}}_M = \hat{\boldsymbol{Y}}_M \times \hat{\boldsymbol{Z}}_M \tag{1.10}$$

Convention dictates that the geomagnetic dipole axis ($\hat{\boldsymbol{Z}}_M$) points northward along $\hat{\boldsymbol{m}} = -\frac{\boldsymbol{m}}{|\boldsymbol{m}|}$, where $\boldsymbol{m} = (g_1^1, h_1^1, g_1^0)^T$ and $|\boldsymbol{m}| =: g_{1M}^0$ is the so-called reduced moment (Figure 1.1, green box). In the MAG system, the expression for the dipole potential reduces to:

$$V_d = \frac{a^3}{r^2} g_{1M}^0 \cos(\theta) \underset{\text{Eq. 1.7}}{=} \frac{\mu_0 M \cos(\theta)}{4\pi r^2} \Leftrightarrow M = \frac{4\pi}{\mu_0} a^3 g_{1M}^0 \tag{1.11}$$

According to the twelfth generation of the International Geomagnetic Reference Field (IGRF-12; Thébault et al., 2015), a fit of globally measured $\boldsymbol{B}$ to the gradient of the internal geomagnetic potential (Equation 1.8, $N = 13$) gives a dipole moment $M = 7.72 \cdot 10^{22}$ Am^2, a dipole axis tilt $\delta = 9.69°$, and the corresponding position of the northern geomagnetic pole (80.31°N, 72.63°W) for 2015. As the dipole field generally accounts for the majority of the internal field ($\sim$93 % for 2015), the $\sim$7 % decline of M since 1900 (Figure 1.3a) and the accelerating migration of the geomagnetic north pole toward the geographic pole (Figure 1.3b) comprehensively reflect the secular variation (SV), i.e., the changes of the internal magnetic field on time scales $\gtrsim$ 1 year.

1.2.1.3 *The sources of the geomagnetic field and their separation*

So far, the geomagnetic field was described as an internal dipole field (Equation 1.11). The predominant dipolar shape of the geomagnetic field is explained by a self-sustaining dynamo process, which operates in Earth's fluid outer core and produces the so-called core field (e.g., Wicht and Sanchez, 2019). While the core field accounts for $\gtrsim$ 95 % of the total field measured at Earth's surface, the permanently magnetized material in Earth's crust (e.g., Lesur et al., 2016) makes up only a few percent on average (e.g., Olsen and Stolle, 2012). Both fields together are denoted as the *internal* sources (Figure 1.4a).

The *external* sources are contributions from electric current systems in the ionosphere, particularly in the E-region (90–120 km altitude), and in the magnetosphere $\sim$2–8 Earth radii (R_E=a) away. Outside of the polar zones, the most relevant ionospheric sources are the solar-quiet (Sq) current vortexes, which are centered at

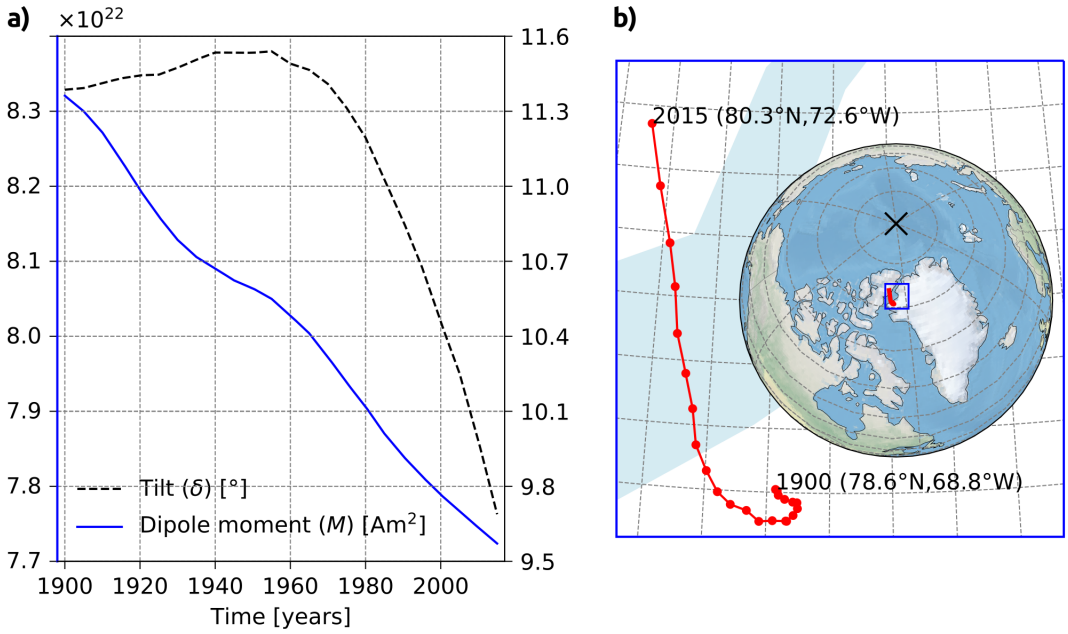

Figure 1.3: Evolution of the geocentric dipole field during 1900–2015 according to IGRF-12. **a)** Dipole moment (M, left Y-axis) and dipole axis tilt (δ, right Y-axis). **b)** Position of the geomagnetic north pole (red dots) in the Vertical Near-Side Perspective projection viewed from 200 km (outer blue box) and 2000 km (inset) above the surface. The black cross marks the geographic north pole.

$\pm 30°$ latitude on both hemispheres, and the equatorial electrojet (EEJ), which flows eastward along the dip equator ($I = 0$). Both are produced by the dayside E-region dynamo whose strength strongly depends on the prevailing solar activity level, which is typically indicated by the solar radio flux at 10.7 cm (F10.7; e.g., Yamazaki and Maute, 2017). The magnetic fields associated with these ionospheric current systems are of comparable amplitude as the crustal field. The magnetospheric currents constitute the second-largest contribution to the overall field after the core field outside of the auroral oval (Lühr et al., 2017). Closest to us are the ring currents in the inner magnetosphere, which are connected to the polar ionosphere via field-aligned currents (FACs). The net ring current encircles the Earth in a westward sense in the magnetic equatorial plane (details in sections 1.2.2.3 and 1.2.2.4).
The external fields are highly variable in time and induce currents in the conducting mantle which, in turn, give rise to a secondary magnetic contribution that is measured at the observatories together with the primary fields. An appropriate distinction between the inducing and the induced fields requires knowledge of the 3-D conductivity structure in Earth's mantle, the modeling of which is a considerable scientific challenge of its own (see chapter 7).
A mathematical separation of internal and external sources to the total magnetic field is possible based on Gauss's formulation for the geomagnetic potential (Equation 1.8). An evaluation of $-\nabla V$ at the Earth's surface gives:

$$B_r = -\frac{\partial V}{\partial r} = \Re \left\{ \sum_{n=1}^{N} \sum_{m=0}^{n} \left[(n+1)\, \iota_n^m - n \epsilon_n^m \right] P_n^m \left(\cos(\theta) \right) e^{im\phi} \right\}$$

$$B_\theta = -\frac{1}{r}\frac{\partial V}{\partial \theta} = -\Re\left\{\sum_{n=1}^{N}\sum_{m=0}^{n}\left[\iota_n^m + \epsilon_n^m\right]\frac{dP_n^m\left(\cos(\theta)\right)}{d\theta}e^{im\phi}\right\}$$
$$B_\phi = -\frac{1}{r\sin(\theta)}\frac{\partial V}{\partial \phi} = -\Re\left\{\sum_{n=1}^{N}\sum_{m=0}^{n}\left[\iota_n^m + \epsilon_n^m\right]\frac{im}{\sin(\theta)}P_n^m\left(\cos(\theta)\right)e^{im\phi}\right\} \quad (1.12)$$

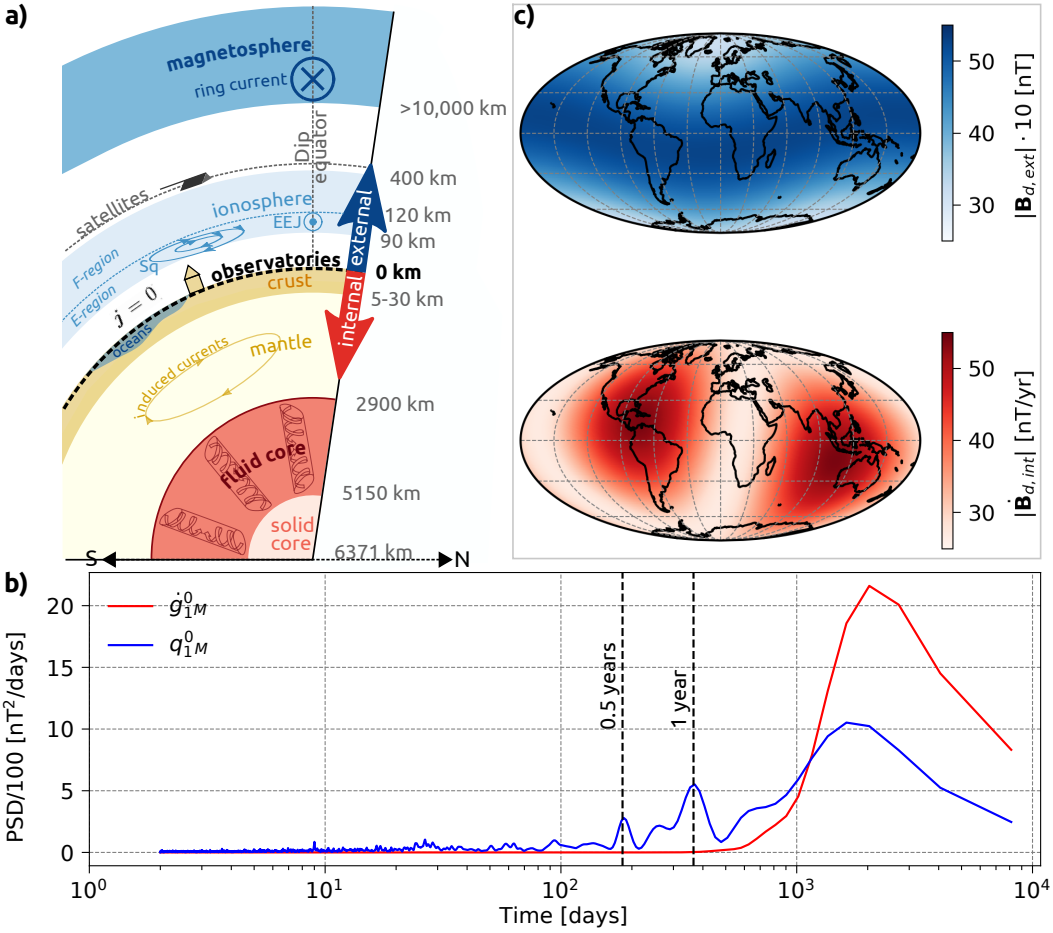

Figure 1.4: **a)** Not-to-scale sketch of internal (reddish) and external (blueish) geomagnetic field sources at low and middle latitudes with respect to Earth's surface (black dashed). Shown is a radial section of the Earth and geospace on the dayside, centered on the plane of the magnetic dip equator ($I = 0$). **b)** Power spectral densities calculated from daily-resolved time series (1997–2019) of dipole coefficients ($N = 1$), representing SV (red) and the magnetospheric field along the Z_M-axis (see Figure 1.1b). The spectra overlap for periods $\gtrsim 1$ year. **c)** Global maps of the dipole field intensity due to SV (bottom) and the magnetospheric field intensity (incl. the mantle-induced field) in the Solar Magnetic coordinate system (SM; Laundal and Richmond, 2017, section 3.5) on April 30, 2019 00:00:00 UT (top). The plots b) and c) are produced from the model CHAOS-6.x9 via "ChaosMagPy" (`http://www.spacecenter.dk/files/magnetic-models/CHAOS-6/`).

Since the horizontal field components (B_θ, B_ϕ) yield information on the sum of the internal and external Gauss coefficients ($b^+ = \iota_n^m + \epsilon_n^m$), while the radial component (B_r) yields information on their difference ($b^- = (n+1)\,\iota_n^m - n\epsilon_n^m$), the internal and external Gauss coefficients are separable:

$$\iota_n^m = \frac{nb^+ + b^-}{2n+1}, \quad \epsilon_n^m = \frac{b^+(n+1) - b^-}{2n+1} \tag{1.13}$$

In this work, the focus lies on the separation of the prime internal and external magnetic fields, i.e., the field contributions from the core and the magnetosphere, in the overlapping large-scale ($N = 1$), low-frequency ($T \gtrsim 1$ year) domain (Figure 1.4b). The strengths of these sources differ by about four orders of magnitude, but the SV is of comparable magnitude as the inner magnetospheric field (Figure 1.4c). The quality of the source separation can be enhanced by several factors: (a) *Obtaining the best possible measurement coverage.* The very uneven distribution of ground observatories (cf. Figure 1.2a) causes spatial aliasing in the source separation procedure (Olsen et al., 2010). This can be avoided by incorporating vector field measurements from the dedicated spacecraft missions (particularly Swarm, cf. Figure 1.2b). A combination of ground- and space-based measurements is used by all state-of-the-art geomagnetic field models that offer a sophisticated source separation, prominent examples of which are the CHAOS (Finlay et al., 2016) and CM (Sabaka et al., 2015) model series. However, because the discontinuous space-based measurements only reach back to 1965, the separation procedure has to rely solely on observatory data beforehand. (b) *Choosing an appropriate coordinate system.* The core field is best organized in the MAG system (cf. Figure 1.1b, blue). Due to the geometry of the magnetospheric ring current, i.e., a westward directed current loop in the magnetic equatorial plane, the corresponding magnetic field is dipolar (cf. section 1.2.1.2) and opposes the dipolar component of the core field. Consequently, the ring current field is also well-organized in MAG. (c) *Choosing an appropriate latitudinal range for the target external current system.* The ground effects of the different magnetospheric current systems are roughly sorted by geomagnetic latitude (sections 1.2.3, 6.3.1). Targeting the ring current system, the complexity of the separation can be reduced by excluding observatory measurements from high geomagnetic latitudes, at which other field sources contribute relevantly to the total field (e.g., FACs).

1.2.2 *Solar wind forcing of magnetospheric current dynamics*

The buildup and decay of magnetospheric electric currents are expressed in the level of "geomagnetic activity", a collective term for short-lived variations of Earth's magnetic field in response to observable phenomena on the Sun. The coupling agent between solar and geomagnetic activity is the solar wind, i.e., the outward expanding flow of ionized solar plasma with an embedded weak remnant of the solar magnetic field ($\boldsymbol{B}_{sw}$), the Interplanetary Magnetic Field (IMF). This section follows the cause-and-effect chain that starts with the highly variable near-Earth solar wind and ends with the particle motion in the magnetosphere and the formation of the ring current.

1.2.2.1 *The solar wind near Earth*

Our current knowledge of the near-Earth solar wind is largely based on in-situ observations of spacecraft positioned on Lissajous orbits around the L1 Lagrangian point (e.g., Wind, ACE, DISCOVR). L1 lies on the Earth-Sun line at a radial distance of ∼1.5 million km from the Earth's center ($r_{L1} \simeq$ 0.01 AU), where the gravitational pull of the Sun and the Earth equals the centripetal force required for the spacecraft to orbit the Sun in sync with the Earth.
The solar wind flow is supersonic with a mean velocity (v_{sw}) of $\sim$ 500 km/s (cf. Prölss, 2004, chapters 6.1, 6.2). It is composed predominantly ($\sim$ 96 %) of protons (H^+) and electrons (e^-) with number densities $\rho_{N_p} \simeq \rho_{N_e} \simeq \rho_N \simeq 6\ \mathrm{cm}^{-3}$ as well as a small contribution of alpha particles (He^{2+}). The large-scale flow structure in the ecliptic is of bimodal nature: Dense, slow solar wind ($v_{sw} \lesssim 400$ km/s) from coronal regions of closed magnetic field configuration (e.g., the streamer belt) alternates with tenuous high-speed streams (HSSs; $v_{sw} \gtrsim 600$ km/s) from coronal holes (e.g., Cranmer et al., 2017). This structure can be described by means of jetlines (cf. Prölss, 2004, p. 296) that connect solar wind particles from the same source region (also called pathlines, e.g., Schatten, 2013). Jetlines form an Archimedian spiral whose curvature depends on v_{sw} ("Parker spiral").
The solar wind plasma is relatively dilute, and can be considered as fully ionized with nearly infinite electrical conductivity and an extremely low collision frequency compared to the gyro- and plasma frequencies ("ideal plasma", e.g., Kivelson, 1995, pp. 40–41). In such a plasma an electric field parallel to $\boldsymbol{B}$ cannot be sustained because it is canceled by electrons that move rapidly along the field lines so that $\boldsymbol{E} \simeq \boldsymbol{E}_\perp$. Also, $\boldsymbol{E}_\perp$ vanishes *in the reference frame of the moving particles* ($\boldsymbol{E}'_\perp = 0$) due to the extremely high conductivity. Using the Lorentz transformation for the non-relativistic case ($|\boldsymbol{v_B}| \ll c$), this is expressed as

$$\begin{aligned} \boldsymbol{E}' &\simeq \boldsymbol{E} + \boldsymbol{v_B} \times \boldsymbol{B} = 0 \\ &\Leftrightarrow \boldsymbol{E} = -\boldsymbol{v_B} \times \boldsymbol{B} \underset{\text{Eq. 1.6}}{\Longrightarrow} \frac{\partial \boldsymbol{B}}{\partial t} = \nabla \times (\boldsymbol{v_B} \times \boldsymbol{B}), \end{aligned} \tag{1.14}$$

where $\boldsymbol{E}$ is the "convection" electric field as seen by an observer on the Earth and $\boldsymbol{v_B}$ is the bulk velocity of the solar wind. Equation 1.14 is termed the "frozen-in flux theorem" because it implies that the magnetic flux through a fluid surface S remains constant as S is convected with the plasma flow, i.e., that the jetlines guide the solar magnetic field lines through interplanetary space.
The average IMF intensity observed at L1 is $F_{sw} \simeq 3.5$ nT. IMF, too, displays strong, short-lived fluctuations in direction and amplitude of which particularly those along the Z-axis of the Geocentric Solar Magnetic coordinate system (GSM) determine the efficiency of the coupling to the Earth's magnetosphere (section 1.2.2.2). In GSM, the X-axis points from the Earth toward the Sun, and the Y-axis is perpendicular to $\hat{\boldsymbol{m}}$ pointing toward dusk (Laundal and Richmond, 2017). The coupling is promoted by sustained periods of southward oriented $B_{sw,z}$, which are effectively caused by two distinct, sporadic solar wind disturbances (e.g., Kilpua et al., 2017): Co-rotating or Stream Interaction Regions (CIRs or SIRs) and Interplanetary Coronal Mass Ejections (ICMEs). While the former are regions of locally compressed plasma which form at the intersection between fast and slow stream sectors (e.g., Gosling

and Pizzo, 1999), the latter are the interplanetary continuations of huge plasma eruptions from active regions of the corona (e.g., Zurbuchen and I. G. Richardson, 2006).
In addition to these disturbances, solar magnetic activity varies regularly with a period of about 27 days (one Bartels rotation), which corresponds to one synodic solar rotation period (SR). The Earth is embedded either in an IMF sector pointing away (GSM $-\boldsymbol{B}_{sw,x}$) or toward (GSM $+\boldsymbol{B}_{sw,x}$) the Sun, at which sector reversals typically occur twice or four times per SR. This is explained by the Earth's position relative to the heliospheric current sheet (HCS; e.g., Smith, 2001), which is inclined with respect to the ecliptic and displays azimuthal warps as does the skirt of a twirling ballerina (Alfvén, 1977). According to the solenoidality of $\boldsymbol{B}$ (Equation 1.3), the toward and away solar magnetic fluxes through a heliocentric sphere are balanced, so that the signed flux (of one radial polarity) should be half the value of the unsigned flux. Since the radial component of the solar magnetic field is virtually independent of latitude, the signed solar magnetic flux, Φ_{sw}, is retrievable from point observations of $B_{sw,x}$ near the Earth at $r_M \approx 1$ AU (L1), averaged over one SR (Lockwood, 2013):

$$\begin{aligned} \Phi_{sw}(r_M) &\approx 2\pi r_M^2 \langle |B_{sw,x}(r_M)| \rangle_{\mathrm{SR}} \\ \mathrm{OSF}(r_S) &= \Phi_{sw}(r_M) - E \end{aligned} \tag{1.15}$$

$\Phi_{sw}(r_M)$ can be taken as an approximation of the coronal source magnetic flux, the open solar magnetic flux (OSF), once the excess flux (E), accounting for additional flux contributions arising between the source surface ($r_S = 2.5 R_\odot$) and the measurement point, has been subtracted (Lockwood et al., 2009a,b).
On even longer time scales, the solar wind structure varies with the SC ($\sim$11 years) phase (summarized by J. D. Richardson and Kasper, 2008). At solar minimum, the solar magnetic field is dipolar, the coronal holes are concentrated at the poles, the HCS (small tilt) and the streamer belt are located near the equator, and few ICMEs occur. In contrast, the magnetic field is disordered at solar maximum with slow solar wind originating from all heliolatitudes and frequent occurrences of ICMEs ($\sim$15 % of the solar wind at the Earth).

1.2.2.2 *Solar wind-magnetosphere interaction*

As the solar wind encounters the Earth's magnetosphere, the plasma flow is abruptly slowed and heated, forming the bow shock and the magnetosheath region in front of the dayside magnetopause. Following from the frozen-in flux theorem, the plasma of the magnetosheath can not simply diffuse into that of the magnetosphere, so that the plasma populations are separated by the magnetopause current sheet (e.g., Birn et al., 1998; Speiser, 1973). However, it is precisely within such a thin current sheet that the frozen-in flux condition can break down, which enables the process of magnetic reconnection, i.e., the merging of two oppositely directed magnetic field lines (see Treumann and Baumjohann, 2013).
In 1961, J. W. Dungey explained the large-scale plasma convection observed inside the magnetosphere by two successive reconnection processes at the dayside magnetopause and the nightside magnetotail (Dungey, 1961, 1965). According to the

"Dungey convection cycle", the merging of antiparallel geomagnetic (northward GSM B_z) and IMF (southward GSM $B_{sw,z}$) field lines at the magnetopause creates two "open" field lines, each with one end attached to the Earth and the other stretching into interplanetary space (see Milan et al., 2017, Fig. 1a for an illustration of the "open magnetosphere"). The solar wind flow pulls the interplanetary parts of the field lines to the nightside at which the plasma on the flux tubes senses the dawn-dusk directed convection electric field, $\boldsymbol{E} = -\boldsymbol{v_{sw}} \times \boldsymbol{B_{sw}}$ (cf. Equation 1.14). In the geotail, the two open field lines reconnect with each other again to reform purely geomagnetic and interplanetary field lines. The cycle is completed once the geomagnetic field lines have drifted back to the dayside through the outer magnetosphere.

The rate at which energy is extracted from the solar wind by the magnetosphere is approximated by so-called coupling functions, of which a great variety has been derived over the years (e.g., Finch and Lockwood, 2007; Newell et al., 2007). They typically combine GSM $B_{sw,z}$ and v_{sw} as well as a geometry factor, $\theta_{sw} = \tan^{-1}\left(\frac{B_{sw,y}}{B_{sw,z}}\right)$, called clock angle.

The small fraction of the magnetosheath plasma that penetrates the magnetosphere via the open magnetic field lines contributes to two relevant plasma reservoirs: one just inside of the magnetopause, the magnetospheric boundary layer (plasma mantle), which, along with the polar ionosphere, feeds the other, prime reservoir, a slab-like plasma sheet in the magnetotail (e.g., Daglis et al., 1999; Welling and Ridley, 2010).

1.2.2.3 *Guiding center particle drift*

Electrons and ions respond to the Lorentz force by gyrating around $\boldsymbol{B}$ in opposite senses with cyclotron frequencies $\omega_C = qBm^{-1}$ (in SI units). The center of the gyration orbits, the *guiding center*, drifts according to the morphology of the electromagnetic field. As a direct consequence of the Lorentz transformation, electrons and ions collectively "$\boldsymbol{E} \times \boldsymbol{B}$" drift with the velocity $\boldsymbol{v_E} = (\boldsymbol{E} \times \boldsymbol{B})\, B^{-2}$ from the plasma sheet toward the Earth in the magnetic equatorial plane. At the same time, they sense a strengthening spatial inhomogeneity of the Earth's (dipole) magnetic field, which eventually causes an azimuthal drift motion. The field's gradient (∇B) and its curved field lines (with local curvature radius $\boldsymbol{R_c} = -R_c^2(\nabla B)B^{-1}$) produce gradient and centrifugal forces, respectively, in response to both of which electrons and ions move into different directions: electrons drift eastward (toward dawn) and ions drift westward (toward dusk) around the Earth as viewed from above. Adding the gradient ($\boldsymbol{v_G}$) and curvature ($\boldsymbol{v_C}$) drifts results in the total magnetic guiding center drift (cf. Baumjohann and Treumann, 1997, p. 30),

$$\boldsymbol{v_D} = \underbrace{\frac{v_\perp^2}{2\omega_C B^2}\,(\boldsymbol{B} \times \nabla B)}_{\boldsymbol{v_G}} + \underbrace{\frac{v_\parallel^2}{\omega_C B R_c^2}\,(\boldsymbol{R_c} \times \boldsymbol{B})}_{\boldsymbol{v_C}} = \left(\frac{1}{2}v_\perp^2 + {v_\parallel}^2\right)\frac{\boldsymbol{B} \times \nabla B}{\omega_C B^2}, \quad (1.16)$$

which is the perpendicular component of a particle's instantaneous velocity, averaged over one cyclotron turn, $\boldsymbol{v_D} = \langle \boldsymbol{v}_\perp \rangle_{\tau_C}$. As such, it plays an important role in the formation of the ring current.

1.2.2.4 *Magnetospheric current systems with a spotlight on the ring current*

Three large scale current systems flow in the magnetosphere: the system of Chapman-Ferraro magnetopause currents and cross-tail currents in the outer magnetosphere as well as the ring current system in the inner magnetosphere (illustrated in Milan et al., 2017, Fig. 1b,c). They are defined by their associated transverse current densities ($\boldsymbol{j}_\perp$, Equation 1.4), which are related to the *collective particle motion* and consist of two relevant contributions. The first one, $\boldsymbol{j_D}$, results from the spatially averaged guiding center drift, $\boldsymbol{v_{G\perp}} = \langle \boldsymbol{v_D} \rangle$, see Equation 1.16. The second one is the "magnetization current", $\boldsymbol{j_M}$, which accounts for the collective diamagnetic effect of the cyclotron motion in form of a plasma magnetization, $\boldsymbol{M} = -P_\perp \boldsymbol{B}/B^2$, where $P_\perp = \frac{1}{2}\rho_N m \left[\langle v_\perp^2 \rangle_{\tau_C} - {v_D}^2\right]$ is the plasma pressure perpendicular to the magnetic field (cf. Roederer and Zhang, 2014, chapters 5.2, 5.3). Adding $\boldsymbol{j_D}$ and $\boldsymbol{j_{M\perp}}$ under the assumed simplifying conditions (force-balanced state, no time dependence on the time scale of interest, no inertial terms) gives the perpendicular current density as derived by Parker, (1957), see also Ganushkina et al., (2018):

$$\boldsymbol{j}_\perp = \underbrace{\rho \cdot \boldsymbol{v_{G\perp}}}_{\boldsymbol{j_D}} + \underbrace{\nabla \times \boldsymbol{M}|_\perp}_{\boldsymbol{j_M}} = \frac{\boldsymbol{B}}{B^2} \times \left[\nabla P_\perp + \left(P_\parallel - P_\perp\right) \frac{(\boldsymbol{B} \cdot \nabla)\, \boldsymbol{B}}{B^2}\right] \tag{1.17}$$

The two terms including $P_\perp$ result from the magnetization current, while the remaining one ($\propto P_\parallel = \rho_N m \langle v_\parallel^2 \rangle_{\tau_C}$) represents curvature drift ($\boldsymbol{v_C}$) that only contributes in the case of anisotropic plasma pressure ($P_\parallel \neq P_\perp$). In this representation, the gradient drift ($\boldsymbol{v_G}$) has dropped out entirely.

According to Equation 1.17, current density is retrievable from plasma pressure measurements. In the inner magnetosphere the plasma pressure is mostly due to ions (protons) up to a few hundreds of keV (e.g., Keika et al., 2018; Williams, 1981). Radial plasma pressure profiles in the equatorial plane show a maximum around 3 R_E (e.g., Lui et al., 1987) which gives rise to a westward flowing symmetric ring current (SRC) at the outer edge and an eastward flowing SRC at the inner edge. Because of the decay of the magnetic field with radial distance, the westward current is stronger than the eastward one, so that the net SRC is westward.

The injection of the plasma from the plasma sheet into the ring current causes azimuthal plasma pressure gradients, which give rise to radially directed currents (cf. Equation 1.17). As these currents cannot close in the magnetosphere, they do so in the ionosphere via "Region-2" FACs (e.g., Vasyliunas, 1970). The part of the current loop that coincides with the symmetric ring current is termed the partial ring current (PRC). Observational evidence of the PRC does not only come from plasma pressure distributions (see also Lui, 2003) but also from estimations of the curl of the measured magnetic field disturbance (Equation 1.4; Le et al., 2004) and the observation of Energetic Neutral Atoms (ENAs; e.g., Perez et al., 2015). While instructive schematics of the whole ring current system are presented in Liemohn et al., (2015, Fig. 1a,b) and Ganushkina et al., (2018, Fig. 5, 6), the historical cause of events regarding ring current research is summed up in the "ring current biography" by Egeland and Burke, (2012).

The specific importance of the ring current is that it decreases the strength of the predominantly northward directed magnetic field observed at low and middle

geomagnetic latitudes on the Earth's surface. This phenomenon is quantified by the Dessler-Parker-Sckopke (DPS) relation (Dessler and Parker, 1959; Sckopke, 1966),

$$\frac{\Delta H(0)}{H_0} = -\frac{2}{3}\frac{W_p}{W_m} \;\Leftrightarrow\; \Delta H(0)\ [\text{nT}] \simeq -4.15 \cdot 10^{-30} W_p\ [\text{keV}], \tag{1.18}$$

which states that the horizontal intensity of the disturbance field at the Earth's center, $\Delta H(0)$, is related to the total energy of any steady distribution of trapped particles in the magnetosphere (W_p). H_0 is the horizontal, equatorial surface strength of the Earth's dipole field (Equation 1.2) and $W_m = \frac{4\pi}{3\mu_0} H_0^2 a^3$ is the total energy of the Earth's dipole field beyond the Earth's surface.

1.2.3 *Geomagnetic indices for ring current activity*

Geomagnetic indices quantify the level of geomagnetic activity measured at geomagnetic observatories. Depending primarily on the geomagnetic latitudes of the underlying observations, their correlations with the solar wind-magnetosphere coupling functions (section 1.2.2.2) vary (e.g., Finch et al., 2008) along with their sensitivities to particular parts of the coupled system of magnetospheric currents. The DPS relation (Equation 1.18) provides the physical basis upon which H-disturbances, observed at low- and mid-latitude geomagnetic observatories, are interpreted in terms of the total energy of ring current particles. An approximation to $\Delta H(0)$ is given by the Disturbed storm time index (Dst; Sugiura, 1964; Sugiura and Kamei, 1991) which is based on hourly-averaged ΔH estimations from four observatories at geomagnetic latitudes $\theta'_{M,i}$, spaced roughly equally in local time (LT), cf. Figure 1.2a:

$$\text{Dst}(t) = \sum_{i=1}^{4} \underbrace{[H_i(t) - H_{i\text{base}}(t) - H_{i\text{Sq}}(t)]}_{\Delta H_i} \cdot \left(\sum_{i=1}^{4} \cos\left(\theta'_{M,i}\right) \right)^{-1} \tag{1.19}$$

$H_i(t)$ are the horizontal field intensities measured at the four observatories at UT hour t, and $H_{i\text{base}}(t)$ are the corresponding baselines, i.e., the core field contributions. They are approximated by quadratic fits to annual means of H from the five geomagnetically "quietest" days per month (Q-days; Johnston, 1943) of the current and the four preceding years. Q-days are days at which the level of global geomagnetic activity is lowest according to the Kp index (`https://www.gfz-potsdam.de/en/kp-index/`).

Apart from the baselines, the Sq daily variation, $H_{i\text{Sq}}(t)$, is subtracted from the observatory measurements. It is approximated for each observatory and each year by averaging H of those five local days per month (m) that have the largest overlap with the five Q-days. Before averaging, the linear trends are calculated from the hours just before and after the identified days and removed. The resulting twelve monthly-averaged Sq time series, each spanning 24 hours in LT (t'), are expanded in a double Fourier series $f_i(t', m)$, from which $H_{i\text{Sq}}(t)$ is determined for each UT hour of that year.

Dst has become the standard measure of geomagnetic storms (e.g., Borovsky and Shprits, 2017). Over time it has been extended to 1932 (Karinen and Mursula, 2005)

and revised, particularly concerning the so-called "non-storm component" (e.g., Mursula and Karinen, 2005) and the Sq-field elimination procedure (Gannon and Love, 2011; Love and Gannon, 2009), see Table 1.1. Despite its frequent use in space weather applications (e.g., Rastätter et al., 2013), Dst's original derivation scheme causes problems which restrain its use, especially with regard to long-term studies:

1. *Ignorance of quiet-time currents*
   Dst's baseline is defined such that the index averages to zero on the Q-days of a particular year. This is physically incorrect because the current systems are *always* flowing. Consequently, Dst underestimates the total energy of the trapped particles by an amount that varies with the SC.

2. *Underestimation of solar-cycle variability*
   Because Q-days of solar minimum years are systematically quieter than those of solar maximum years (e.g., Vennerstroem, 2000), the calculation of the baseline from Q-days results in an underestimation of the true SC variation in Dst.

Both problems are solved in the framework of customary geomagnetic field models. Firstly, the utilization of vector magnetic field measurements from dedicated spacecraft missions (see Figure 1.2b) allows a proper internal-external field separation during geomagnetically quiet conditions (see section 1.2.1.3). Secondly, the quiet conditions are defined consistently over time, e.g., by setting a fixed threshold for an independent geomagnetic activity index like Kp. To these effects, a 25 nT quiet-time ring current level was first estimated by Langel et al., (1980) using Magsat observations from two quiet days in November 1979. This was later refined by Lühr and Maus, (2010), who determined a quiet-time range of 0–15 nT in dependence of the phase of SC 23 (2000–2009), based on measurements from CHAMP and Ørsted. The iterative inversion schemes of up-to-date geomagnetic field model series thus allow a co-estimation of improved Dst-like indices with respect to problems 1 and 2. The index baseline results from the internal-external field separation, which, in turn, is performed for quiet conditions defined according to the very same index, usually in combination with Kp and a variant of a coupling function (e.g., E_m from Newell et al., 2007). Examples of such indices are VMD (Thomson and Lesur, 2007), SVMD (Kunagu et al., 2013), MMA (Hamilton, 2013), and RC (Olsen et al., 2014), summarized in Table 1.1. However, for the ground-based indices (VMD, RC), the static, small-scale crustal field contribution which is not resolved by the corresponding model presents a remaining problem. It locally influences the observatory measurements by varying amounts, the "crustal biases", whose improper consideration in the baselines results in indices which are offset from the true activity level (e.g., Lühr et al., 2017, for RC).
A fundamentally different approach is to model Dst based on measurements of the solar wind plasma and IMF, which was first systematically done by Burton et al., (1975). Their estimation of a 20 nT constant corresponding to the average quiet-day currents is in compliance with the early satellite-based magnetic field models and was also confirmed by more sophisticated solar wind-based models of Dst (Temerin and Li, 2006, 2015). In the latter study spanning 1995–2009, the authors demonstrate that Dst $= 0$ corresponds to a disturbance of -15 nT during solar

minimum and between -20 to -30 nT during solar maximum (problem 1), and that Dst underestimates the SC variability by ~ 13 nT during this time (problem 2). In addition to these specific shortcomings of Dst, the limitations of the general concept of geomagnetic indices cause the following problems:

3. *Ignorance of spatial asymmetry*
   By definition, geomagnetic indices cannot resolve any spatial asymmetry, so that Dst does not capture the longitudinal disturbance field asymmetry associated with the PRC.

4. *Non-specificity regarding the source region*
   Being calculated from ground-level magnetic field measurements, Dst is bound to be affected not only by the SRC but also by other large-scale magnetospheric current systems.

Problem 3 was addressed by Iyemori, (1990) who split the measured mid-latitude disturbance field into a longitudinally symmetric (SYM-H) and a longitudinally asymmetric part (ASY-H), see also Iyemori et al., (2010). SYM-H is very similar to Dst, although one-minute resolved $B_{\theta,M}$-components in the MAG system from six out of 11 observatories are used, and the subtraction of the baseline and the Sq variation is done slightly differently. Once SYM-H is subtracted from the disturbances $\Delta B_{\theta,M_i}$ under consideration of a latitudinal correction at each observatory, the range of the residuals gives ASY-H.
Taking advantage of the increased availability of one-minute data (Figure 1.2b) and the current knowledge of the PRC's geometry, Newell and Gjerloev, (2012) derived the "SuperMag-based partial ring current indices" from 98 observatories, i.e., one index for each six-hour magnetic local time sector (MLT; Laundal and Richmond, 2017, Eq. 93): SMR-00, SMR-06, SMR-12, and SMR-18. These indices show quantitatively that the disturbance attributed to the PRC is largest at dusk (18 MLT), followed by noon (12 MLT) and midnight (0 MLT), and lowest at dawn (6 MLT; roughly half as disturbed as dusk).
During geomagnetically quiet periods, the plasma pressure distribution is largely symmetric, so that variations in Dst-like indices can be ascribed to the SRC. However, it is during geomagnetic storms that problem 4 presents a significant complication to the interpretation of the index' variability in terms of individual current systems. While the main contributors to the storm-time Dst have been identified with broad consensus (SRC, PRC, tail and magnetopause currents), their relative intensity as a function of the storm phase is still debated (e.g., Asikainen et al., 2010; Dubyagin et al., 2014). Ganushkina et al., (2018) summarize the results from the corresponding studies, based on both empirical as well as first-principles drift physics models (e.g., Liemohn et al., 2015): While the asymmetric current systems dominate the storm main phase (first tail current, then PRC), the symmetric ones dominate the later recovery phase (eastward and westward SRC).

Against this background, I tackle problems 1–3 in the publications presented in the following chapters. Problem 4 remains an open question in solar-terrestrial physics.

Table 1.1: Geomagnetic index family for symmetric and partial ring current activity, sorted according to the year they are first available. The upper (lower) table part lists indices which are developed in the space weather (geomagnetic field modeling) community. Vector quantities are denoted by bold-faced characters. The asterisks indicate uncertain limits of the time series.

Index	Time span	Cadence	Purpose	Reference(s)
Dcx	since 1932	1 hour	**C**orrected, extended Dst	Karinen and Mursula, (2005), Mursula and Karinen, (2005)
Dst	since 1957	1 hour	Storm detection	Sugiura, (1964), Sugiura and Kamei, (1991)
$D_{st}^{5807\text{-}4SH}$	since 1958	1 hour	Improved removal of Sq-field	Love and Gannon, (2009)
SMR	since 1980	1 min	MLT-sector-resolved Dst	Newell and Gjerloev, (2012)
SYM/ASY	since 1981	1 min	Sizes of sym[asym]-metric disturbances	Iyemori, (1990)
$D_{st}^{8507\text{-}4SM}$ (USGS-Dst)	since 1985	1 min	1-min version of $D_{st}^{5807\text{-}4SH}$	Gannon and Love, (2011)
RC	since 1997	1 hour	Stable baseline	Olsen et al., (2014)
VMD	2001–2005*	20 min	Stable baseline	Thomson and Lesur, (2007)
SVMD	2001–2010*	∼90 min (orbit av.)	Adaptation of VMD to CHAMP data	Kunagu et al., (2013)
MMA	since 2014	∼90 min (orbit av.)	Dedicated to Swarm	Hamilton, (2013)

2 OVERVIEW OF PUBLICATIONS

PUBLICATION 1: Pick, L. and M. Korte (2017). "An annual proxy for the geomagnetic signal of magnetospheric currents on Earth based on observatory data from 1900–2010." In: *Geophysical Journal International* 211.2, pp. 1223–1236. DOI: 10.1093/gji/ggx367

PUBLICATION 2: Pick, L., Korte, M., Thomas, Y., Krivova, N., and C.-J. Wu (2019). "Evolution of large-scale magnetic fields from near-Earth space during the last 11 solar cycles." In: *Journal of Geophysical Research: Space Physics* 124.4, pp. 2527–2540. DOI: 10.1029/2018JA026185

PUBLICATION 3: Pick, L., Effenberger, F., Zhelavskaya, Z., and M. Korte (2019). "A statistical classifier for historical geomagnetic storm drivers derived solely from ground-based magnetic field measurements." In: *Earth and Space Science* 6.10, pp. 2000–2015. DOI: 10.1029/2019EA000726

2.1 PUBLICATION 1

In the first study, the basic strategy for addressing the first part of the research question is laid. The Annual Magnetospheric Currents index (AMC) is derived, which can be seen as an annually resolved version of Dst, extended to 1900 and improved with regards to problems 1 and 2.
I conceptualized the study in close collaboration with my supervisor and coauthor, Monika Korte, who provided assistance regarding the utilization and comparison of the already existing geomagnetic field models and the interpretation of the results. I wrote the programming code concerning the derivation of AMC, performed the calculations, produced the figures, and wrote the manuscript.

2.2 PUBLICATION 2

In the subsequent study, the AMC derivation scheme from publication 1 is refined so that an index with hourly resolution can be calculated, the Hourly Magnetospheric Currents index (HMC), on the basis of which the second part of the research question is addressed.
The longer list of coauthors reflects the increased complexity of the task. I lead the conceptualization of the study, at which I was supported by Monika Korte, who provided continuous feedback throughout the whole working process. Yannik Thomas, a student assistant, helped to get an overview of the digitally available, heterogeneous record of hourly observatory measurements. He manually checked and revised the geomagnetic field time series under my supervision. Upon my request, Natalie Krivova proposed the comparison of HMC with the OSF and consulted on its acquisition and interpretation. Her Ph.D. student at the time, Chi-Ju

Wu, provided an extension of the F10.7 index to the past, which I used to estimate the Sq-field.
As before, I wrote the programming code concerning the derivation of HMC, performed the calculations, produced the figures, and wrote the manuscript. Additionally, I handled the publication of the HMC index and the underlying geomagnetic data set in compliance with the Enabling FAIR Data Project guidelines (`http://www.copdess.org/enabling-fair-data-project/`):

Pick, L. and M. Korte (2018). "HMC index." In: *GFZ Data Services*. DOI: 10.5880/GFZ.2.3.2018.006

2.3 PUBLICATION 3

In the last study, HMC and the underlying observatory data set from publication 2 are used to address the second part of the research question, this time with a focus on problem 3. Based on machine-learning techniques, it is analyzed which features of the storm-time asymmetric disturbance field are indicative of the solar wind driver type in a statistical sense.
I lead the conceptualization of the study, at which I was supported by all three coauthors. I consulted Frederic Effenberger regarding the existing catalogs of already classified solar wind structures, and in general matters of solar and space physics. Irina Zhelavskaya contributed with her expertise in the field of machine learning.
Again, I wrote the programming code in form of a Jupyter notebook, performed the calculations, produced the figures, and wrote the manuscript. As before, I also handled the publication of the software and the corresponding input data set:

Pick, L. (2019). "ClassifyStorms - an automated classifier for geomagnetic storm drivers based on machine learning techniques. V. 1.0.1." In: *GFZ Data Services*. DOI: 10.5880/GFZ.2.3.2019.003

Part II

PUBLICATIONS

3

AN ANNUAL PROXY FOR THE GEOMAGNETIC SIGNAL OF MAGNETOSPHERIC CURRENTS ON EARTH BASED ON OBSERVATORY DATA FROM 1900– 2010

Leonie Pick[1,2] and Monika Korte[1]

[1] Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany
[2] University of Potsdam, Am Neuen Palais 10, House 9, 14469 Potsdam, Germany

 doi: 10.1093/gji/ggx367

summary We introduce the Annual Magnetospheric Currents index as long-term proxy for the geomagnetic signal of magnetospheric currents on Earth valid within the time span 1900–2010. Similar to the widely used disturbance storm time and "Ring Current" indices, it is based on geomagnetic observatory data, but provides a realistic absolute level and uncertainty estimates. Crucial aspects to this end are the revision of observatory crustal biases as well as the implementation of a Bayesian inversion accounting for errors in the main field estimate, both required for the index derivation. The observatory choice is based on a minimization of index variance during a reference period spanning 1960–2010. The new index is capable of correcting observatory time series from large-scale external signals in a user-friendly manner. At present the index is only available as annual mean values. An extension to hourly values for the same time span is in progress.

3.1 INTRODUCTION

The magnetic field of the Earth is highly irregular in space and time with relatively slowly varying internal sources in Earth's outer core, nearly static contributions from the upper lithosphere and more rapidly changing extraterrestrial sources. Over 90 % of this field is explained by a tilted axial dipole, that is produced by a self-sustaining geodynamo in the fluid core and changes on timescales of several months and beyond (secular variation, SV). This main field is overlain by the temporally stable crustal field from permanently magnetized rocks and comparably weak but highly variable induced fields in the electrically conducting parts of the crust and upper mantle. These originate as secondary contributions from the rapidly changing magnetic fields generated by electrical current systems in the ionosphere and magnetosphere.

Strategies to separate internal from external sources have been investigated since 1839 when Carl-Friedrich Gauss applied spherical harmonic analysis (SHA) to his observations of the magnetic vector field (details in Olsen et al., 2010). Since then the measurement network has steadily grown, with some observatories having recorded SV for up to $\sim$175 years. In order to analyze the long-term time series with regard to core flow dynamics, all external contributions need to be eliminated from the measurements. The main influence that does not average out on SV timescales is the near-Earth signature of the large-scale magnetospheric field (e.g. Korte and Lesur, 2012). This is commonly removed by fitting of a core field model rather than by directly cleaning the data as we propose here.

There are three current systems that regularly contribute to the magnetospheric signal recorded in observatory data. These are the ring current in the inner magnetosphere as well as the magnetopause current (Chapman-Ferraro current) and tail currents in the outer magnetosphere (details in e.g., McPherron, 1995; Wolf, 1995). They collectively produce a magnetic field at Earth that is globally southward directed and aligned with the dipole axis of the main field (Z_M axis, see appendix A for details on coordinate systems). In particular it opposes the horizontal component of the core field at low to mid geomagnetic latitudes, especially when geomagnetic activity is high. During geomagnetic storms the disturbance field is dominated by an intensified ring current circling Earth in the equatorial plane at radial distances of 2–7 Earth radii (e.g. Daglis et al., 1999).

The Dst index (Sugiura, 1964) combines the degree of main field depression measured at four mid-latitude observatories (HON, SJG, HER, KAK) to characterize the global strength of the magnetospheric disturbance field. To enhance its use for geomagnetic field modeling, it was broken down into a direct external (Est) and an indirect induced (Ist) signal (Maus and Weidelt, 2004; Olsen et al., 2005). Dst was reconstructed for the years 1932–2002 and corrected for semiannual and seasonal variations, a normalization according to observatory latitude and an unequal weighting of the stations in a series of publications between 2005 and 2011 (see Mursula et al., 2011). The outcome of these changes is called Dcx index (Mursula and Karinen, 2005) and is publicly available at the Dcx index server (see appendix B for external data references). Furthermore the removal of the solar-quiet variation (Sq) was updated and the temporal resolution of the Dst increased to one minute

(Gannon and Love, 2011).
Despite all these improvements Dst still suffers from baseline instabilities (Olsen et al., 2005) that limit its adequacy for long-term studies. Therefore a new index called RC (appendix B) was proposed as part of the CHAOS-4 geomagnetic field model (Olsen et al., 2014). It was originally derived from observatory hourly means (OHMs) at 21 globally distributed locations at low and mid geomagnetic latitudes ($-41^\circ \leq \theta'_M \leq 54^\circ$), excluding regions close to the geomagnetic equator that are affected by the equatorial electrojet. The OHMs were stripped of internal field contributions by subtracting a previous version of the CHAOS core field model. Secondly the resulting OHMs were averaged over quiet times ($\mathrm{Kp} \leq 2^0$, $|\mathrm{dDst/dt}| \leq 2$ nT/h) to approximate and subsequently remove lithospheric offsets at each site. The observatory residuals were then transformed into the spherical geomagnetic coordinate system (*mag*). At each universal time hour a SHA was performed on the residual southward components ($B_{\theta,M}$) from night-side observatories (local time hours between 18 and 6). The RC index was then defined as the negative central external dipole term ($-q_1^0$), with the minus sign indicating a southward directed field at the dipole equator. It can be split into a direct (RCE) and an induced (RCI) part following the same strategy as for the Dst index (see above).
According to a recent assessment by Lühr et al., (2017), RC is much more consistent with direct observations of the ring current effect made by the CHAMP satellite than Dst is. Thus it has replaced Dst in up-to-date geomagnetic field models such as the Swarm Initial Field Model (Olsen et al., 2015) as an hourly proxy for the magnetospheric field intensity. However, none of these indices assesses the absolute baseline of the quite-time magnetospheric field. Based on Magsat vector data the quiet-time magnetospheric field level was first estimated as 20 nT by Langel and Estes, (1985a) and Langel and Estes, (1985b). This first approximation was then specified using nine years of CHAMP data that revealed a stable quiet-time field of 13 nT with an additional variable part of up to 15 nT depending on the solar cycle phase (Lühr and Maus, 2010). In the RC index derivation the crustal bias estimates include the average quiet-time magnetospheric field, that is consequently erroneously subtracted, causing a shift of RC to relatively less negative values.
In this study we propose a new index called AMC (*A*nnual *M*agnetospheric *C*urrents), that describes the large-scale magnetospheric field disturbances for 1900–2010 with improved absolute level and uncertainty estimates taking into account uncertainties in the used core field model. A simple extension of the RC index scheme to the past is not possible because many of the observatories used in the RC index derivation did not exist and the CHAOS core field model series is not valid before 1997.
We first give an overview of the involved observatory data (section 3.2.1) and the geomagnetic field models (section 3.2.2). We then explain the estimation of observatory crustal biases (section 3.3.1) that are needed to form observatory residuals (section 3.3.2) as the index basis. The actual derivation of the index is detailed in section 3.3.3, followed by the presentation of results that are partitioned in a reference period 1960–2010 (section 3.4.1) and the long-term period 1900–2010 (section 3.4.2). The subsequent discussion covers the index' absolute level (section 3.5.1), our reasoning regarding data choice (section 3.5.2) and an example application (section

3.5.3), before ending with conclusions (section 3.6). The appendices A–D provide information on the coordinate systems, the data sources and the geomagnetic observatories as well as two extra figures. The AMC index is available in the online supplement.

3.2 DATA AND MODELS

3.2.1 *Data*

We base our study on observatory annual means (OAMs). They are available from the World Data Centre for Geomagnetism Edinburgh, hosted by the British Geological Survey (appendix B), that also provides information on the corresponding locations, possible relocations and general measurement issues. Working with annual means has the main advantage that short-lived ionospheric disturbances, primarily the Sq daily variation, can be neglected. Although the Sq signal does not cancel in an annual average, its amplitude stays below 2 nT at a typical mid latitude observatory as Niemegk according to a study spanning the years 1960–2001 by Verbanac et al., (2007).
To identify a suitable database for our new AMC index, we consider the 21 observatories selected for the original RC index (Fig. 3.1). Due to the better data quality in the second half of the 20th century and the fact that some of the observatories were established not until the International Geophysical Year 1957–1958, we initially restrict our analysis to 16 candidates (details in appendix C) covering the period 1960–2010 without gaps greater than 10 years (see also Fig. 3.5). Based on a sensitivity analysis (section 3.5.2) we identified a set of nine observatories as ideal choice for this period ("reference set"). Finally, five observatories from the reference set offer time series reaching back to the early 1900s and form the "final set" for the AMC index (section 3.4.2).

3.2.2 *Models*

Our requirements for modeled field contributions cannot be fulfilled by one model alone: First, we need a long-term model of the main field that preferably covers the complete observatory era ($\sim$1840–present) to eliminate the core field influence from the OAMs. Secondly, we need a recent global field model including satellite data to estimate the large-scale external field at each observatory location. This is essential in order to ensure that our proxies for the high-degree crustal field contribution are free from external influences. Below, we briefly justify our decision in favor of the COV-OBS and GRIMM-3.2 field models (sources in appendix B).

3.2.2.1 *Long-term core field model*

COV-OBS (Gillet et al., 2013) spans the period 1840–2010 and thereby is the only available model covering the current century and all years back to the beginning of observatory records. The ground-based data is essentially that used for the *gufm1* field model (Jackson et al., 2000), which is supplemented by satellite data from

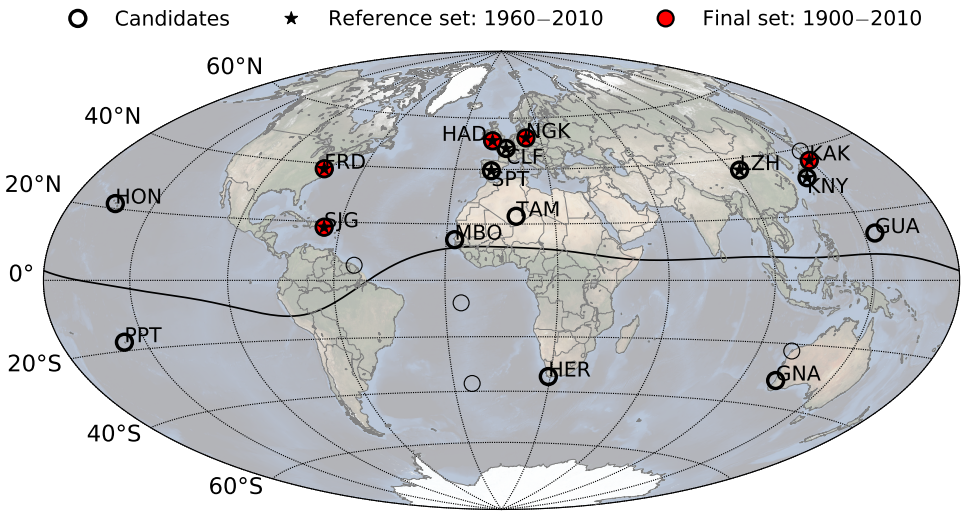

Figure 3.1: Locations of the 21 observatories used to derive the original RC index in relation to the geomagnetic equator (black line) from IGRF-12 for epoch 2015 (Thébault et al., 2015). 16 observatories are candidates for the AMC index (bold style, IAGA codes) and nine are found to be ideal during the reference period 1960–2010 (stars). Five observatories are finally chosen for the AMC index (red fillings).

the missions marked in Fig. 3.2. It is a "stochastic" field model that assumes a Gaussian prior probability distribution for the spherical harmonic (SH) coefficients of the geomagnetic potential that is expanded up to degree and order (d/o) 14. The dependence of the coefficients on time is parameterized by cubic B-splines with a knot spacing of two years. COV-OBS calculates an ensemble solution of up to 100 members whose statistics equip the user with uncertainties, that are not provided by other core field models.

To demonstrate the usefulness of these uncertainties we compare COV-OBS to model CM4 (Sabaka et al., 2004), spanning 1960–2002, as a representative of other existing decadal field models. Following a "comprehensive approach" CM4 estimates all internal and external sources of the geomagnetic field in a joint inversion of OHMs and satellite data. The difference in the northward core field component, Xc, from these two models is shown in Fig. 3.2 for four widely separated observatory locations from the reference set. While mean model differences (red) stay within $\sim$10 nT at European and East-Asian locations, e.g. NGK, they climb to $\sim$20 nT at other observatories onshore, e.g. GNA, and finally reach up to $\sim$30 nT at remote island-based stations, e.g. HON and GUA. COV-OBS' confidence interval is given by three standard deviations, $\pm 3\sigma$, from the mean (gray band). It is reduced noticeably by the inclusion of vector field satellite data from Magsat and missions in the 2000s (orange bars 2 & 4). At most years (green horizontal bars), half the confidence interval is wider than the absolute mean model difference. This means that COV-OBS includes CM4 within its uncertainty, which seems reasonable overall.

3.2.2.2 *Large-scale external field model*

Seeking the most detailed description of external fields, the POMME (Maus et al., 2010, version six) model series would be a reasonable choice. It incorporates a

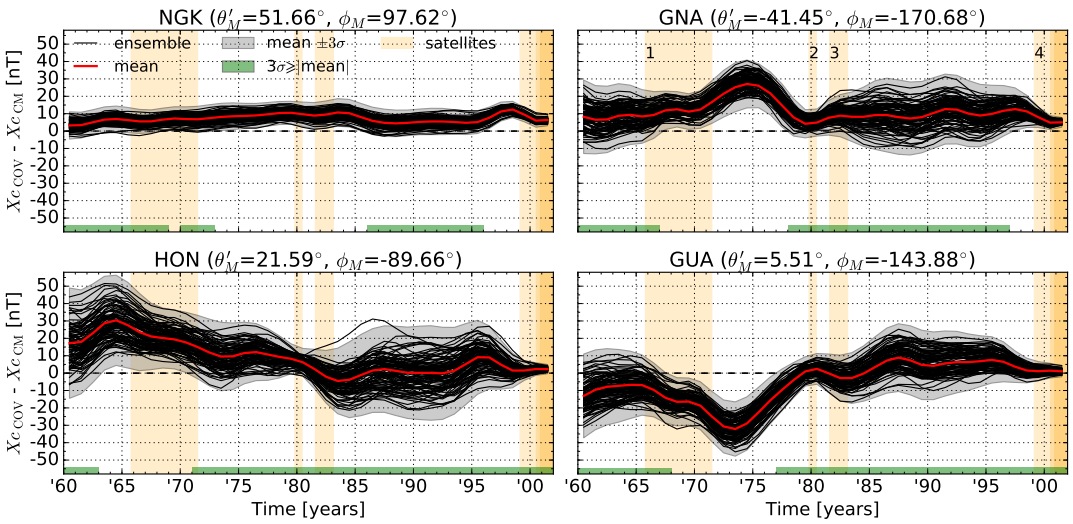

Figure 3.2: Yearly averaged differences between Xc components (core field, north, geodetic ned_d) of COV-OBS and CM4 up to d/o 14 calculated for 100 COV-OBS realizations (black) and their mean (red). The grey area marks the 99.73 % confidence interval. Green bars indicate years when COV-OBS' uncertainty envelope includes CM4. Satellite missions are 1: POGO series 2: Magsat, 3: DE-2, 4: Ørsted, CHAMP & SAC-C successively. Observatory locations are given in geomagnetic coordinates (mag).

dedicated magnetospheric field model (Lühr and Maus, 2010) that requires five solar-terrestrial system parameters as input. Among them is the F10.7 index measuring solar radio flux per unit frequency at 10.7 cm wavelength. It complements Est in parameterizing the solar cycle dependence of the ring current. We use model version eight for testing, but any version from six to 11 could have been chosen as they all use the same magnetospheric field model.

The GRIMM-3.2 field model covers 2001–2010 and does not aim at describing distinct external field sources. Instead the whole range of vector CHAMP data at all local times is used to separate the rapidly varying external field from the internal field. The external field parameterization closely follows that of the IGRF-12 candidate by Lesur et al., (2015), in which Est and IMF By control the rapidly varying part of the field. In the version we use, this is done by the SVMD index, a satellite-based modification (Kunagu et al., 2013) of the VMD index (Thomson and Lesur, 2007).

Further options apart from these are the CHAOS (Finlay et al., 2016, version six) and CM (Sabaka et al., 2015, version five) model series. We reject both in order to preserve independence of the RC index, which is used in the latest versions of both these series. The Dst and RC indices provide additional possibilities of comparison to check the plausibility of the two considered external field models. Fig. 3.3 shows a common shape of the magnetospheric signal from models and indices at the same four locations as in Fig. 3.2. The signal's amplitude decreases with distance from the geomagnetic equator (GUA→NGK) and so does the offset between models and indices. Most importantly POMME-8 and GRIMM-3.2 are in remarkable agreement despite their different modeling techniques. This suggests that the models provide the correct absolute level of the magnetospheric signal as opposed to the indices

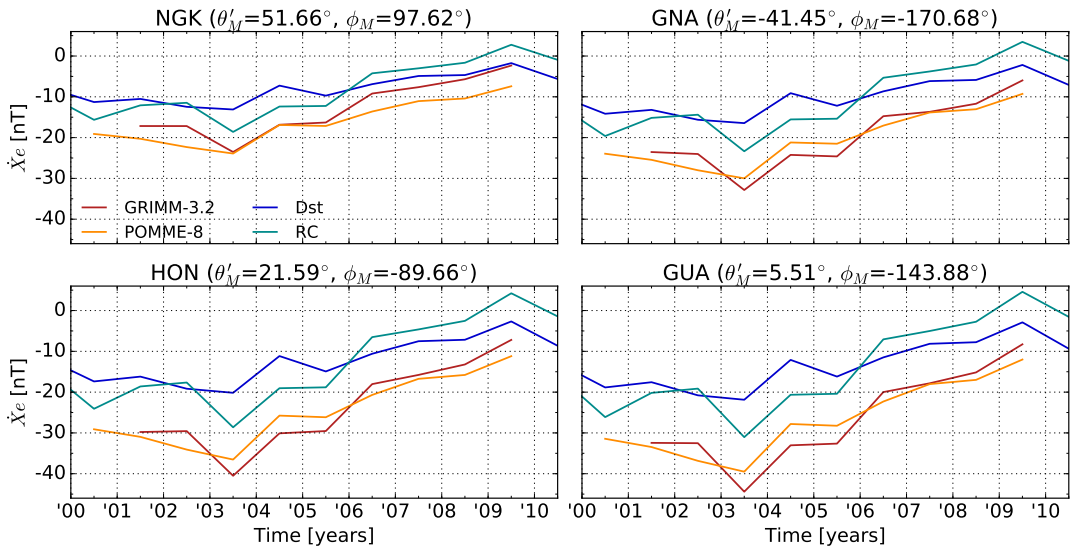

Figure 3.3: Yearly averaged $\dot{X}e$ components (external field, north, geocentric ned_c) of the large-scale external fields from GRIMM-3.2 (red) and POMME-8 (orange) in comparison with Dst (blue) and RC (green) indices. The decomposition of the indices into vector components is described at the end of section 3.3.3. Observatory locations are given in geomagnetic coordinates (*mag*).

(see section 3.3.1). We favor GRIMM-3.2 as external field model given the simplicity of running the forward code compared to POMME-8.

3.3 METHODS

In order to extract the large-scale external signal from the observatory measurements the core field as well as local lithospheric anomalies have to be subtracted at each site. While the main field is given by the COV-OBS field model (section 3.2.2.1), crustal offsets need to be estimated individually for each observatory. Below we describe the calculation of these biases followed by the calculation of measurement residuals, on the basis of which the AMC index is derived.

3.3.1 *Observatory crustal biases*

Time-dependent core field models are generally limited to SH degrees that include at most the large-scale lithospheric field (e.g. CM4 max. d/o 65). Crustal field models (e.g. Enhanced Magnetic Model EMM2017, Meyer et al., 2017) reach SH degrees up to 790 corresponding to a horizontal wavelength of 51 km. But even these dedicated models cannot resolve crustal anomalies in the direct vicinity of the observatories with amplitudes that may considerably exceed the external signal in question. The magnetic field contribution of such anomalies is commonly termed observatory crustal bias.

We assume the crustal biases to be constant over the period 1900–2010, thereby neglecting the possible variation of lithospheric anomalies that are induced. Between 1960 and 2002 the crustal field changes in the range $15 \leq \mathrm{d/o} \leq 19$ are reported

to be as weak as 0.06–0.12 nT/yr on average, not exceeding 1.3 nT/yr in South America at observatory Kourou (Thébault et al., 2009).
Temporal invariability of the crustal contribution allows us to base the calculation of biases on the GRIMM-3.2 field model (section 3.2.2.2) although it only covers nine years out of the considered 110-year period. Biases $\mathbf{Bb}_i$ are calculated for each observatory location $i = 1, \ldots, n$ by averaging the differences between the OAMs $\mathbf{Bo}_i(t)$ and core field $\mathbf{Bc}_i(t)$ (d/o ≤ 14) along with external field $\mathbf{Be}_i(t)$ (d/o ≤ 2) estimates from GRIMM-3.2:

$$\mathbf{Bb}_i = \frac{1}{9} \sum_{t=2001}^{2009} \mathbf{Bo}_i(t) - (\mathbf{Bc}_i(t) + \mathbf{Be}_i(t)) \tag{3.1}$$

These biases (values listed in appendix C) include contributions from the transitional spectrum range $15 \leq$ d/o ≤ 18. For these degrees GRIMM-3.2 yields a maximum root mean square core field intensity change of 0.08 nT/yr considering the locations of the 16 AMC candidate observatories (Fig. 3.1). This is negligible compared to the error introduced when subtracting the core field model and in line with the study cited above.
The advantage of the calculation method in Eqn. 3.1 is that we remove an estimation of the magnetospheric field $\mathbf{Be}_i(t)$ prior to averaging. This way the constant part of the external field signal is not included in the static lithospheric contribution $\mathbf{Bb}_i$, which is not achieved in the calculation of the Dst and RC indices. Both these indices underestimate the ring current effect by a constant amount (see Fig. 3.3), that corresponds to the signal of the quiet-time background current and was erroneously attributed to the observatory offsets.
This effect is also evident when comparing our biases to biases from studies that do not consider the large-scale external field. Fig. 3.4 shows differences between our biases $\mathbf{Bb}_i$ and values from three other studies, namely Langel et al., (1982), Mandea and Langlais, (2002) and Verbanac et al., (2015). While a detailed explanation of disagreements between these biases at specific locations is beyond the scope of our study, we want to highlight a systematic trend that is consistent across all studies. The components Xb_i (red) of our biases are larger than the reference biases by $\sim$27 nT on average (dashed red line). As the large-scale magnetospheric background field is reducing the main field's northward component, crustal biases calculated from uncorrected data will display lower amplitudes than those calculated from corrected data such as our offsets. On the basis of the new observatory biases we consequently expect the resulting AMC index to have the correct absolute level.

3.3.2 *Observatory residuals*

Observatory residuals $\mathbf{B}_i(t)$ can now be estimated by subtracting the modeled core field $\mathbf{Bc}_i(t)$ (COV-OBS, d/o ≤ 14) together with the corresponding crustal biases $\mathbf{Bb}_i$ (Eqn. 3.1) from the OAMs at each station $i = 1, \ldots, n$ and for each year in the long-term period:

$$\mathbf{B}_i(t) = \mathbf{Bo}_i(t) - \mathbf{Bc}_i(t) - \mathbf{Bb}_i \tag{3.2}$$

The residual vectors $\mathbf{B}_i(t)$ should then essentially consist of the external magnetospheric signal. The major part of this signal can be represented by only one

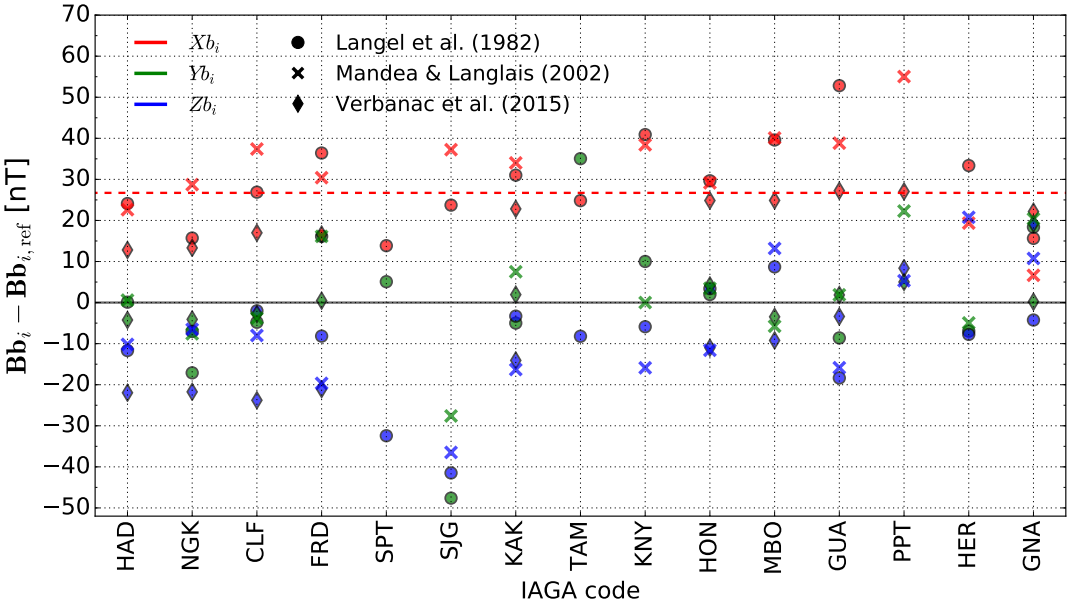

Figure 3.4: Differences between new biases (geodetic ned_d) and biases from three other studies (symbols). For these comparisons the effect of all jumps that occurred either after 1982 or after 2002 were removed (see also Verbanac et al., 2015). The values originally published by Langel et al., (1982) were taken from Gubbins and Bloxham, (1985). Biases attributed to Mandea and Langlais, 2002 are the means of their biases calculated either from Magsat or Ørsted data sets. The new Xb components (crustal bias, north, geodetic ned_d) are larger than those from the reference studies by 26.72 nT on average (red dashed line).

vector component, $Z_{M,i}(t)$, that points northward along the main field's dipole axis. Starting from local geodetic (i.e. spheroidal) coordinates (ned_d), $\mathbf{B}_i(t)$ is adjusted to geocentric (i.e. spherical) coordinates (ned_c or seu_c) before being transformed into the Cartesian geographic (GEO) and geomagnetic (MAG: X_M,Y_M,Z_M) systems. The position and field vectors are finally converted to a spherical system (mag: $B_{\theta,M}$,$B_{\phi,M}$,$B_{r,M}$) ready for further analysis:

$$\mathbf{B}_{ned_d} \underbrace{\overset{1}{\rightarrow}}_{\text{WGS-84}} [\mathbf{B}_{ned_c},\ \mathbf{B}_{seu_c}] \overset{2}{\rightarrow} \mathbf{B}_{\mathrm{GEO}} \underbrace{\overset{3}{\rightarrow}}_{\text{IGRF-12}} \mathbf{B}_{\mathrm{MAG}} \overset{4}{\rightarrow} \mathbf{B}_{mag} \tag{3.3}$$

Refer to appendix A for the nomenclature and to Hapgood, (1992) for coordinate transformations.

3.3.3 *Index derivation*

The pending task is to transfer the individual observatory residuals into a time-dependent index with global validity. The simplest option is to consider an average external signal $\overline{Z_M}(t)$ as the arithmetic mean of residual components $Z_{M,i}(t)$ from all contributing observatories at each year. This requires the residuals to be entirely independent from the observatory location, which is analyzed in section 3.4.1. A more sophisticated option is to apply a year-by-year SHA to the residuals $\mathbf{B}_{mag}$, called $\mathbf{B}$ hereafter, in which case their dependency on location (θ_M, ϕ_M, r -

geomagnetic colatitude, longitude, radius) is considered. The observatory residuals are taken to be the negative gradient of a scalar potential, V_e, that fulfills Laplace's equation:

$$\mathbf{B} = -\nabla V_e \,, \qquad \nabla^2 V_e = 0 \tag{3.4}$$

We consider an external large-scale geomagnetic potential whose expansion into spherical harmonics is truncated at degree one:

$$\begin{aligned} V_e(\theta_M, \phi_M, r) = a \sum_{m=0}^{1} \left(\frac{r}{a}\right) \left[q_1^m \cos\left(m \cdot \phi_M\right)\right. \\ \left. + s_1^m \sin\left(m \cdot \phi_M\right)\right] P_1^m\left(\cos\left(\theta_M\right)\right) \end{aligned} \tag{3.5}$$

The potential increases with radial distance from Earth's center (r) towards the source region in the magnetosphere and is evaluated at a mean radius of $a = 6371$ km. Above, $\{{}^{q}_{s}\}_1^m$ denote the three Gauss coefficients of degree one and order m and P_1^m are the Schmidt semi-normalized Associated Legendre Polynomials (Schmidt, 1917) including the Condon-Shortley phase factor.
The combined observatory residual vector $\mathbf{B}$ ($k = 1, \ldots, N_D = 3 \cdot n$ entries) contains three vector components from each of the n observatories. It is connected to the vector of Gauss coefficients $\mathbf{x}$ (3 entries) via the ($N_D \times 3$) matrix $\mathbf{A}$:

$$\underbrace{\begin{pmatrix} B_{\theta 1} \\ B_{\phi 1} \\ B_{r1} \\ \vdots \\ B_{rn} \end{pmatrix}}_{\mathbf{B}} = \underbrace{\begin{pmatrix} \sin(\theta_1) & \cos(\phi_1)\cos(\theta_1) & \sin(\phi_1)\cos(\theta_1) \\ 0 & -\sin(\phi_1) & \cos(\phi_1) \\ -\cos(\theta_1) & \cos(\phi_1)\sin(\theta_1) & \sin(\phi_1)\sin(\theta_1) \\ \vdots & \vdots & \vdots \\ -\cos(\theta_n) & \cos(\phi_n)\sin(\theta_n) & \sin(\phi_n)\sin(\theta_n) \end{pmatrix}}_{\mathbf{A}} \cdot \underbrace{\begin{pmatrix} q_1^0 \\ q_1^1 \\ s_1^1 \end{pmatrix}}_{\mathbf{x}} \tag{3.6}$$

The inversion of this linear forward problem for the coefficients $\mathbf{x}$ should consider the propagation of errors from $\mathbf{B}$ that are mainly due to uncertainties in the COV-OBS main field results (section 3.2.2.1). We therefore adopt a Bayesian approach (method detailed in e.g. Gelman et al., 2013) in which we assume an uninformative prior probability density function (PDF), $\rho(\mathbf{x}) = 1$, and a likelihood distribution $\rho(\mathbf{B}|\mathbf{x})$ of multivariate Gaussian type $\mathbf{B} \propto \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The latter describes the spread of the residuals, $\mathbf{B}$, around reality, modeled as $\mathbf{A} \cdot \mathbf{x}$, according to uncertainties from COV-OBS. The residual mean, $\boldsymbol{\mu}$ (N_D entries), and covariance matrix, $\boldsymbol{\Sigma}$ ($N_D \times N_D$), are calculated from the COV-OBS main field ensemble comprising $j = 1, \ldots, N_E = 100$ residual vectors, $\mathbf{B}_j$, per time step. We neglect the off-diagonal elements in $\boldsymbol{\Sigma}$ because errors in the components of $\mathbf{B}$ and among the observatories are assumed to be uncorrelated. This way $\boldsymbol{\Sigma}$ is diagonal and contains the squares of standard deviations σ:

$$\begin{aligned} \mu_k &= \frac{1}{N_E} \sum_{j=1}^{N_E} B_{jk} \\ \Sigma_{k,k} = {\sigma_k}^2 &= \frac{1}{N_E - 1} \sum_{j=1}^{N_E} \left(B_{jk} - \mu_k\right)^2 \end{aligned} \tag{3.7}$$

The posterior PDF $\rho(\mathbf{x}|\mathbf{B})$ is Gaussian as well with $\mathbf{x} \propto \mathcal{N}(\boldsymbol{\mu}_\mathrm{x}, \boldsymbol{\Sigma}_\mathrm{x})$, so its maximization yields the mean of the Gauss coefficients:

$$\boldsymbol{\mu}_\mathrm{x} = \left(\mathbf{A}^T\boldsymbol{\Sigma}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^T\,\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \tag{3.8}$$

The posterior covariance matrix, $\boldsymbol{\Sigma}_\mathrm{x}$ (3×3 entries), is symmetric with negligible off-diagonal elements. It contains the variances of the coefficients on the diagonal:

$$\boldsymbol{\Sigma}_\mathrm{x} = \left(\mathbf{A}^T\boldsymbol{\Sigma}^{-1}\mathbf{A}\right)^{-1} = \begin{pmatrix} \sigma_\mathrm{x}(q_1^0)^2 & \bullet & \circ \\ \bullet & \sigma_\mathrm{x}(q_1^1)^2 & \diamond \\ \circ & \diamond & \sigma_\mathrm{x}(s_1^1)^2 \end{pmatrix} \tag{3.9}$$

Because the strongest magnetospheric signal is aligned with the dipole axis, we define the AMC index as:

$$\mathrm{AMC} := -q_1^0 \propto \mathcal{N}\left(-\mu_\mathrm{x}(q_1^0), \sigma_\mathrm{x}(q_1^0)^2\right) \tag{3.10}$$

The negative sign makes our index definition comparable with those of the Dst and RC indices as southward disturbances that are given with respect to the Z_M axis pointing northward along the dipole axis. In further agreement with Dst and RC, the AMC index contains the direct external signal as well as the signal that is induced in the conductive upper layers of Earth's interior. A separation of these two parts is possible following the strategy of Maus and Weidelt, (2004) or Olsen et al., (2005), but is not necessary for our purposes at this stage.
The AMC index can be used to correct observatory data (Bo_θ, Bo_ϕ, Bo_r) for large-scale external disturbances. To do so, AMC intensity needs to be split in three vector components (AMC_θ, AMC_ϕ, AMC_r) according to the following steps:

(i) Calculate the position of the geomagnetic dipole ($\theta_P{}'$,ϕ_P) and accordingly convert the geocentric observatory position (θ_c, ϕ, r) to geomagnetic coordinates (θ_M, ϕ_M, r). Colatitudes are named θ whereas latitudes are labeled θ'.

(ii) Determine the angle (D) between geographic and geomagnetic north:

$$D = s \cdot \cos^{-1}\left(\frac{\sin(\theta_P{}') - \sin(\theta_c{}')\cos(\theta_M)}{\cos(\theta_c{}')\sin(\theta_M)}\right), \quad s = \begin{cases} -1 & \text{for } 0^\circ \leqslant \phi - \phi_P \leqslant 180^\circ \\ 1 & \text{otherwise} \end{cases} \tag{3.11}$$

(iii) The index components are:

$$\begin{aligned} \mathrm{AMC}_r &= \cos(\theta_M) \cdot \mathrm{AMC} \\ \mathrm{AMC}_H &= -\sin(\theta_M) \cdot \mathrm{AMC} \\ \mathrm{AMC}_\theta &= \cos(D) \cdot \mathrm{AMC_H} \\ \mathrm{AMC}_\phi &= -\sin(D) \cdot \mathrm{AMC_H} \end{aligned} \tag{3.12}$$

The corrected data is obtained by subtracting the appropriate index component in Eqn. 3.12 from the corresponding observatory measurement.

3.4 RESULTS

We begin with an analysis of the annual residual mean ($\boldsymbol{\mu}$, Eqn. 3.7) time series from the 16 observatories in the reference set (Fig. 3.1), that leads the way to the presentation of the AMC index.

3.4.1 *Reference period: 1960–2010*

Fig. 3.5 (top) shows residual mean components, Z_M, that are sorted by the geomagnetic latitudes of the corresponding observatories from north (top) to south (bottom). The vertical stripes illustrate a pronounced similarity of the residuals at all sites. There is no distinct dependence on location.
The degree of the signal's global uniformity improves with time. The middle panel of Fig. 3.5 shows the scatter in amplitude (error bars) around the average external signal $\overline{Z_M}(t)$ from the 16 AMC candidate stations (red IAGA codes). For the period 1999–2010, the mean scatter is only 49 % of what it was in the previous period 1960–1998. We attribute this reduction to the fact that COV-OBS profits from continuous vector satellite data since the turn of the century (Fig. 3.2). Note that the scatter in amplitude would have been larger by $\sim$11 nT on average, if we had used CM4 (57 nT) instead of COV-OBS (46 nT) as main field model.
The variations of the average external signal show clear resemblances to other geomagnetic indices shown in Fig. 3.5 (bottom, sources in appendix B). Altogether the extracted signal is consistent with the globally southward directed (negative Z_M values) magnetic field signature of the magnetospheric current system in response to the solar cycle. Low-frequency variations thus respond to the F10.7 index with some time-lag that cannot be estimated reasonably based on annual means but should be about 20 months according to (Lühr and Maus, 2010). Under solar minimum conditions the signal is expected to be particularly small, albeit still negative. An extremely quiet year was 1965 when the only positive residual is calculated at TAM (11 nT). If not a data issue, this could be caused by a non-vanishing contribution from ionospheric currents, too small to be recognized at other times. The exceptional character of that year is also expressed by the uniquely positive Dst value (dashed line). This was corrected for the revised Dcx index (black line) by Karinen and Mursula, (2005). However, our average external signal (middle panel, blue line) correlates better with the Ap index (Pearson coefficient r = -0.94) than with the Dcx index (r = 0.87).
In the top panel of Fig. 3.6a) the average external signal, $\overline{Z_M}(t)$ (dashed line), is compared to the AMC index (solid line) calculated either from the reference set of nine observatories (green) or the final set of five observatories (blue). For recent years, the temporal variability of all these results agrees with that of the annual RC index (black). Likewise it agrees with an average over external field estimates from GRIMM-3.2 (section 3.2.2.2) for the locations of observatories in the final set (red). As expected (section 3.3.1) there is an offset between the RC index and the final AMC index, which amounts to -12.2 nT $\pm$ 1.2 nT and is discussed further in section 3.5.1.

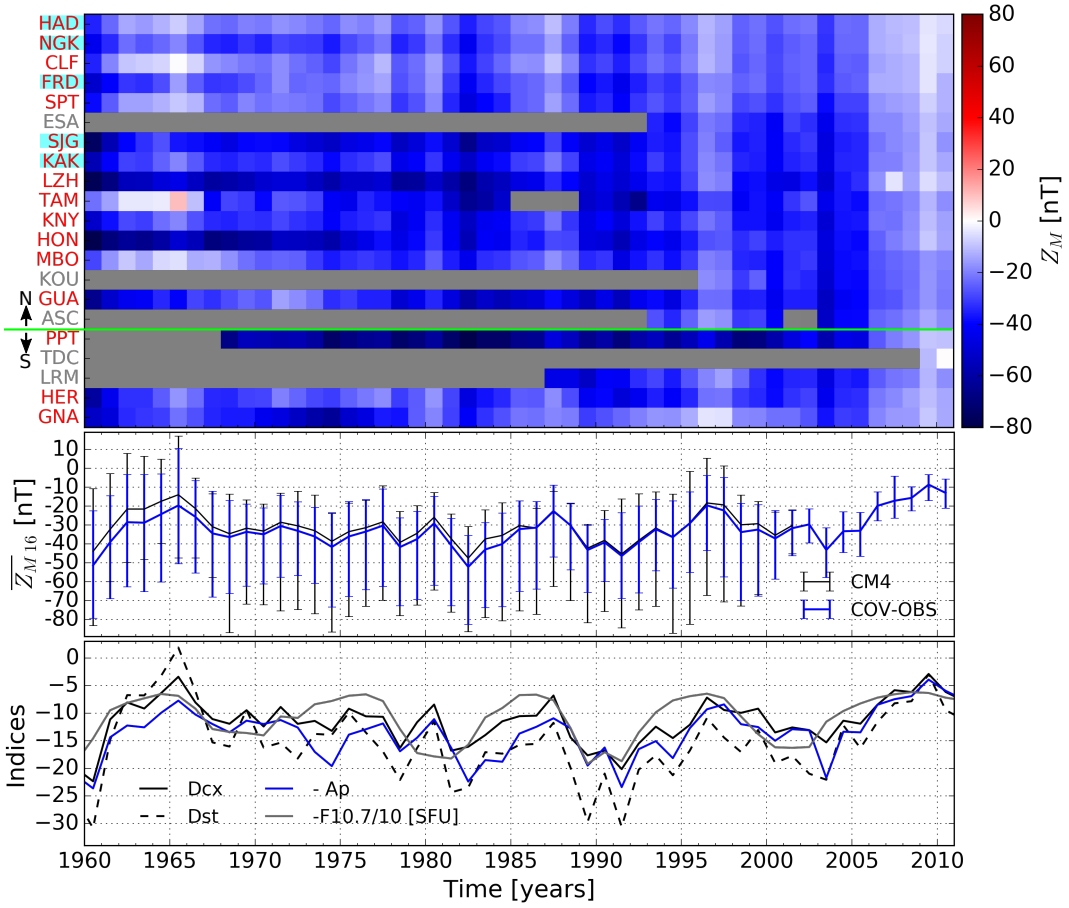

Figure 3.5: Top: Observatory residuals Z_M (north, geomagnetic MAG) from 21 observatories (Fig. 3.1) sorted by geomagnetic latitude from north (top) to south (bottom). The green horizontal line separates observatories north and south of the geomagnetic equator. Missing annual means are marked in gray. IAGA codes of the 16 AMC candidate observatories (red) and the five final observatories (cyan) are highlighted. Middle: Mean external signal $\overline{Z_M}$ from AMC candidate stations calculated using either COV-OBS or CM4. Error bars indicate the maximum spread from the mean. Bottom: Annual means of the Dcx, Dst, Ap (negative) and absolute solar flux index F10.7 (negative & scaled).

The error bars of the two AMCs overlap at all times within the 50-year period (orange shaded region), except in the year 2003. The corresponding average external signals tend to lie at comparable or more negative values than the AMCs and fall within the overlapping region at most years. However, they are noticeably cutting across the AMCs during the years 1979–1980 and 1999–2010, clearly surpassing the common uncertainty boundaries in the latter case.

This behavior is linked to the variances of the core field estimates from COV-OBS, that are located on the diagonal of the covariance matrix $\mathbf{\Sigma}$ (Eqn. 3.7). The normalized trace of $\mathbf{\Sigma}$ measures the spatially averaged core field uncertainty and is shown in the bottom panel of Fig. 3.6a) for both location sets. If the core field uncertainty decreases, so does the uncertainty of the AMC index. This relation scales with the number of contributing observatories, n, such that a gain in information, i.e. a greater n, results in smaller AMC uncertainties. The efficiency of the uncertainty

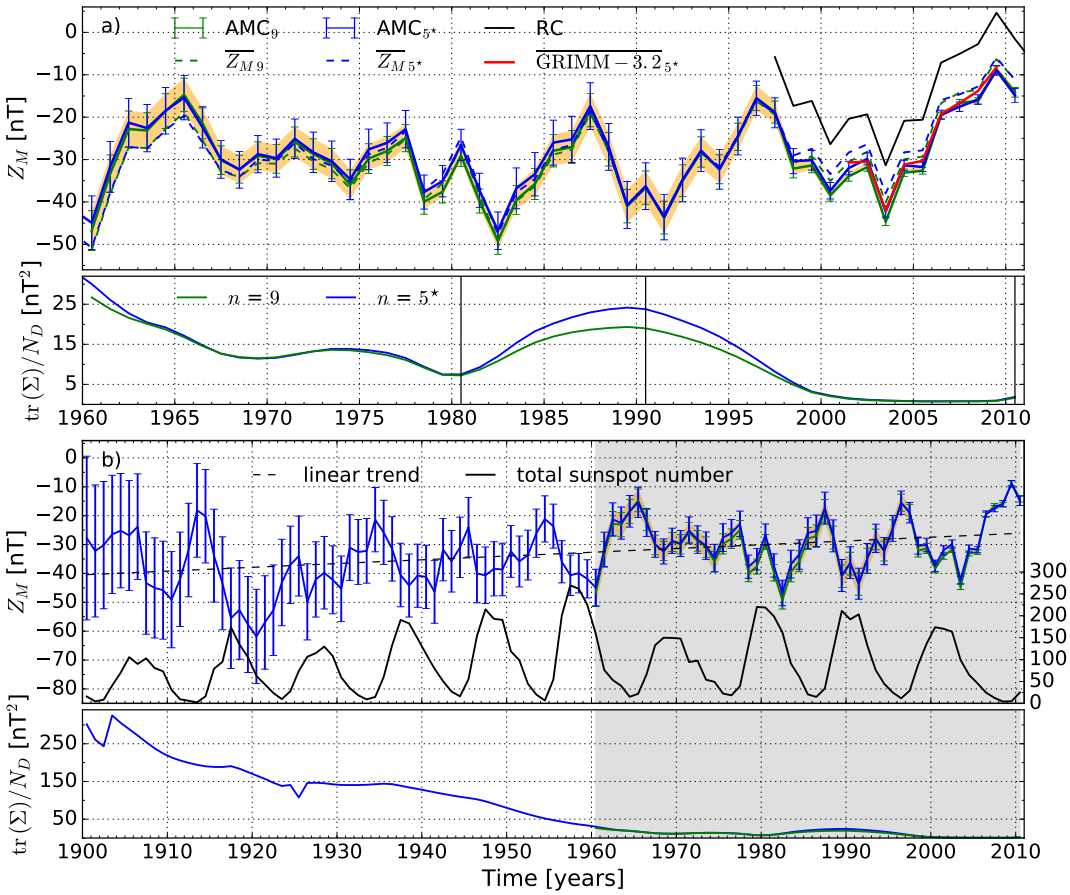

Figure 3.6: a) Top: Mean external signal $\overline{Z_M}$ (dashed line) and AMC index with ∼99.73 % confidence (solid line with error bars) for the reference set (green) and the final set (blue) of observatories (Fig. 3.1). The times when the error bars overlap are shaded in orange. Also shown are the RC index (black) and external field estimates from GRIMM-3.2 averaged over the locations of stations in the final set (red). a) Bottom: Trace of the COV-OBS covariance matrix $\mathbf{\Sigma}$ normalized by the number of data points N_D from the same two data sets. Vertical black lines indicate years with distinct global patterns of COV-OBS main field uncertainties that are presented in appendix D, Fig. 3.D1. b): Gray shaded region is the same as in a) but here the complete time span 1900–2010 is shown with the total sunspot number for comparison (right ordinate). Estimated linear trend for the AMC index is 0.13 nT/yr (dashed black line).

reduction depends on the specific choice of observatories with regard to the global distribution of core field uncertainty. If satellite data is available to COV-OBS, e.g. in years 1980 and 2010 (Fig. 3.2), differences in core field uncertainty amongst mid latitude observatories are marginal, so that their distribution is irrelevant to the AMC index. However, for all other times, e.g. year 1990, COV-OBS' core field uncertainties are globally heterogeneous, favoring observatories in regions with dense coverage, especially from Europe and Japan. Maps showing the core field uncertainty from COV-OBS for the mentioned years can be found in appendix D, Fig. 3.D1.

3.4.2 *Long-term period: 1900–2010*

The final AMC index cannot be calculated from the OAMs in the reference set (Fig. 3.1), because three out of the nine observatories (KNY, LZH, SPT) do not provide long enough time series. Additionally, special care has to be taken of observatory relocations. Observatory residuals in the style of Fig. 3.5 with marked relocations can be found in appendix D, Fig. 3.D2 for the long-term period. The figure indicates that there are serious issues with the OAMs from Val Joyeux (VAL), the predecessor of CLF, persisting roughly from 1906–1931. The complete time series was therefore discarded which led to the final data set.
Based on core field estimations from COV-OBS we found that the secular variations at subsequent locations of an observatory can differ substantially. In particular, discrepancies are large between Greenwich (GRW) and Hartland (HAD), which are separated by more than 300 km. We therefore calculated observatory residuals individually for each location and reassembled the sections by applying an offset accounting for the different observatory biases. This offset is estimated as the difference between the jump modeled with COV-OBS (not including biases) and the jump in the OAMs (including biases) at the year of the relocation. Data jumps are documented in the BGS annual records, but in the case of POT→SED→NGK the POT and SED values were already adjusted to the NGK location and jump values had to be taken from (Bock, 1950).
The resulting long-term AMC index is shown in Fig. 3.6b) including the two panels from a) in the gray shaded region. The increase of AMC uncertainties at earlier times is expected, as is the fact that local minima closely follow peaks in the total sunspot number (black, right ordinate). The AMC shows a small positive trend over the 110-year period amounting to $\sim$0.13 nT/yr, that is not matched by the solar activity evolution. In view of the large uncertainties we cannot confirm this trend to be physically real at this point.

3.5 DISCUSSION

In this section we assess the presented AMC index in several regards. First, we concentrate on the absolute level of the index and secondly discuss our observatory choice as constrained by the index variance. Lastly, we apply the AMC index to observatory measurements from the reference period and illustrate implied scientific opportunities.

3.5.1 *AMC absolute level*

The excellent agreement of signal amplitudes between AMC and the GRIMM-3.2 estimate (Fig. 3.6a) is not conclusive given the fact that this model was used to free our crustal biases from the average magnetospheric field (section 3.3.1). However, the $\sim$13 nT amplitude of this field (Lühr and Maus, 2010, section 3.1) is adequately met by the average offset between AMC and RC indices of -12.2 nT (Fig. 3.6a). Furthermore, this offset is in the range of recently reported offsets amounting to -9.1 nT needed for the RCE to match annual averages of CHAMP Z_M values

(Lühr et al., 2017) and -10 nT needed to match the $-q_1^0$ SH coefficient of the CM5 geomagnetic model (Sabaka et al., 2015). The significance of both these comparisons is partially restricted by the fact that AMC includes secondary induced signals in addition to the direct external field and did not undergo any subjective quiet-time selection procedure. Yet the reasonable agreements imply that the new crustal biases enable us to determine the correct baseline for the AMC index. Lastly, it is striking that the average external signal $\overline{Z_M}(t)$ surpasses the AMC index only in periods when vector satellite data is available (section 3.4.1). This suggests that COV-OBS might overweight this data with respect to ground measurements, leading to an underestimation of uncertainties (see also Fig. 3.2). Obviously, this effect is especially strong since the beginning of the 20th century.

3.5.2 *Observatory selection*

The selection of the observatories in the reference set is based on the finding that the information extracted from different observatory measurements is redundant (see Fig. 3.5, top). Deviating AMC results for different observatory sets consequently arise from the input variance only and are manifest in the AMC variance (Fig. 3.6). Consequently the reference set should

(i) minimize the average variance over the period 1960–2010;

(ii) maximize the robustness against observatory replacement;

(iii) minimize computational complexity;

To address requirement (i), we calculated the AMC variance ($\sigma_x(q_1^0)^2$, Eqn. 3.9) for all possible combinations of candidate observatories, normalized by the average AMC variance from all these 16 stations, and averaged over the time span 1960–2010 (α). Picking the observatories associated with the minimum α ("optimum set") for each set size leads to a characteristic relation describing the most efficient reduction of index variance with data increase (Fig. 3.7a). Although not imposed a priori, the optimum sets are generated by cumulatively adding the observatories in a specific order starting with NGK (Fig. 3.7c, top). Regarding requirement (ii), we then calculated the coefficient of variation (CV) for the α distributions of each set size (Fig. 3.7c). CV increases linearly for sets of 15 down to nine, then rises more quickly as the observatory number decreases to one. Owing to requirement (iii), we ultimately decided for the reference set with nine stations as a compromise between a low CV and a minimum observatory number (circle in Fig. 3.7c).
Although the five stations in the final AMC set (section 3.4.2) are not those in the optimum set of size five {NGK,KAK,CLF,KNY,HAD}, they still fall below the 10th percentile of the corresponding α distribution (Fig. 3.7b). This means that they sufficiently, yet not optimally, reduce the index variance during 1960–2010 while being as robust as possible against observatory drop outs during 1900–2010.

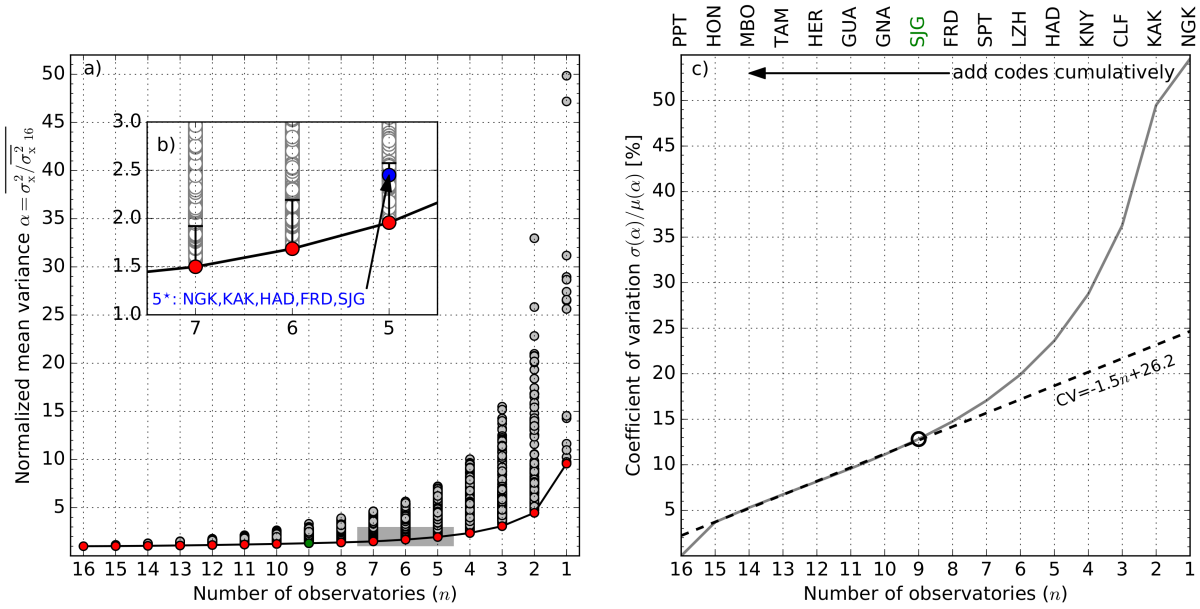

Figure 3.7: a): Mean posterior variance (α) during 1960–2010 for all combinations of a particular number of observatories (gray circles) ranging from 16 to one. The black line connects station sets with minimum α for each set size ("optimum sets", red circles). The green circle marks the chosen reference set. b): Blow-up of shaded area in a). The final AMC set (5⋆, blue circle) falls below the 10th percentile (black horizontal bar) of the corresponding α distribution. c): Coefficient of variation (CV) for the α distributions (gray line) and linear fit to n=9,...,15 observatories (dashed line). The optimum set size is fixed at n = 9 (black circle). IAGA codes at the top indicate which n observatories (from right to left) form the optimum sets.

3.5.3 *AMC application*

We now turn to a demonstration of AMC index performance regrading the removal of external field signal from observatory data. Fig. 3.8 shows SV in the local southward component (gray), because we expect the strongest regular external field influence in this direction. SV is calculated from first differences of OAMs measured at four locations. This is compared to modeled SV from COV-OBS (blue) and CM4 (green). The final AMC index is decomposed into its vector components (section 3.3.3) and the AMC_θ time series is subtracted from the observatory measurements prior to the calculation of corrected first differences (red). In general this reduces the short-term variability of the measured SV time series so that the corrected curves display less variability than the originals, but greater variability than the smoothed global model results. This allows the identification of times at which the models might suffer locally from an incomplete separation between internal and large-scale external field components. For instance, CM4 seems to show some external field leakage around the NGK location during 1995–2001, while COV-OBS might be affected by a similar effect around GNA during 1975–1980. The latter also seems to be the case for the region around HON.
The corrected SV time series are of particular interest when trying to identify the typical "V"-shaped signature of geomagnetic jerks. This is often done on the basis

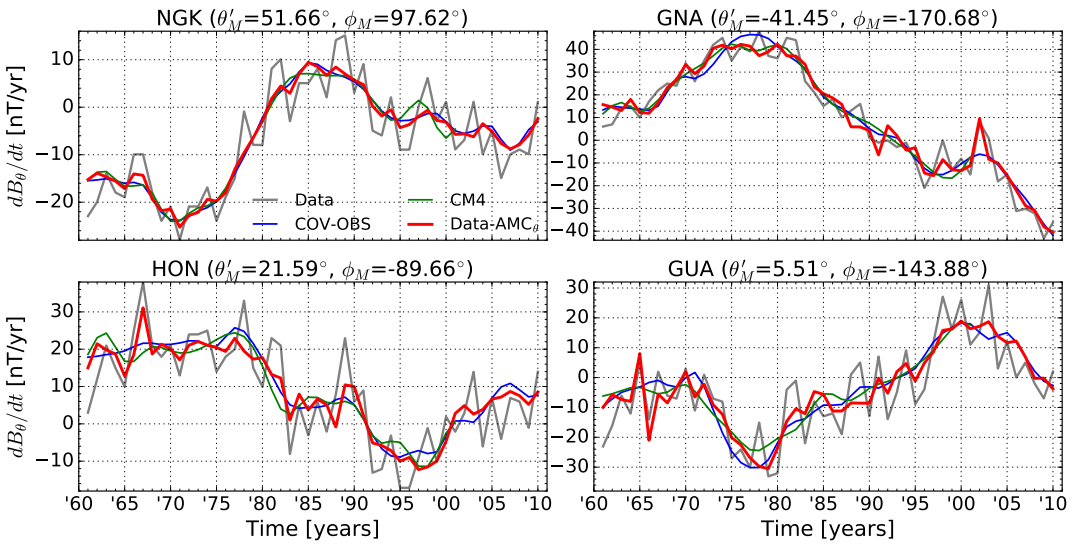

Figure 3.8: Comparison between SV in B_θ components (south, geocentric seu_c) from observatory measurements (gray) and AMC-corrected measurements (red) compared to modeled SV from COV-OBS (blue) and CM4 (green). The AMC index was decomposed into its vector components according to the observatory locations. These are given here in geomagnetic coordinates (mag).

of monthly SV as in Brown et al., (2013) who use a method developed by Wardinski and Holme, (2011) to remove external signal from the SV measurements. This signal is approximated by the SV residual between observatory data (a popular choice is NGK) and an internal magnetic field model. Compared to this procedure the application of the AMC index with improved time resolution would be less labor-intensive and more easily applicable regarding a worldwide use.
Our next aim is to increase the AMC index' temporal resolution to hourly means. This requires a dedicated removal of the ionospheric disturbance field from the observatory residuals. One approach would be to consider the magnetic field of the Sq current system as modeled by CM4 on the basis of F10.7, which reaches back to 1947. Annual reconstructions of F10.7 are available since the 17th century (Krivova et al., 2010). These could be sufficient to parametrize the solar cycle dependence even for the hourly index. Alternatively, Empirical Orthogonal Functions could be applied to the observatory residuals in order to extract the spatiotemporal pattern of the ionospheric disturbance field as in Shore et al., (2016). In either case, an AMC index with hourly resolution could readily be used by researchers investigating geomagnetic jerks as well as by global geomagnetic field modelers.

3.6 CONCLUSIONS

The AMC index is based on five northern hemispheric observatories (HAD, NGK, FRD, SJG, KAK) providing annual mean data from 1900 to 2010. Compared to the Dst and RC indices it does include the average magnetospheric background signal as a consequence of recalculated crustal biases, which have systematically larger northward components than previously published values. The average offset of -12.2 nT from the RC index during 1997–2010 seems plausible when compared to

other studies indicating that the intensity of the near-Earth magnetospheric field is represented correctly. Index uncertainties originate from those in the COV-OBS main field model, that can be relatively large and globally heterogeneous during the observatory era but decrease and converge notably in years with vector satellite data. At such times we suppose that COV-OBS systematically underestimates uncertainties. The final observatory combination is chosen because it provides a reasonably low average AMC variance while being robust against station replacements in case of data gaps or observatory drop outs. The preference of clustered stations over a globally uniform distribution is an interesting result in view of the signal's large-scale character. We plan to investigate this matter further when we upgrade the temporal resolution of the AMC index to hourly means in order to meet current demands for SV studies and geomagnetic field modeling.

ACKNOWLEDGMENTS

We sincerely thank Julien Baerenzung for his advice particularly regarding the implementation of the inversion scheme and the interpretation of the results. We also thank the institutions that operate the geomagnetic observatories and host geomagnetic field models and activity indices. Special thanks go to Vincent Lesur for sharing a version of GRIMM including external sources. We appreciate the constructive comments from two anonymous reviewers, that helped to improve the manuscript.
This study is funded by the Priority Program 1788 "Dynamic Earth" of the German Research Foundation under grant KO2870/6-1.

REFERENCES

Bock, R. (1950). "Ergebnisse der Beobachtungen am Adolf-Schmidt-Observatorium für Erdmagnetismus in Niemegk in den Jahren 1932 und 1933." In: *Akademie Verlag Berlin*, p. 26 (cit. on p. 39).

Brown, W. J., J. E. Mound, and P. W. Livermore (2013). "Jerks abound: An analysis of geomagnetic observatory data from 1957 to 2008." In: *Physics of the Earth and Planetary Interiors* 223, pp. 62–76. DOI: 10.1016/j.pepi.2013.06.001 (cit. on p. 42).

Daglis, I. A., R. M. Thorne, W. Baumjohann, and S. Orsini (1999). "The terrestrial ring current: Origin, formation, and decay." In: *Reviews of Geophysics* 37.4, pp. 407–438. DOI: 10.1029/1999RG900009 (cit. on p. 26).

Finlay, C. C., N. Olsen, S. Kotsiaros, N. Gillet, and L. Tøffner-Clausen (2016). "Recent geomagnetic secular variation from Swarm and ground observatories as estimated in the CHAOS-6 geomagnetic field model." In: *Earth, Planets and Space* 68.112. DOI: 10.1186/s40623-016-0486-1 (cit. on p. 30).

Gannon, J. L. and J. J. Love (2011). "USGS 1-min Dst index." In: *Journal of Atmospheric and Solar-Terrestrial Physics* 73.2, pp. 323–334. DOI: 10.1016/j.jastp.2010.02.013 (cit. on p. 27).

Gelman, A. et al. (2013). *Bayesian Data Analysis*. 3rd ed. Chapman and Hall/CRC. ISBN: 9781439840955 (cit. on p. 34).

Gillet, N., D. Jault, C. C. Finlay, and N. Olsen (2013). "Stochastic modeling of the Earth's magnetic field: Inversion for covariances over the observatory era." In: *Geochem. Geophys. Geosyst.* 14.4, pp. 766–786. DOI: 10.1002/ggge.20041 (cit. on p. 28).

Gubbins, D. and J. Bloxham (1985). "Geomagnetic field analysis – 3. Magnetic fields on the core – mantle boundary." In: *Geophys. J. Roy. Astron. Soc.* 80.3, pp. 695–713 (cit. on p. 33).

Hapgood, M. A. (1992). "Space physics coordinate transformations: A user guide." In: *Planetary and Space Science* 40.5, pp. 711–717. DOI: 10.1016/0032-0633(92)90012-D (cit. on p. 33).

Jackson, A., A. R. T. Jonkers, and M. R. Walker (2000). "Four centuries of geomagnetic secular variation from historical records." In: *Philosophical Transactions of the Royal Society A* 358.1768, pp. 957–990. DOI: 10.1098/rsta.2000.0569 (cit. on p. 28).

Karinen, A. and K. Mursula (2005). "A new reconstruction of the Dst index for 1932-2002." In: *Annales Geophysicae* 23.2, pp. 475–485. DOI: 10.5194/angeo-23-475-2005 (cit. on p. 36).

Korte, M. and V. Lesur (2012). "Repeat station data compared to a global geomagnetic field model." In: *Annals of Geophysics* 55.6, pp. 1101–1111. DOI: 10.4401/ag-5410 (cit. on p. 26).

Krivova, N. A., L. E. A. Vieira, and S. K. Solanki (2010). "Reconstruction of solar spectral irradiance since the Maunder minimum." In: *Journal of Geophysical Research: Space Physics* 115.A12112. DOI: 10.1029/2010JA015431 (cit. on p. 42).

Kunagu, P., G. Balasis, V. Lesur, E. Chandrasekhar, and C. Papadimitriou (2013). "Wavelet characterization of external magnetic sources as observed by CHAMP satellite: evidence for unmodelled signals in geomagnetic field models." In: *Geophysical Journal International* 192.3, pp. 946–950. DOI: 10.1093/gji/ggs093 (cit. on p. 30).

Langel, R. A. and R. H. Estes (1985a). "Large-scale, near-field magnetic fields from external sources and the corresponding induced internal field." In: *Journal of Geophysical Research: Solid Earth* 90.B3, pp. 2487–2494. DOI: 10.1029/JB090iB03p02487 (cit. on p. 27).

– (1985b). "The near-Earth magnetic field at 1980 determined from Magsat data." In: *Journal of Geophysical Research: Solid Earth* 90.B3, pp. 2495–2509. DOI: 10.1029/JB090iB03p02495 (cit. on p. 27).

Langel, R. A., G. D. Mead, and R. H. Estes (1982). "Some new methods in geomagnetic field modeling applied to the 1960-1980 epoch." In: *J. Geomag. Geoelec.* 34.6, pp. 327–349 (cit. on pp. 32, 33).

Lesur, V. et al. (2015). "Parent magnetic field models for the IGRF-12 GFZ-candidates." In: *Earth, Planets and Space* 67.1, pp. 87–102. DOI: 10.1186/s40623-015-0239-6 (cit. on p. 30).

Lühr, H. and S. Maus (2010). "Solar cycle dependence of quiet-time magnetospheric currentsand a model of their near-Earth magnetic fields." In: *Earth, Planets and Space* 62, pp. 843–848. DOI: 10.5047/eps.2010.07.012 (cit. on pp. 27, 30, 36, 39).

Lühr, H., C. Xiong, N. Olsen, and G. Le (2017). "Near-Earth Magnetic Field Effects of Large-Scale Magnetospheric Currents." In: *Space Science Reviews* 206.1, pp. 521–545. DOI: 10.1007/s11214-016-0267-y (cit. on pp. 27, 40).

Macmillan, S. and N. Olsen (2013). "Observatory data and the Swarm mission." In: *Earth, Planets and Space* 65.11, pp. 1355–1362. DOI: 10.5047/eps.2013.07.011 (cit. on p. 48).

Mandea, M. and B. Langlais (2002). "Observatory crustal magnetic biases during MAGSAT and Ørsted satellite missions." In: *Geophysical Research Letters* 29.15, ORS 4-1–ORS 4-4. DOI: 10.1029/2001GL013693 (cit. on pp. 32, 33).

Maus, S., C. Manoj, J. Rauberg, I. Michaelis, and H. Lühr (2010). "NOAA/NGDC candidate models for the 11th generation International Geomagnetic Reference Field and the concurrent release of the 6th generation Pomme magnetic model." In: *Earth, Planets and Space* 62.10, pp. 729–735. DOI: 10.5047/eps.2010.07.006 (cit. on p. 29).

Maus, S. and P. Weidelt (2004). "Separating the magnetospheric disturbance magnetic field into external and transient internal contributions using a 1D conductivity model of the Earth." In: *Geophysical Research Letters* 31.12, pp. 1–4. DOI: 10.1029/2004GL020232 (cit. on pp. 26, 35).

McPherron, R. L. (1995). "Introduction to Space Physics." In: ed. by M. G. Kivelson and C. T. Russell. Cambridge University Press. Chap. Magnetopsheric Dynamics, pp. 400–458. ISBN: 9780521457149 (cit. on p. 26).

Meyer, B., R. Saltus, and A. Chulliat (2017). *EMAG2: Earth Magnetic Anomaly Grid (2-arc-minute resolution) Version 3. National Centers for Environmental Information. NOAA. Model* (cit. on p. 31).

Mursula, K., L. Holappa, and A. Karinen (2011). "Uneven weighting of stations in the Dst index." In: *Journal of Atmospheric and Solar-Terrestrial Physics* 73.2-3, pp. 316–322. DOI: 10.1016/j.jastp.2010.04.007 (cit. on p. 26).

Mursula, K. and A. Karinen (2005). "Explaining and correcting the excessive semiannual variation in the Dst index." In: *Geophysical Research Letters* 32.14. DOI: 10.1029/2005GL023132 (cit. on p. 26).

Olsen, N., K.-H. Glassmeier, and X. Jia (2010). "Separation of the Magnetic Field into External and Internal Parts." In: *Space Science Reviews* 152.1, pp. 135–157. DOI: 10.1007/s11214-009-9563-0 (cit. on p. 26).

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. DOI: 10.1186/BF03351897 (cit. on pp. 26, 27, 35).

Olsen, N. et al. (2015). "The Swarm Initial Field Model for the 2014 geomagnetic field." In: *Geophysical Research Letters* 42.4, pp. 1092–1098. DOI: 10.1002/2014GL062659 (cit. on p. 27).

Olsen, N. et al. (2014). "The CHAOS-4 geomagnetic field model." In: *Geophysical Journal International* 197.2, pp. 815–827. DOI: 10.1093/gji/ggu033 (cit. on p. 27).

Sabaka, T. J., N. Olsen, and M. E. Purucker (2004). "Extending comprehensive models of the Earth's magnetic field with Ørsted and CHAMP data." In: *Geophysical Journal International* 159.2, pp. 521–547. DOI: 10.1111/j.1365-246X.2004.02421.x (cit. on p. 29).

Sabaka, T. J., N. Olsen, R. H. Tyler, and A. V. Kuvshinov (2015). "CM5, a pre-Swarm comprehensive geomagnetic field model derived from over 12 yr of CHAMP, Ørsted, SAC-C and observatory data." In: *Geophysical Journal International* 200.3, pp. 1596–1626. DOI: 10.1093/gji/ggu493 (cit. on pp. 30, 40).

Schmidt, A. (1917). *Erdmagnetismus*. Ed. by P. H. Furtwängler and E. Wiechert. Encyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen. B. G. Teubner Verlag, pp. 266–396 (cit. on p. 34).

Shore, R. M. et al. (2016). "Decadal period external magnetic field variations determined via eigenanalysis." In: *Journal of Geophysical Research: Space Physics* 121.6, pp. 5172–5184. DOI: 10.1002/2015JA022066 (cit. on p. 42).

Sugiura, M. (1964). "Hourly values of equatorial Dst for the IGY." In: *Annales of the International Geophysical Year* 35, pp. 4–45 (cit. on p. 26).

Thébault, E., K. Hemant, G. Hulot, and N. Olsen (2009). "On the geographical distribution of induced time-varying crustal magnetic fields." In: *Geophysical Research Letters* 36.1, pp. 1–5. DOI: 10.1029/2008GL036416 (cit. on p. 32).

Thébault, E. et al. (2015). "International Geomagnetic Reference Field: the 12th generation." In: *Earth, Planets and Space* 67.79. DOI: 10.1186/s40623-015-0228-9 (cit. on pp. 29, 48).

Thomson, A. W. P. and V. Lesur (2007). "An improved geomagnetic data selection algorithm for global geomagnetic field modelling." In: *Geophysical Journal International* 169.3, pp. 951–963. DOI: 10.1111/j.1365-246X.2007.03354.x (cit. on p. 30).

Verbanac, G., M. Mandea, M. Bandić, and S. Subašić (2015). "Magnetic observatories: biases over CHAMP satellite mission." In: *Journal of Geophysical Research: Solid Earth* 6.2, pp. 775–781. DOI: 10.5194/se-6-775-2015 (cit. on pp. 32, 33).

Verbanac, G., H. Lühr, M. Rother, M. Korte, and M. Mandea (2007). "Contributions of the external field to the observatory annual means and a proposal for their corrections." In: *Earth, Planets and Space* 59.4, pp. 251–257. DOI: 10.1186/BF03353102 (cit. on p. 28).

Wardinski, I. and R. Holme (2011). "Signal from noise in geomagnetic field modelling: denoising data for secular variation studies." In: *Geophysical Journal International* 185.2, pp. 653–662. DOI: 10.1111/j.1365-246X.2011.04988.x (cit. on p. 42).

Wolf, R. A. (1995). "Introduction to Space Physics." In: ed. by M. G. Kivelson and C. T. Russell. Cambridge University Press. Chap. Magnetopsheric Configurations, pp. 288–329. ISBN: 9780521457149 (cit. on p. 26).

APPENDIX A: COORDINATE SYSTEMS

Tab. 3.A1 summarizes the used coordinate systems and the nomenclature of the associated components in the position and field vectors.

Table 3.A1: Nomenclature of field and position vector components in the different coordinate systems used. *ned* abbreviates *n*orth, *e*ast, *d*own and *seu* abbreviates *s*outh, *e*ast, *u*p. Note that $\dot{X} = -B_\theta$, $Y = \dot{Y} = B_\phi$ and $\dot{Z} = -B_r$. Latitudes are labeled θ' while colatitudes are labeled θ.

	Position vector			Field vector		
System	Pos.1	Pos. 2	Pos. 3	Field 1	Field 2	Field 3
Spheroidal ned_d	θ'_d	ϕ	h	X	Y	Z
Spherical ned_c	θ'_c	ϕ	r	$\dot{X}$	$\dot{Y}$	$\dot{Z}$
Spherical seu_c	θ_c	ϕ	r	B_θ	B_ϕ	B_r
Cartesian GEO	x_G	y_G	z_G	X_G	Y_G	Z_G
Cartesian MAG	x_M	y_M	z_M	X_M	Y_M	Z_M
Spherical *mag*	θ_M	ϕ_M	r	$B_{\theta,M}$	$B_{\phi,M}$	$B_{r,M}$

APPENDIX B: DATA SOURCES

All sources were last accessed on September 10, 2017.

Data:

- OAMs – BGS (`http://www.geomag.bgs.ac.uk/data_service/data/annual_means.shtml`)

Activity indicators:

- Dst – NOAA (`ftp://ftp.ngdc.noaa.gov/STP/GEOMAGNETIC_DATA/INDICES/EST_IST`)
- Dcx – University of Oulu, Finland (`http://dcx.oulu.fi/dldatadefinite.php`)
- Ap – GFZ Potsdam via WDC Kyoto (`http://wdc.kugi.kyoto-u.ac.jp/kp/index.html`)
- RC – DTU Space (`http://www.spacecenter.dk/files/magnetic-models/RC/`), version from 12/06/17
- F10.7 – NOAA (`ftp://ftp.ngdc.noaa.gov/STP/space-weather/solar-data/solar-features/solar-radio/noontime-flux/penticton/penticton_absolute/listings`), "Series D Flux"
- Total sunspot number – SILSO (`http://www.sidc.be/silso/datafiles`)

Models:

- COV-OBS – DTU Space (`http://www.spacecenter.dk/files/magnetic-models/COV-OBS/`)
- GRIMM series – GFZ Potsdam (`https://www.gfz-potsdam.de/en/section/geomagnetism/data-products-services/geomagnetic-field-models/`), up to version 3.0
- CM4 – NASA (`https://denali.gsfc.nasa.gov/geomag.html`), not available anymore
- POMME-8 – CIRES (`http://geomag.org/models/pomme8.html`)

APPENDIX C: OBSERVATORY INFORMATION

Tab. 3.C1 lists observatory locations in geodetic and geomagnetic coordinates as well as our recalculated crustal biases.

Table 3.C1: Locations of candidate observatories and corresponding crustal biases. Locations in geomagnetic coordinates were calculated for mid 2010 by linear interpolation between coefficients from IGRF-12 (Thébault et al., 2015). OAMs for 1997–2010 were checked against annual means from OHMs revised by Macmillan and Olsen, (2013). Three OAMs were replaced by these corrected data (KNY X&Y at 2003.5 and ESA X at 2001.5). Observatories used to derive the AMC index have bold IAGA codes.

Geodetic coord. (ned_d)				Geomagnetic coord. (mag)			Crustal bias (ned_d)		
Code	lat [°] θ'_d	lon [°] ϕ	alt [km] h	lat [°] θ'_M	lon [°] ϕ_M	rad [km] r - 6371	Xb [nT]	Yb [nT]	Zb [nT]
HAD	51.0	-4.48	0.095	53.58	80.15	-5.64	-38.89	14.1	69.32
NGK	52.07	12.68	0.078	51.66	97.62	-6.05	-18.28	-3.1	-84.02
CLF	48.03	2.26	0.145	49.56	85.72	-4.49	-65.11	-15.79	101.0
FRD	38.21	-77.37	0.069	47.93	-6.02	-0.93	73.42	-50.88	116.84
SPT	39.55	-4.35	0.922	42.45	76.13	-0.57	9.86	9.09	-30.43
SJG	18.12	-66.15	0.424	27.9	6.57	5.51	-47.27	135.39	127.52
KAK	36.23	140.19	0.036	27.49	-150.73	-0.26	8.02	12.01	-100.3
LZH	36.09	103.85	1.56	25.97	176.48	1.32	23.05	0.43	-103.64
TAM	22.79	5.53	1.373	24.4	82.06	5.32	59.83	-220.96	-50.2
KNY	31.42	130.88	0.107	22.03	-158.76	1.47	5.91	53.01	-49.88
HON	21.32	-158.0	0.004	21.59	-89.66	4.33	-160.34	84.97	-331.6
MBO	14.38	-16.97	0.007	19.78	57.83	5.84	125.59	35.24	66.69
GUA	13.59	144.87	0.14	5.51	-143.88	6.11	135.82	76.39	41.64
PPT	-17.57	-149.57	0.357	-15.03	-74.5	5.56	-894.97	-1034.71	-390.18
HER	-34.43	19.23	0.026	-33.91	84.72	0.37	30.36	9.05	31.25
GNA	-31.78	115.95	0.060	-41.45	-170.68	1.30	-13.36	-100.61	115.75

APPENDIX D: ADDITIONAL FIGURES

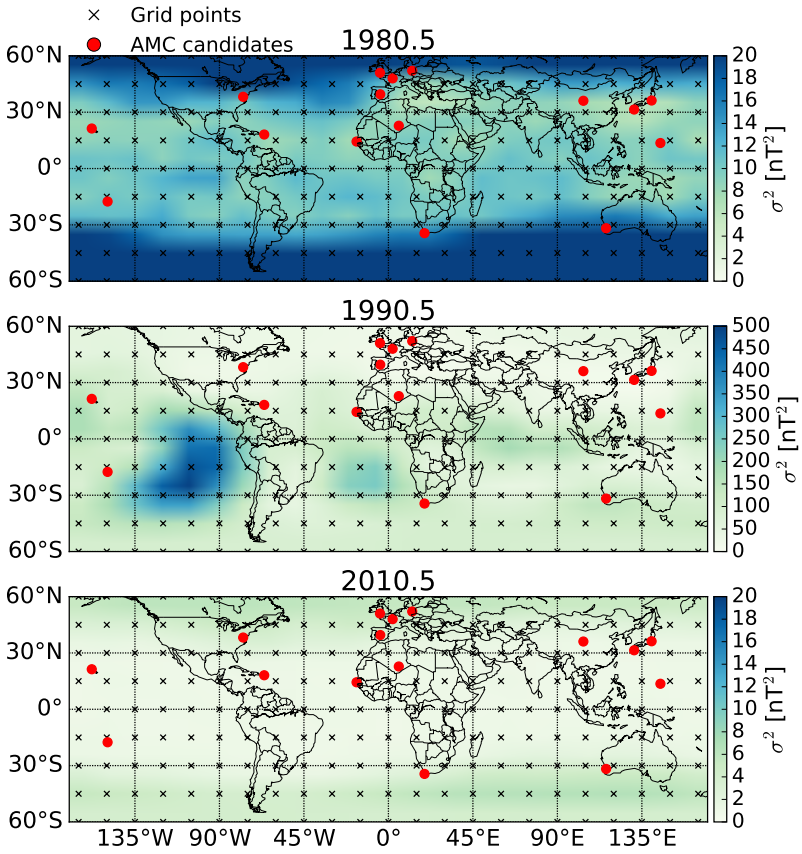

Figure 3.D1: Interpolated main field variances σ^2 on a 15°lat/15°lon grid (black crosses) from COV-OBS ensemble (100 members). For each location variances of all three vector components X, Y & Z were calculated and then averaged. Red dots indicate the locations of the 16 AMC candidate observatories. The chosen times 1980.5, 1990.5 & 2010.5 correspond to those indicated by black vertical lines in Fig. 3.6a). Note the different scale of the color bars for 1980.5/2010.5 & 1990.5.

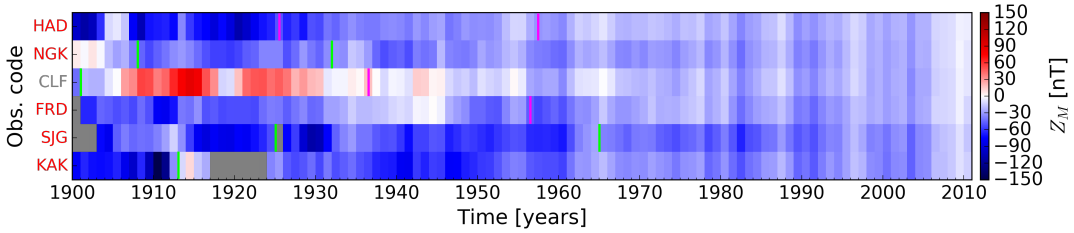

Figure 3.D2: Observatory residuals Z_M (north, geomagnetic MAG) from six observatories spanning 1900–2010 sorted by geomagnetic latitude from north (top) to south (bottom). Red IAGA codes indicate observatories chosen for the AMC data set. Vertical bars indicate relocations of the observatories that either offer an overlap of OAMs (magenta) or not (green). The consecutive IAGA codes are GRW→ABN→**HAD**, POT→SED→**NGK**, PSM→VLJ→**CLF**, CLH→**FRD**, VQS→SJG1→**SJG2**, TOK→**KAK**.

4

EVOLUTION OF LARGE-SCALE MAGNETIC FIELDS FROM NEAR-EARTH SPACE DURNG THE LAST 11 SOLAR CYCLES

Leonie Pick[1,2], Monika Korte[1], Yannik Thomas[1,2], Natalie Krivova[3], Chi-Ju Wu[3]

[1] Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany
[2] University of Potsdam, Am Neuen Palais 10, House 9, 14469 Potsdam, Germany
[3] Max Planck Institute for Solar System Research, Justus-von-Liebig-Weg 3, 37077 Göttingen

 DOI: 10.1029/2018JA026185

ABSTRACT We use hourly mean magnetic field measurements from 34 mid latitude geomagnetic observatories between 1900 and 2015 to investigate the long-term evolution and driving mechanism of the large-scale external magnetic field at ground. The Hourly Magnetospheric Currents (HMC) index is derived as a refinement of the annual version AMC (Pick and Korte, 2017, https://doi.org/10.1093/gji/ggx367). HMC requires an extensive revision of the observatory hourly means. It depends on three third party geomagnetic field models used to eliminate the core, the crustal and the ionospheric solar-quiet field contributions. We mitigate the dependency of HMC on the core field model by subtracting only non-dipolar components of the model from the data. The separation of the residual (dipolar) signal into internal and external (HMC) parts is the main methodological challenge. Observatory crustal biases are updated with respect to AMC and the solar-quiet field estimation is extended to the past based on reconstructed solar radio flux (F10.7). We find that HMC has more power at low frequencies (periods $\geq$ one year) than the Dcx index, especially at periods relevant to the solar cycle. Most of the slow variations in HMC can be explained by the open solar magnetic flux. There is a weakly decreasing linear trend in absolute HMC from 1900 to present, which depends sensitively on the data rejection criteria at early years. HMC is well suited for studying long-term variations of the geomagnetic field.

4.1 INTRODUCTION

Models of Earth's internally generated magnetic field have experienced considerable progress in the last two decades when high spatial resolution satellite data from dedicated missions (Ørstedt, CHAMP and Swarm) became available to complement the geomagnetic observatory record. This advance has now stalled due to the difficulties in properly accounting for external fields, i.e. magnetic fields originating from outside of Earth (Finlay et al., 2017). On the timescale of the "historical era" ($\sim$1840–now) this applies to large-scale fields that fluctuate at rates comparable to relatively rapid, sub-decadal secular variation (SV) of the internal field (e.g. Gillet et al., 2010). These fields originate from electrical current systems in the inner and outer magnetosphere (recent reviews by Ganushkina et al., 2018; Milan et al., 2017).

Magnetospheric fields are often parameterized in spherical harmonic based global field models using time variations from "geomagnetic activity" indices, derived from ground observatory data (Kauristie et al., 2017). For signals from the inner magnetosphere the parameter traditionally used is the Dst index (Sugiura, 1964) or its corrected and extended version Dcx (Mursula and Karinen, 2005). However, it is well known that Dst (Dcx) is inappropriate for long-term modeling of the geomagnetic field, because its instable baseline (e.g. Olsen et al., 2005; Temerin and Li, 2015) causes an underestimation of the solar cycle variation in magnetospheric magnetic fields. In their magnetospheric field model, Lühr and Maus, (2010) compensate for this effect using the 81-day running mean of the 10.7 cm solar radio flux (F10.7), lagged by $\sim$20 months. In the context of this study it is important to understand the connection between solar magnetic and geomagnetic activity. The solar magnetic field emerges at the Sun's surface in form of active and smaller ephemeral magnetic regions (Harvey, 1994; Krivova and Solanki, 2004). Active regions dominate the variability on time scales up to a solar cycle ($\sim$11 years), whereas ephemeral regions are responsible for longer-term variations (Harvey, 1992; Solanki et al., 2000, 2002). Both sunspots (SN is the total sunspot number) and F10.7 originate from active regions and are good proxies of solar magnetic activity in general (e.g. Tapping, 2013). However, they do not represent well the longer-term changes in the solar magnetic field. Furthermore, SN and F10.7 have no physical relation with the variability of magnetospheric magnetic fields, measured by geomagnetic activity indices like Dst. Geomagnetic activity ultimately results from the interaction of the near-Earth Interplanetary Magnetic Field (IMF) with Earth's magnetic field. The source of the IMF is the so-called open solar flux (OSF), which is the part of the total solar magnetic field that leaves the Sun and enters the heliosphere in radial direction (e.g. Lockwood, 2013, details in section 4.2.2.2). This part includes contributions from both active and ephemeral regions. Consequently, OSF physically links solar and geomagnetic activity, while being able to capture the long-term variability (> 11 years) in the solar magnetic field.
The mentioned shortcomings of Dst motivated the creation of index alternatives for specific satellite-based model series, most importantly the VMD index (Thomson and Lesur, 2007) for the GRIMM series (most recent version by Lesur et al., 2015) and the scalar RC index (Olsen et al., 2014) for the CHAOS series (most recent

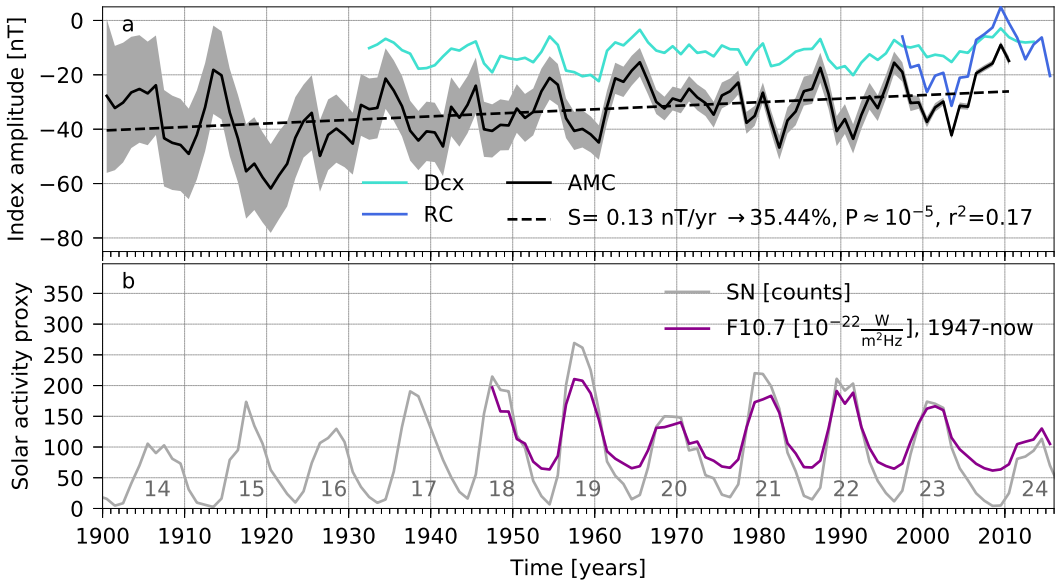

Figure 4.1: **a:** AMC $\pm 3\sigma$ (black & gray; Pick and Korte, 2017) with linear fit (dashed) in comparison to annual means of Dcx (turquoise) and RC (blue). S, P, and r^2 are the slope, the P-value for testing the null hypothesis, and the coefficient of determination. **b:** Annual means of SN (gray) and absolute F10.7 flux (magenta) as proxies for solar magnetic activity. Solar cycles are numbered in gray.

version by Finlay et al., 2016). Such efforts were expanded to the historical era by Pick and Korte, (2017), who defined the scalar Annual Magnetospheric Currents index (AMC) covering the years 1900 to 2010 (Fig. 4.1). While VMD and RC are developed to address the particular needs of their corresponding models, AMC is a stand-alone proxy for the absolute strength of magnetospheric currents at ground including uncertainty estimates.

AMC is particularly useful to reduce the contribution of magnetospheric disturbances in geomagnetic observatory data for studies of SV. However, a higher index cadence is needed to allow adequate use in combination with e.g. investigations of geomagnetic jerks (monthly resolution) or other standard geomagnetic field models and indices (hourly resolution). This is why we move from the AMC to the HMC (Hourly Magnetospheric Currents) index. Wardinski and Holme, (2011) and subsequently Brown et al., (2013) developed more elaborate methods to "denoise" observatory monthly means, but the advantages of the method are only fully brought to bear if a long-term high-quality reference station is available in the close vicinity of the observatory whose data are to be denoised. Going back in time such data become increasingly scarce which is when HMC can contribute. Furthermore, HMC can help to address two open questions with respect to AMC that we discuss in this study:

Q1 A linear fit to AMC reveals a statistically significant decrease (absolute values) of 0.13 nT/yr or 35 % with respect to the level at 1900 (Fig. 4.1a, dashed line). *Is the trend real, i.e. did the magnetic effect of magnetospheric currents on Earth get weaker over time?*

Q2 A validation of [A/H]MC solar cycle dependent variability by a comparison to SN or F10.7 (Fig. 4.1b) is unsatisfactory, as explained above. Given the established

link between OSF and geomagnetic activity, *can the variability in [A/H]MC be traced back to that in OSF?*

4.2 DATA

4.2.1 *Geomagnetic data used for HMC*

The HMC index is based on hourly means of magnetic field measurements from 34 observatories obtained from WDC Edinburgh with observatory locations made available by the IAGA. We refer to specific observatories by their three-letter IAGA codes (`www.bgs.ac.uk/iaga/vobs`).
For the best possible retrieval of magnetospheric current signal the observatories should cover different magnetic local times (MLT), while avoiding high latitude disturbances (i.e. ionospheric auroral currents) and the ionospheric equatorial electrojet flowing along the magnetic equator on the dayside. Also, both hemispheres should be covered by data in order to mitigate hemispheric differences in the signal (e.g. Gannon, 2012). This places potential observatories in two mid-latitudinal regions $10^\circ \leq |\theta'_{\mathrm{M}}| \leq 50^\circ$ (Fig. 4.2, gray shaded), with dipole latitude θ'_{M} in the geomagnetic coordinate system (*mag*, definition in Russell, 1971). Within these boundaries we identified the longest time series from all observatories operating during the time span in question. These include composite time series from observatory clusters. On the one hand, such clusters consist of IAGA-designated predecessor-successor stations, e.g. Cheltenham (CLH, Canada) and Fredericksburg (FRD, Canada). On the other hand, clusters are formed by observatories that are separated less than 300 km but are not officially replacing each other, e.g. Wien Auhof (WIA, Germany) and Hurbanovo (HRB, Slovakia). The search radius of 150 km was selected as a typical upper bound between official predecessor-successor observatories (CLH and FRD are exceptionally far apart).
The observatories and clusters finally chosen to contribute to HMC are labeled in Fig. 2. Caution is needed in case of the observatory M'Bour (MBO, Senegal) that is located within the realm of the South Atlantic Anomaly where the geomagnetic equator $\theta'_{\mathrm{M}} = 0$ (gray line) deviates significantly from the "true" magnetic equator $\theta'_{\mathrm{true}} = 0$ (black lines), calculated from IGRF-12 (Thébault et al., 2015). We verified, that the characteristic features of HMC are not particularly sensitive to data from MBO (section 4.5.1).
An overview of the selected data is given in Fig. 4.3. For this diagnostic plot we subtracted the temporally varying core field contribution (from model COV-OBS.x1; Gillet et al., 2015) and static lithospheric biases (section 4.3.1) from the time series to facilitate the detection of spikes, jumps and drifts by visual inspection. Spikes are single standout values, which we corrected if caused by obvious typos or else set to NAN. We eliminated sudden steps in the time series ("jumps", e.g. Fig. 4.3a, SVD) by joining the lower and upper edges of the step. Lastly, we identified times of smooth and gradual changes, spanning months to several years, that are unique to one particular observatory ("drifts", e.g. Fig. 4.3b, SSH). Such structures cannot be produced by magnetospheric sources and are consequently eliminated by subtracting either a Polynomial or a Fourier fit to the data segment in question.

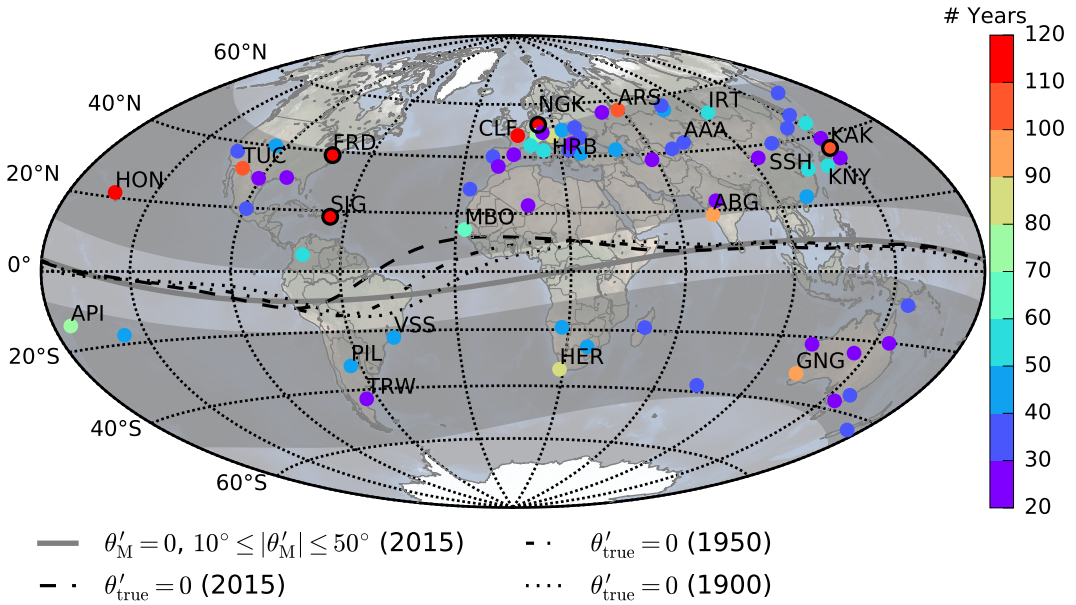

Figure 4.2: The 70 observatories and observatory clusters with the longest time series during 1900–2015 (color coded, max. 116 years). For clusters only the most recently active observatory is shown. HMC observatories (labeled) fall within 50° of *mag* latitude (θ'_M, gray shaded), excluding the geomagnetic equator (gray line for 2015). The position of the true magnetic equator ($\theta'_{\mathrm{true}} = 0$) is calculated from IGRF-12 for 1900, 1950 and 2015 (black lines). AMC observatories are framed in black for reference.

Extensive modifications have been carried out particularly on the Z components from observatories VLJ, EKT, CLH, TOK, VQS & VSS at early years. There are cases where the revised data still show characteristics not seen elsewhere in the same time series nor in any of the other observatories (dashed boxes in Fig. 4.3). The influence of these segments on HMC as well as the drifts is discussed in section 4.5.1.

Finally, core and lithospheric contributions were again added to the altered residua to give a revised set of hourly means. Note that our data processing is tailored to isolate large-scale external field contributions. Signatures of either residual SV or artificial drifts cannot be differentiated and are both removed from the time series. Consequently, our revised set of hourly means is not appropriate for investigations of the internal geomagnetic field.

4.2.2 *Proxies of solar magnetic activity*

4.2.2.1 *The 10.7 cm solar radio flux*

The F10.7 flux is widely used in the field of geomagnetism to describe the influence of solar activity on the ionospheric magnetic field (Olsen, 1993). It is measured consistently on a day-to-day basis in Canada since February 1947, first by a radio telescope near Ottawa and then at the Dominion Astrophysical Observatory near Penticton. We have extended the record back to 1610 using the SATIRE-T model

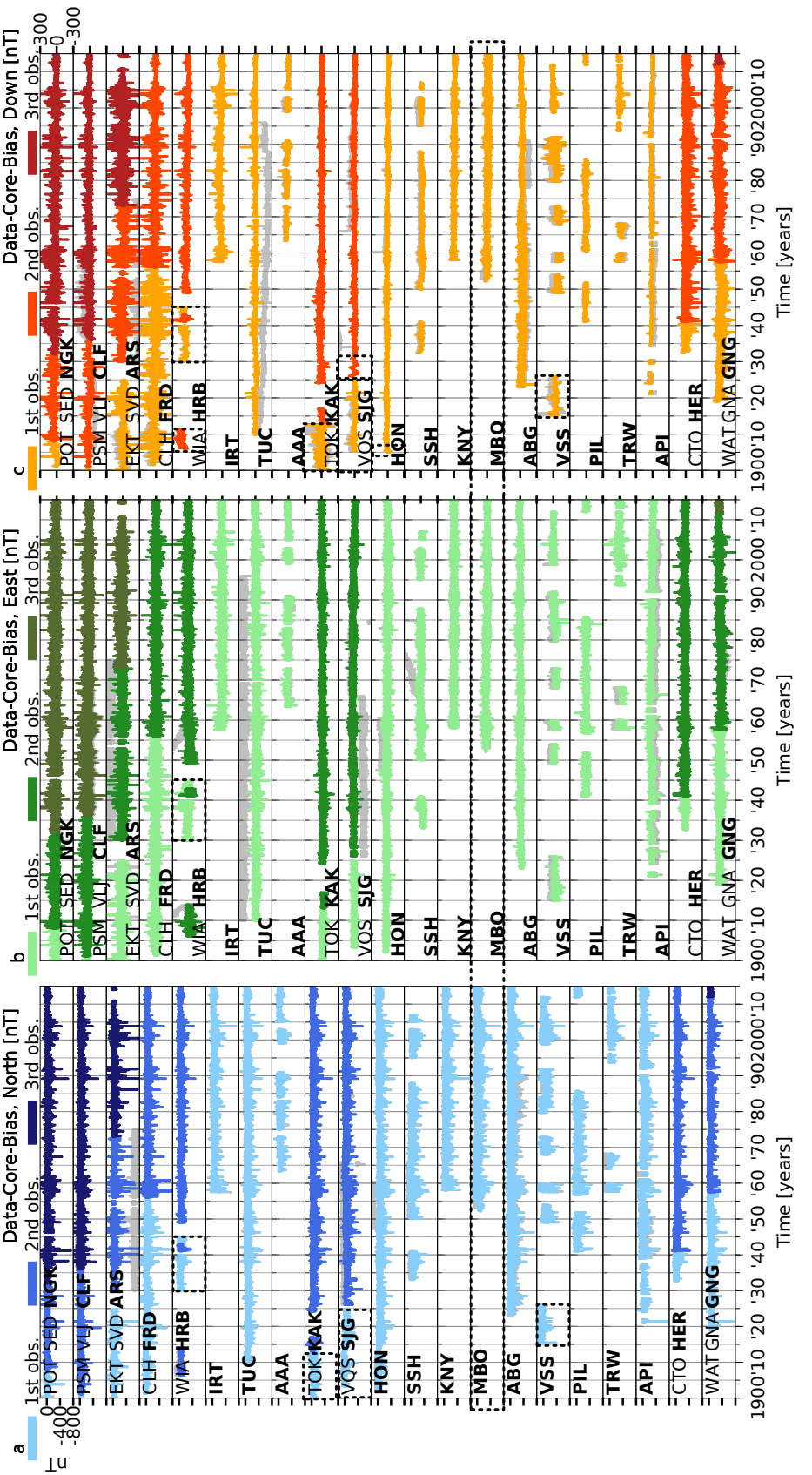

Figure 4.3: Residual magnetic field measurements (hourly) from observatories contributing to HMC. Local geodetic vector components point toward north (**a**, -800 to 200 nT), toward east (**b**, -300 to 300 nT), and vertically down (**c**, -300 to 300 nT). Time series from observatory clusters are color coded from light (earliest station) to dark (present station). Residua without modifications are shown in gray. Dashed boxes contain data segments of uncertain quality.

(Krivova et al., 2010; Wu et al., 2018). This model employs observations of sunspot numbers (since 1610) and areas (since 1876) to reconstruct the past evolution of the solar surface magnetic field as well as changes in the solar total and spectral irradiance. Past changes in F10.7 are then reconstructed through a regression analysis between the solar surface magnetic field and the measured F10.7 flux.

4.2.2.2 *The open solar magnetic flux*

The OSF is formed by that part of the solar magnetic field, which is blown out into the heliosphere where it threads a “coronal source surface", usually taken to be a heliocentric sphere at r_S = 2.5$R_\odot$ ($R_\odot$ is mean solar radius). The “signed" OSF, i.e. the flux of one radial field polarity, can be calculated as follows:

$$\mathrm{OSF}(r_S) = 2\pi r_M^2 < |B_r(r_M)| >_{27\ \mathrm{d}} - E. \tag{4.1}$$

B_r is the radial component of the IMF pointing away from the Sun, equivalent to $-B_x$ in the Geocentric Solar Ecliptic (GSE) frame. Averaging of $|B_r|$ over 27 days (one Bartels rotation) eliminates variations with coronal longitude. The excess flux E accounts for magnetic field contributions generated between the source surface (r_S) and the measurement point (r_M), which do not reflect the coronal source field. It is calculated from the longitudinal IMF component and the solar wind speed (v_{sw}) following Lockwood et al., (2009a,b) to give the “kinematically corrected" OSF (Eq. 1). All required quantities are taken from the OMNI-2 low resolution (hourly-averaged) data set, comprising measurements from near-Earth satellites at $r_M = 1$ AU (currently Wind, ACE) since November 1963. Due to data gaps in early years, a continuous record of the 27-day averaged kinematically corrected OSF starts in the beginning of 1974.
OSF is linked to HMC as follows (see review by Lockwood, 2013):

$$\mathrm{OSF} \propto |B_r| \xrightarrow[\text{theory}]{\text{1. Parker spiral}} B, v_{sw} \xrightarrow[\text{coupling}]{\text{2. sw–magnetosphere}} \mathrm{HMC} \tag{4.2}$$

B_r is connected to the magnitude of the IMF (B) and v_{sw} by Parker’s spiral theory (Parker, 1958, 1963). The product $B \cdot v_{sw}^n$ quantifies the coupling between the solar wind and the magnetosphere (e.g. Finch and Lockwood, 2007), whereby an influx of solar wind particles alters magnetospheric current systems and triggers a response in HMC. Since the sensitivity to magnetospheric current systems varies among different indices, their correlations with $B \cdot v_{sw}^n$ peak for different n (Lockwood et al., 2013b). Relation 2 is valid for (a) an averaging time greater than one year (Stamper et al., 1999) and (b) a homogeneous response of the geomagnetic data to v_{sw}, i.e. a constant optimum n (Lockwood et al., 2013a).
Lockwood et al., (2014) reconstruct OSF back to 1845 with an annual resolution in three steps: First, B and v_{sw} are reconstructed from the geomagnetic indices IDV(1d) (Lockwood et al., 2013a) and a corrected version of aa (aa_c), using connection 2 of relation 2 above. Secondly, $|B_r|$ is calculated using connection 1. Lastly, $|B_r|$ is regressed against the observed kinematically corrected OSF (Eq. 4.1) to give the kinematically corrected, reconstructed OSF (see Supporting Information, Text S3).

4.3 METHODOLOGY

The modeling strategy follows the derivation scheme of AMC, which assumes the field produced by magnetospheric currents to be large-scale, i.e. being represented by spherical harmonic (SH) degree N=1, and axisymmetric with respect to the dipole axis. We use all local times, including daytime observatories, and the full vector, including the vertical downward component (Z). This is because we prioritize the use of all available data carrying the signal in question over the avoidance of ionospheric fields (during daytime) and the induced signal (Z), which are globally heterogeneous. Furthermore, a proper treatment of induced signal requires an appropriate conductivity model, possibly a 3-D model to account for the "ocean effect" (Grayver et al., 2017), and we would like to leave this choice to the user of HMC.

4.3.1 *Processing*

First, we calculate residua $\vec{\mathrm{B}}_i(t)$ for all observatory time series, $i = 1, \ldots, N_o$ (max. N_o is 21), by subtracting estimates for the core field, $\vec{\mathrm{Bc}}_i(t)$, the mid-latitudinal ionospheric field, $\vec{\mathrm{Bi}}_i(t)$, and the static, high-degree crustal field, $\vec{\mathrm{Bb}}_i$, from the revised observatory data $\vec{\mathrm{Bo}}_i(t)$:

$$\vec{\mathrm{B}}_i(t) = \vec{\mathrm{Bo}}_i(t) - \left(\vec{\mathrm{Bc}}_{\mathrm{ND},i}(t) + \vec{\mathrm{Bi}}_i(t) + \vec{\mathrm{Bb}}_i\right) \tag{4.3}$$

$$\vec{\mathrm{Bb}}_i = \langle \vec{\mathrm{Bo}}_i(t) - \left(\vec{\mathrm{Bc}}_i(t) + \vec{\mathrm{Bi}}_i(t) + \vec{\mathrm{Be}}_i(t)\right) \rangle_{[1999,\ldots,2015],\ \text{quiet \& dark}} \tag{4.4}$$

For the core field, the choice of COV-OBS.x1 is practically without alternative as there is no other published core field model covering the historical era up to the present. HMC is particularly sensitive to the dipole component ($N = 1$) of this core field model. We therefore do not subtract the complete model but only the non-dipole (ND) model part corresponding to SH degrees $N = 2, \ldots, 14$, $\vec{\mathrm{Bc}}_{\mathrm{ND},i}(t)$. This way we do not need to assume that COV-OBS's internal dipole component is correct and we can perform the separation of signal into internal and external contributions ourselves.

As we include daytime data we need to remove as far as possible the mid latitude ionospheric field $\vec{\mathrm{Bi}}_i(t)$ peaking around local noon when photoionization by the Sun maximizes. This field causes the so-called solar-quiet (Sq) variations (e.g. Yamazaki and Maute, 2017) and we consider it in a climatological sense by the CM4 model (Sabaka et al., 2004). CM4 can be evaluated outside of the validity interval of the corresponding core field description (1960–2002) with appropriate values for the monthly mean F10.7 flux (three month moving averages). To this end we supplement the record of measured F10.7 flux (starting in 1947) with reconstructed F10.7 flux starting in 1900 (section 4.2.2.1). An example of Sq variations in the northward component (X) at Apia observatory (API, Western Samoa) for 15 days in September 2014 is shown in Fig. 4.4a before (gray dashed) and after (gray solid) CM4 has been subtracted. The amplitude of the variation clearly reduces, although CM4 can not handle the day-to-day variability.

For the crustal biases, $\vec{\mathrm{Bb}}_i$, we use the CHAOS model (version six, Finlay et al., 2016)

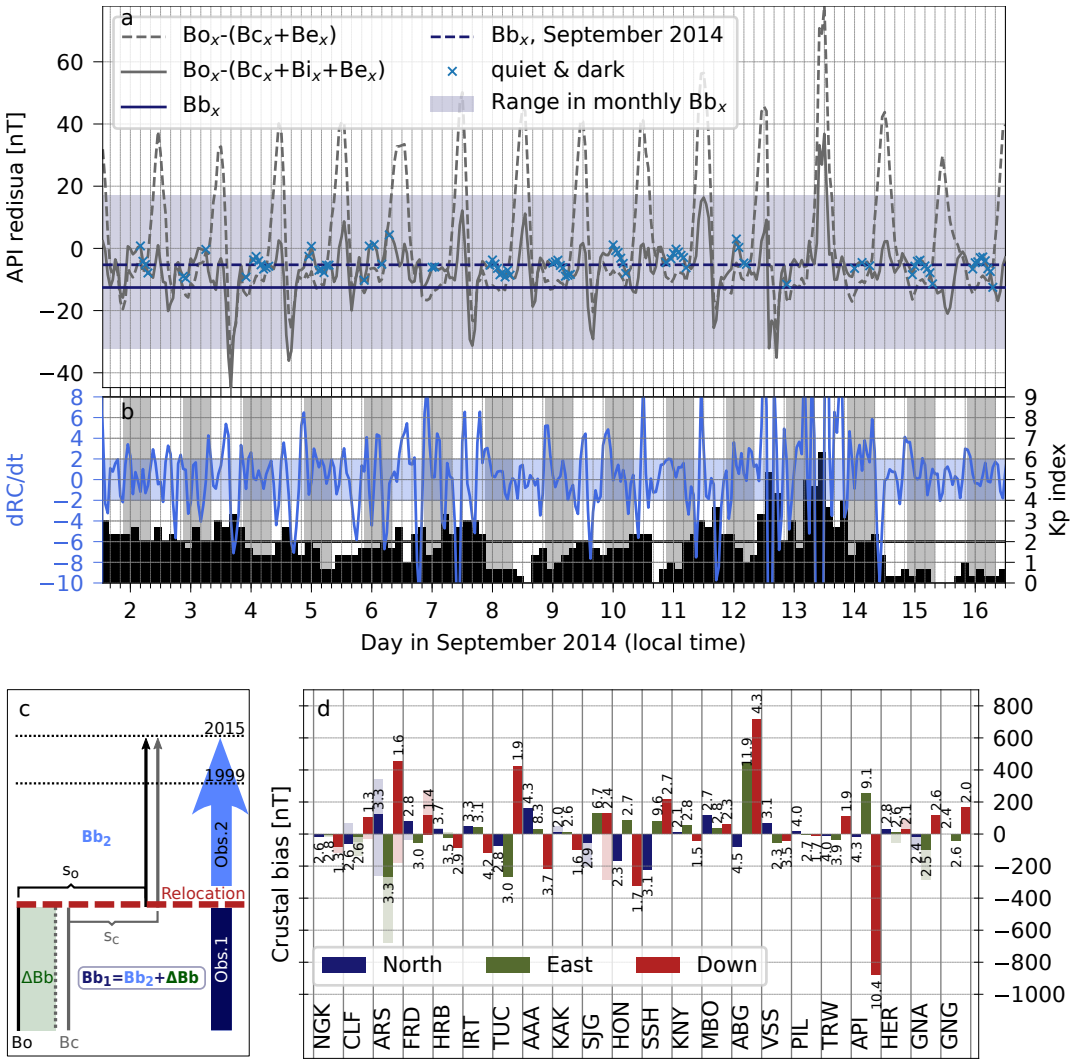

Figure 4.4: **a:** Measurement residua and crustal bias at API. Residua before (gray dashed) and after (gray solid) CM4 was removed with data points for crustal bias calculation (blue crosses). Final bias (blue solid) and short-term bias (blue dashed) fall within a 49.2 nT wide range (blue shaded) spanned by all short-term biases. **b:** Quiet ($|\mathrm{dRC/dt}| \leq 2.0$, blue; $\mathrm{Kp} \leq 2^0$, black) and dark (solar elevation angle $\leq -10°$, gray) selection criteria for crustal bias data. **c:** Sketch illustrating the calculation of the crustal bias ($\vec{\mathrm{Bb}}_1$) for an observatory (Obs. 1), that was relocated prior to 1999 (thereafter Obs. 2) from differences in measurements ($\mathrm{s_o}$) and core field estimates ($\mathrm{s_c}$) at the two sites (see text and Eq. 4.5). **d:** Crustal biases (bars) with median absolute deviations (small labels), see Table 4.S1 of the Supporting Information. Biases for predecessors are shown in faded colors.

to estimate the recent internal ($\vec{\mathrm{Bc}}_i(t)$) and external magnetospheric ($\vec{\mathrm{Be}}_i(t)$) field contributions in Equation 4.4. Considering $\vec{\mathrm{Be}}_i(t)$ ensures a correct baseline level of HMC (see Pick and Korte, 2017, section 3.3.1). We apply the same criteria for selecting geomagnetically quiet ($\mathrm{Kp} \leq 2^0$, $|\mathrm{dRC/dt}| \leq 2.0$) and dark (Sun at least 10° below horizon) periods as are applied in CHAOS-6 for vector field data at mid latitudes. The bias vector is calculated for each observatory from the median of the selected data points within the period 1999–2015. This procedure is illustrated

for the X component by the example of API in Fig. 4.4a,b. If the northward bias had instead been calculated from the first 15 days of each month within the period 1999–2015 (“short-term" bias), there would be a $\sim$50 nT wide spread in the results for API representing the dependency on season and solar cycle phase. API is the observatory most sensitive to these modulations (across all components) which is why we show it as “worst case" in terms of bias estimation. However, when considering the whole time span, 1999–2015, the median absolute deviation (MAD) of biases for all observatories is about an order of magnitude smaller, i.e. $\leq$ 4.5 nT for X, $\leq$ 11.9 nT for Y and $\leq$ 10.4 nT for Z (Fig. 4.4d; Supporting Information, Tab. 4.S1). We conclude, that our estimation of crustal biases is robust and justify the use of the CHAOS-6 core field estimate by a slightly smaller cumulative MAD as compared to the COV-OBS.x1 estimate.

A problem arises for those observatories in clusters that are not active during 1999–2015. In such cases we follow a procedure illustrated in Fig. 4.4c. For official predecessor - successor observatories the step in measurements ($\vec{\mathrm{s}}_\mathrm{o}$) between observatory 1 ($\vec{\mathrm{Bb}}_1$ not known) and observatory 2 ($\vec{\mathrm{Bb}}_2$ known) at the time of the relocation is documented in the observatory annual means file provided by the British Geological Survey. In the case of WIA and HRB we calculated $\vec{\mathrm{s}}_\mathrm{o}$ directly from the annual mean data for 1941 and check the influence of this alternative method in section 4.5.1. Since the external fields should not vary significantly between the two locations, the step can be assumed to result from a difference in the core and crustal field (crustal bias) contributions. We estimate the difference in the core field contribution with COV-OBS.x1 ($\vec{\mathrm{s}}_\mathrm{c}$) and find the crustal bias of observatory 1 as:

$$\vec{\mathrm{Bb}}_1 = \vec{\mathrm{Bb}}_2 + (\vec{\mathrm{s}}_\mathrm{o} - \vec{\mathrm{s}}_\mathrm{c}) = \vec{\mathrm{Bb}}_2 + \Delta\mathrm{Bb} \tag{4.5}$$

In a last processing step the observatory residua $\vec{B}_i$ are transformed from the measurement coordinate system (spheroidal *ned*: latitude-θ'_d, longitude-ϕ, height-h & North-X, East-Y, Down-Z) into a geocentric system (spherical *seu*: colatitude-θ_c, longitude-ϕ, radius-r & South-B_θ, East-B_ϕ, Up-B_r). All coordinate system transformations involved follow the nomenclature detailed in both Equation 3.3 and Table 3.A1 of Pick and Korte, (2017).

4.3.2 *Modeling*

Having calculated the residua time series for each observatory, we combine them in one data vector $\vec{\mathrm{B}}$ per universal time hour (in the *seu* system). We assume the residua to arise from a Laplacian vector field that is expressed as the gradient of a scalar potential V:

$$\vec{\mathrm{B}} = -\nabla \mathrm{V} \tag{4.6}$$

The potential is expanded in spherical harmonics up to degree (N) and order (m) one and composed of an internal ($\sim (a/r)^{n+1}$) and an external ($\sim (r/a)^n$) part according to the radial distance of the observation from the reference level at the Earth's surface (a =6371.2 km):

$$\mathrm{V} = V_\mathrm{int} + V_\mathrm{ext} = \Re\left\{ a \sum_{n=1}^{N} \sum_{m=0}^{n} \left[\iota_n^m \left(\frac{a}{r}\right)^{n+1} + \epsilon_n^m \left(\frac{r}{a}\right)^n \right] P_n^m(\cos\theta_c) e^{im\phi} \right\} \tag{4.7}$$

P_n^m are the Schmidt quasi-normalized associated Legendre polynomials and ι_n^m, ϵ_n^m are the internal and external Gauss coefficients:

$$\iota_n^m = g_n^m - ih_n^m, \quad \epsilon_n^m = q_n^m - is_n^m \tag{4.8}$$

The forward problem is

$$\mathbf{G}\vec{\mathrm{x}} + \vec{\epsilon} = \vec{\mathrm{B}} \tag{4.9}$$

where matrix $\mathbf{G}$ relates the observations $\vec{\mathrm{B}}$, including uncertainties $\vec{\epsilon}$, to the model coefficients $\vec{\mathrm{x}}$. After the processing described by Equation 4.3, we expect that the residua result from an N=1 internal field overlain by an external field contribution best described in geomagnetic coordinates (*mag*: θ_M, ϕ_M, r_M & $B_{\theta,M}$, $B_{\phi,M}$, $B_{r,M}$). Thus, the coefficients vector consists of the three internal dipole coefficients (*seu* system) and the external axial dipole coefficient (*mag* system):

$$\vec{\mathrm{x}} = \left(g_1^0, g_1^1, h_1^1, q_{1M}^0\right)^T \tag{4.10}$$

$$\mathrm{HMC} := -q_{1M}^0 \tag{4.11}$$

The latter is what we define as HMC index, where the minus sign indicates that it is a southward directed field.
The corresponding matrix $\mathbf{G}$ has dimensions ($3N_o \times 4$, N_o: number of observatories) with the following entries:

$$\mathbf{G} = \begin{pmatrix} -\left(\frac{a}{r}\right)^3 \frac{dP_1^0}{d\theta} & -\left(\frac{a}{r}\right)^3 \cos(\phi) \frac{dP_1^1}{d\theta} & -\left(\frac{a}{r}\right)^3 \sin(\phi) \frac{dP_1^1}{d\theta} & -\frac{dP_1^0}{d\theta} \\ 0 & \left(\frac{a}{r}\right)^3 \sin(\phi) \frac{P_1^1}{\sin\theta} & -\left(\frac{a}{r}\right)^3 \cos(\phi) \frac{P_1^1}{\sin\theta} & 0 \\ \left(\frac{a}{r}\right)^3 2P_1^0 & \left(\frac{a}{r}\right)^3 2\cos(\phi) P_1^1 & \left(\frac{a}{r}\right)^3 2\sin(\phi) P_1^1 & -P_1^0 \\ \underbrace{\vdots}_{\frac{d}{dg_1^0}, seu} & \underbrace{\vdots}_{\frac{d}{dg_1^1}, seu} & \underbrace{\vdots}_{\frac{d}{dh_1^1}, seu} & \underbrace{\vdots}_{\frac{d}{dq_{1M}^0}, mag \to seu} \end{pmatrix} \tag{4.12}$$

In order to co-estimate the internal and external Gauss coefficients, the column of $\mathbf{G}$ which relates to the external coefficient (column 4) has to be estimated using *mag* coordinates for the observatory positions. Subsequently, the vectors corresponding to each observatory i (rows 1–3,4–6,...) have to be transferred back to the *seu* system. We invert the problem in Equation 4.9 using a simple iterative least squares fit (details in the Supporting Information, Text S1):

$$\hat{\vec{\mathrm{x}}}_{k+1} = \left(\mathbf{G}(\hat{\vec{x}}_k)^T \mathbf{G}(\hat{\vec{x}}_k)\right)^{-1} \mathbf{G}^T(\hat{\vec{x}}_k)\vec{\mathrm{B}} \tag{4.13}$$

We do not treat the uncertainties of the observatory residua $\vec{\epsilon}$ here. For AMC we have assumed that data uncertainties are normally distributed and simply weighted matrix $\mathbf{G}$ with the residua covariance matrix at each inversion time step. For HMC this procedure comes at a high computational cost given that the inversion is performed at each UT hour (1.016.832 time steps). This effort is unnecessary provided the long-term uncertainty characteristics remain unchanged when stepping down to hourly means (i.e. AMC uncertainties apply) or the method to calculate index uncertainties is not profoundly revised.

4.4 RESULTS

The questions **Q1** & **Q2** target long-term characteristics of HMC, i.e. low-frequency variations in HMC. We analyze HMC's spectral content in terms of the power spectral density (PSD) and use the magnitude squared coherence to compare HMC to Dcx, RC and q^0_{1M} from COV-OBS.x1 (section 4.4.1). Low-pass filtered versions of HMC and the other quantities are then investigated in the time domain (section 4.4.2). Text S2 of the Supporting Information provides information on the internal dipole coefficients.

4.4.1 *HMC in the frequency domain*

The PSDs of the different magnetospheric field measures (Fig. 4.5a) are calculated from the maximum available time series, so that frequency bins for HMC (since 1900) are most narrow, followed by Dcx (since 1933) and RC (since 1997). The shape of the PSDs for the different proxies are similar with known peaks, prominently at one, nine and ∼27 days as well as half a sidereal year (365.256 days). All PSDs show a broad maximum centered approximately at half a solar cycle. HMC shows greater power at periods ≥ one year than the other quantities. Thus, we fix the the cutoff frequency for the low-pass filter at $1/(365.256\ \text{days}) \sim 3.2 \cdot 10^{-8}$ Hz.
Dcx shows high power between ∼9 days and one month, but the power decreases notably at longer periods relative to the other measures. This confirms that Dcx lacks long-term variability as a consequence of its erroneous baseline (section 4.1). RC is the most variable, showing multiple peaks, that are either absent or very small for the other quantities, especially at periods below 1 day and greater than ∼50 days. This could be connected to the use of only night side data. q^0_{1M} COV-OBS is part of a model parameterized in time by cubic B-splines with a two year knot spacing. Therefore, its energy content is meaningful for periods ≥ two years and becomes comparable to that of the other indices not until periods ≥ three years.
The coherence analysis (Fig. 4.5b) reveals a very high degree of similarity (close to one) in the frequency contents of HMC and RC, except for some notches caused by RC variability not present in HMC (most prominently at a one year period). The coherence for HMC and Dcx exhibits a plateau-like shape with large values for periods between few days and a month and a rapid decrease to below 0.5 for larger periods. Due to the different temporal resolutions, the coherence of HMC and q^0_{1M} COV-OBS does not get larger than ∼0.75 at periods relevant to the solar cycle.

4.4.2 *HMC in the time domain*

We now compare the low-pass filtered HMC to the other indices and q^0_{1M} COV-OBS in the time domain. Regarding the long-term trend (**Q1**) we find, that much of the linear decrease in absolute AMC (35 %) is gone leaving only a small, albeit significant, decrease in absolute HMC of 8.41 % (Fig. 4.5c). We attribute the weakening of the trend to the data modifications performed and discuss the sensitivity of HMC to these alterations as well as the robustness of the long-term trend in section 4.5.1. HMC is generally more negative than RC and Dcx, with a recent absolute minimum

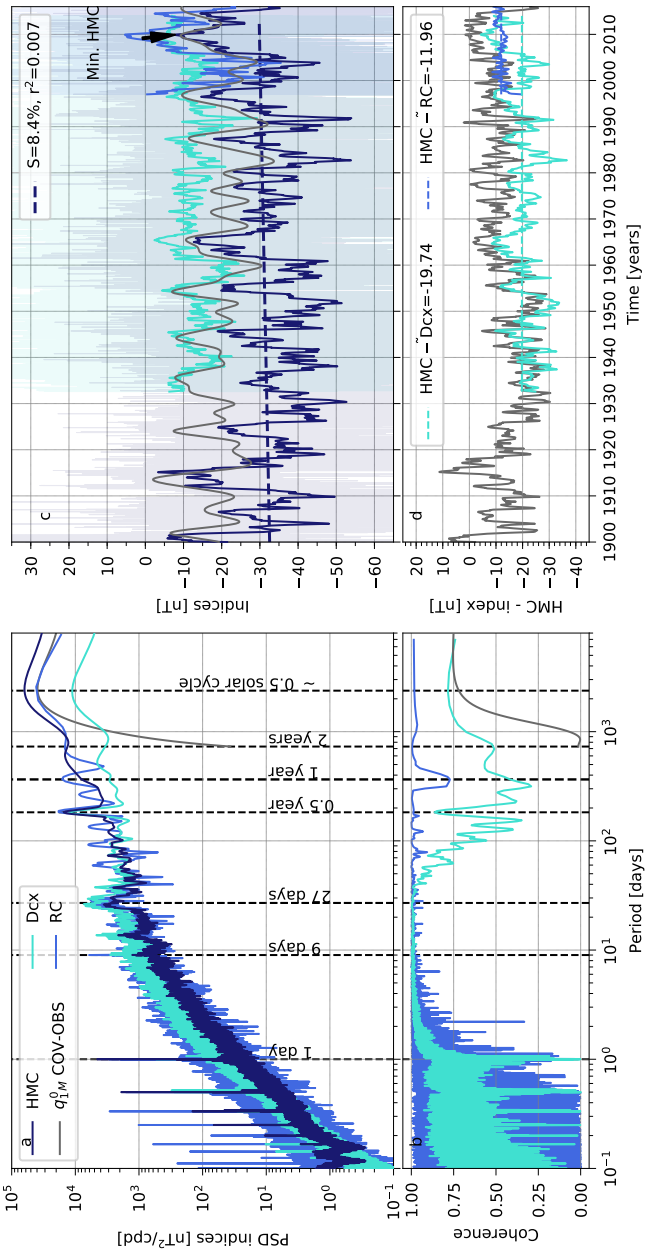

Figure 4.5: **a:** Power spectral densities (PSD) of HMC, Dcx, RC and q^0_{1M} COV-OBS ($\geq$ 2 years). **b:** Magnitude squared coherence of HMC and either Dcx, RC or q^0_{1M} COV-OBS ($\geq$ 2 years). **c:** Low-pass filtered (cutoff 1 year) HMC with linear fit (dashed), Dcx, RC and q^0_{1M} COV-OBS atop the corresponding hourly values (faded). HMC minimum (absolute value) of -7.78 nT at 2009 (arrow). **d:** Differences between HMC and either Dcx, RC or q^0_{1M} COV-OBS together with the median of the differences (dashed).

of -7.78 nT reached on December 3, 2009 at the beginning of solar cycle 24 (Fig. 4.5c). The smallest absolute AMC is only $\sim$1 nT larger that year, although the underlying observatory annual means include geomagnetically disturbed times. This means, that occurrences of geomagnetic storms do not effectively alter the magnetospheric signal's background level. The differences between HMC and the other quantities are shown in Fig. 4.5d. The median difference between RC and HMC is -11.96 nT, which agrees with the offset found between AMC and RC and thereby implies that HMC's absolute level is reasonable (Pick and Korte, 2017). The offset between Dcx and HMC is considerably larger (-19.74 nT). Solar cycle dependent offsets calculated for Dst in the order of -15 nT (Lühr and Maus, 2010; Temerin and Li, 2015) do not

apply, because Dcx corrects for the so called “nonstorm component" (Karinen and Mursula, 2006), shifting the index to more positive values on average.
Differences between q^0_{1M} COV-OBS and HMC (Fig. 4.5d, gray) are structured in three parts with a moderate difference level until ∼1935, large differences up to 30 nT between 1935–1965 and comparably small differences afterwards. These “steps" in the offsets could be due to the different data bases for COV-OBS.x1 and HMC. COV-OBS.x1 uses annual mean observatory measurements which do not reveal data problems as clearly as hourly means do. The deviations prior to 1935 could therefore be caused by our data modifications that are not accessible to COV-OBS.x1. Satellite data is used in COV-OBS.x1 starting with the POGO series, which could explain the step in the 1960s.

4.5 DISCUSSION

In this section we discuss the influence of the data modifications on HMC’s low-frequency variability (**Q1**, section 4.5.1) and to what extent this variability can be traced back to OSF (**Q2**, section 4.5.2).

4.5.1 *Sensitivity of HMC to data modifications and robustness of long-term trend*

First, we define $\mathrm{HMC}_{\mathrm{no\ cor}}$, which includes no data manipulations apart from the removal of large spikes in the original .wdc files. $\mathrm{HMC}_{\mathrm{no\ cor}}$ and the final HMC (including all modifications) are compared in Fig. 4.6. These versions differ significantly during 1905–1947, moderately during 1947–2003 and insignificantly afterwards. Special attention has to be paid to the influence of our drift removal (section 4.2.1) on the inferred long-term trend in HMC. $\mathrm{HMC}_{\mathrm{no\ drift}}$ gives the spread in results obtained from data with drift removal (HMC) and without drift removal in either X, Y, Z or all components (5 versions). In accordance with the final HMC, the inferred trends consistently indicate a decrease in absolute value, varying between 7.3 % and 16.2 % since 1900. We can thus conclude, that the detected weakly positive trend in HMC is most likely genuine, i.e. not an artifact of our data modifications.

While it is expected, that the data modifications as a whole improve the result, the effect of some problematic data segments or specific observatories (Fig. 4.3, dashed boxes) is not obvious. Figure 4.6 shows versions of HMC that do not include unreliable looking data ($\mathrm{HMC}_{\mathrm{no\ bad}}$), data from WIA observatory ($\mathrm{HMC}_{\mathrm{no\ WIA}}$) and data from MBO observatory ($\mathrm{HMC}_{\mathrm{no\ MBO}}$). Based on these test runs we can rule out, that the alternative method to determine the crustal bias for WIA or the possible presence of disturbances from the equatorial electrojet in the data of MBO significantly alter the result. However, discarding the unreliable data segments significantly changes the index prior to ∼1935 (HMC vs. $\mathrm{HMC}_{\mathrm{no\ bad}}$) so that the overall trend flips from positive (blue dashed) to negative (orange dashed). The fact, that $\mathrm{HMC}_{\mathrm{no\ bad}}$ reaches unrealistically large positive values made us keep all data and associated data modifications in the final version of HMC.

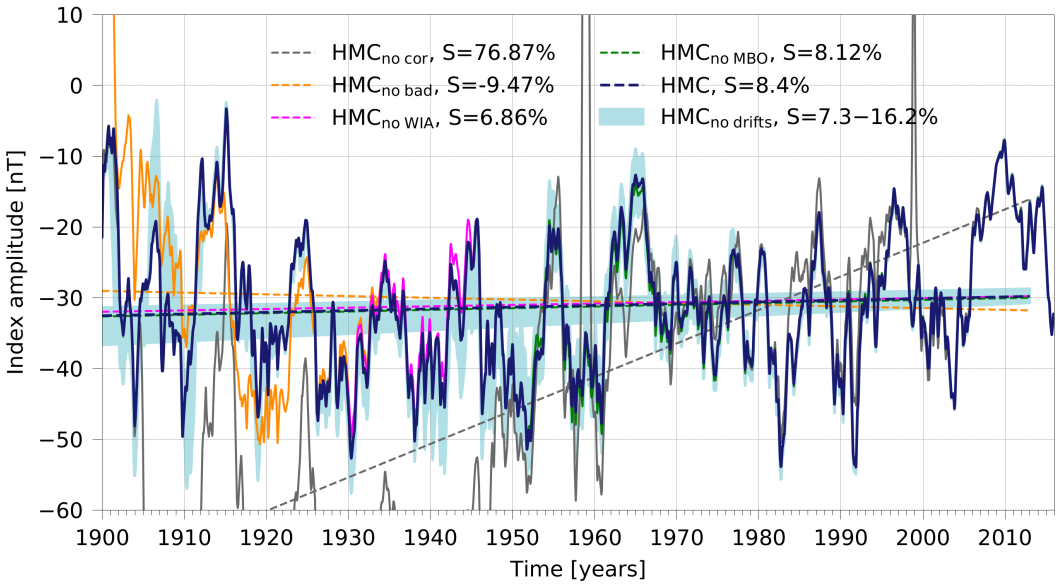

Figure 4.6: Comparison of HMC (dark blue) to index versions based on different data sets with linear fits. **no cor:** No data manipulations as shown in Fig. 4.3, gray. **no bad:** Data segments circled in Fig. 4.3 are discarded (apart from MBO and WIA). **no WIA & no MBO:** All data from WIA or MBO observatories are discarded. **no drifts:** Spread in results for data with drift removal (HMC) and without drift removal in either X, Y, Z or all components.

4.5.2 *OSF as driver for HMC variability*

A detailed comparison between HMC and OSF (**Q2**) is shown in Fig. 4.7. In contrast to HMC the reconstructed OSF shows a linear increase (absolute values) of 31.46 % since 1900 (Fig. 4.7a,b). The major difference between these trends originates from the period 1935–1965, during which HMC and OSF exhibit oppositely directed trends (Tab. 4.1).
Fig. 4.7c,d shows the coefficient of determination (r^2) between OSF and either HMC, q^0_{1M} COV-OBS, Dcx or RC for annual means (c, OSF reconstructed) and 27-day means (d, OSF measured). r^2 is calculated on different time scales, starting with values since 1900 (left) and ending with values since 1997 (right). An example of the underlying linear regressions is given in Fig. 4.7e, showing HMC vs. OSF since 1997 (measured) and since 1900 (reconstructed) with corresponding linear fits. In both cases, ∼60 % of the variability in HMC can be explained by OSF. Three pieces of information can be extracted from Fig. 4.7c–e in total. First, OSF is related to annual means of geomagnetic indices more closely (greater r^2) than to 27-day means (e.g. 0.96 for annual means and 0.62 for 27-day means since 1997). This is expected, given that the link between OSF and geomagnetic activity (section 4.2.2.2, relation 2) deteriorates as the averaging time scale falls below one year. Secondly, OSF is related to HMC more closely than to the other proxies (e.g. 0.62 for HMC and 0.54 for RC since 1997), which supports the usefulness of HMC for describing low-frequency changes in magnetospheric magnetic fields. Lastly, r^2 increases as the time series decrease in length, i.e. as earlier times are excluded (Fig. 4.7c). If the long-term trends (Tab. 4.1) are removed from the annual means of HMC and the

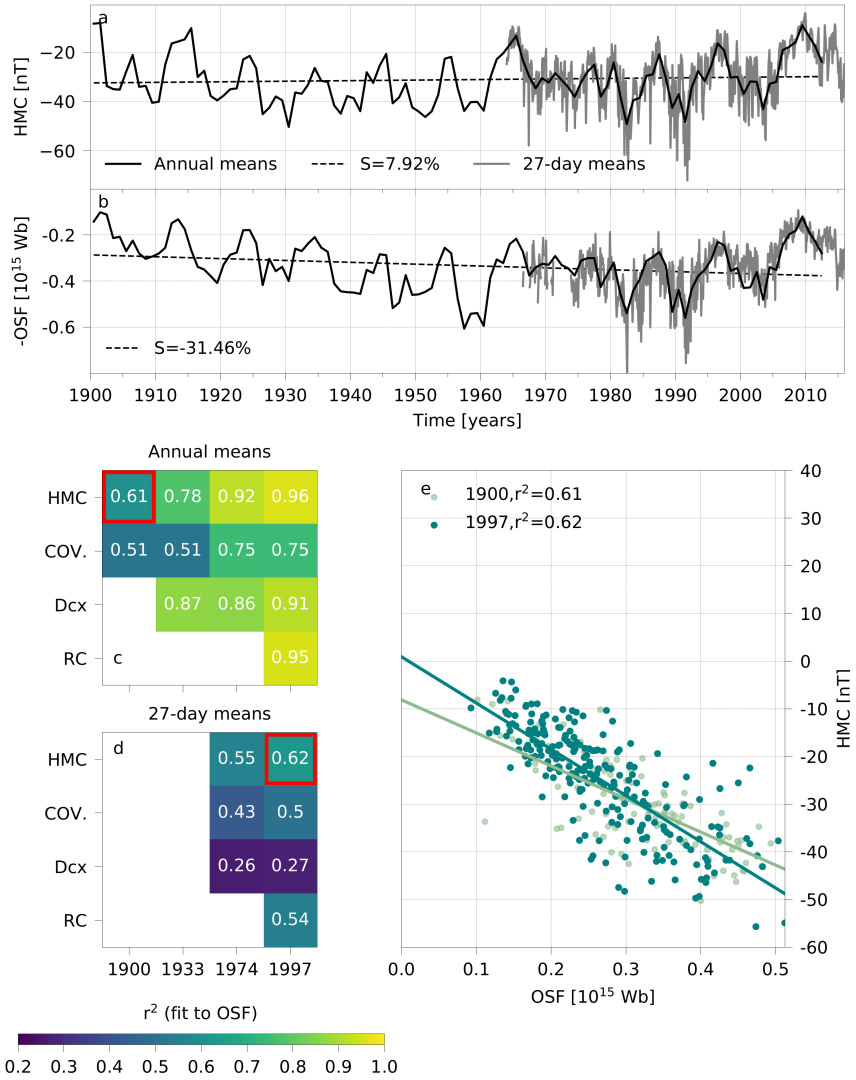

Figure 4.7: **a,b:** Annual and 27-day means of HMC and the kinematically corrected OSF with linear fits. Annual values for OSF are taken from its reconstruction (section 4.2.2.2). **c,d:** Coefficients of determination (r^2) for OSF (c: reconstructed, d: measured) and either HMC, q^0_{1M} COV-OBS.x1 (COV), Dcx or RC for four different time intervals – since 1900 (full), since 1933 (start Dcx), since 1974 (start continuously measured OSF) and since 1997 (start RC). **e:** Linear regressions for HMC and OSF corresponding to the r^2 in the red framed boxes.

reconstructed OSF prior to the calculation of r^2, this effect weakens (0.68 instead of 0.61 since 1900), but still remains present. Apparently, there are mechanisms at work limiting r^2 by causing not only the trends, but also a different kind of disagreeing variability among HMC and OSF.

Before we discuss two such mechanisms, it is important to recall, that HMC and the reconstructed OSF are based on the same type of data, as OSF is reconstructed from the IDV(1d) and aa_c geomagnetic indices. Therefore, r^2 in Fig. 4.7c is larger than it would be for annual means of the independently measured OSF ($r^2 = 0.8$ since 1974 and $r^2 = 0.88$ since 1997). Nevertheless, HMC and the reconstructed OSF do not share the same information, because the sets of observatories used to derive

Table 4.1: Slopes of piece-wise linear fits to annual means of HMC and $HMC_{no\ bad}$ (or to low-pass filtered versions in brackets), SN, F10.7 (reconstructed) and OSF (reconstructed). Slopes are given in percent and refer to the level at the beginning of the time interval. Arrows indicate the direction of the slope in Fig. 4.5c, Fig. 4.6 and Fig. 4.7a,b.

Years	HMC [%]	$HMC_{no\ bad}$ [%]	SN [%]	F10.7 [%]	OSF [%]
1900–1935	-59.9 (-60.1) ↓	-371.8 (-407.3) ↓	21.6	19.6	55.6 ↓
1935–1965	27.6 (28.2) ↑	27.6 (28.2) ↑	-21.9	7.1	1.6 ↓
1965–2012	25.5 (26.2) ↑	25.5 (26.2) ↑	-64.6	-26.0	-17.3 ↑
1900–2012	7.9 (8.4) ↑	-9.4 (-9.4) ↓	50.0	29.9	31.5 ↓

IDV(1d), aa_c and HMC are disjunct, except for the years 1900–1910, when data from POT and SED are used for IDV(1d).
The first mechanism is tied to the data from these geomagnetic observatories. The increasing degradation in geomagnetic data quality when moving to the past means, that the magnetic signal from magnetospheric currents is increasingly “buried" under station specific measurement noise. This causes an artificial deviation between geomagnetic indices based on data from different stations. Thus, this mechanism can explain the limitation of r^2 between HMC and the measured/reconstructed OSF as well as the increase of r^2 with time.
The second mechanism is tied to the derivation method of HMC. Obviously, HMC and OSF are not as directly connected as suggested in section 4.2.2.2, likely because the second connection in relation 2 is compromised: The response of HMC to v_{sw} is probably not homogeneous, neither in space, nor in time. Finch et al., (2008) showed, that geomagnetic data from high latitudes have a stronger dependence on v_{sw} than that from low latitudes. Since HMC includes data from stations at latitudes $14^\circ \leq \theta'_d \leq 75^\circ$, it is likely that the response to v_{sw} varies among them. Furthermore, the combination of observatories contributing to HMC varies irregularly with time (Fig. 4.3). Thus, this mechanism can explain disagreements between OSF and HMC not caused by the first mechanism. In particular, it can help to explain the exceptional period 1935–1965, when (a) the trends in OSF and HMC are oppositely directed (Tab. 4.1), (b) r^2 between OSF and Dcx is greater than between OSF and HMC (Fig. 4.7c) and (c) HMC is further offset from COV-OBS.x1 than at all other times (Fig. 4.5d, gray).
The above points provide strong motivation for a future study focused on the individual responses of HMC observatories to the solar wind ($B \cdot v_{sw}^n$).

4.6 CONCLUSIONS

We have constructed the geomagnetic index HMC to investigate low-frequency variations in the magnetic signature of magnetospheric currents on ground since 1900. The greatest advantage of HMC over already existing indices is its reasonable absolute level and enhanced power at periods roughly between one and 11 years (solar cycle). In this frequency range HMC is an improvement over the only other

comparable long-term index, Dcx (since 1933). Annual means of HMC show a weak linear decrease in magnitude of 7.9 % between 1900 and 2012, indicating that the strength of the magnetospheric currents, to which HMC is sensitive, may have decreased with time. The weakening of this long-term trend as compared to that in AMC (-35.4 %) can be attributed to the data modifications performed, especially prior to ∼1947. However, the long-term trend in HMC remains highly sensitive to the choice of early observatory data (prior to ∼1935) and should be interpreted with caution.

We find that most of the low-frequency variability in HMC can be attributed to the reconstructed open solar flux (∼60 % since 1900, ∼96 % since 1997), and that OSF is more closely related to HMC than to similar proxies (RC, Dcx, q^0_{1M} COV-OBS). The agreement between HMC and OSF is limited by the degrading geomagnetic data quality in the past and, probably, the inhomogeneous response of HMC observatories to solar wind forcing. Hence, these mechanisms could be responsible for the disagreement of trends in HMC and OSF (+31.46 % since 1900). However, most of this disagreement originates from an exceptional period roughly covering 1935 to 1965, when OSF agrees better with Dcx than with HMC, which in turn deviates noticeably from the COV-OBS.x1 external field measure. At this stage we can not explain these peculiarities, which are subject of further investigation.

ACKNOWLEDGMENTS

L. Pick thanks N. Olsen and C. Finlay for helpful discussions. The kinematically corrected OSF was kindly provided by M. Lockwood. We thank the anonymous referees for their suggestions that helped to improve the manuscript. L. Pick, M. Korte and Y. Thomas acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under project SPP1788 "Dynamic Earth"–DFG KO 2870/6-1. All publicly available third party material used is listed in the Supporting Information, Text S3. The HMC index, the modified observatory data and a documentation of all modifications performed can be accessed via GFZ Data Servies (Pick and Korte, 2018, https://doi.org/10.5880/GFZ.2.3.2018.006).

REFERENCES

Brown, W. J., J. E. Mound, and P. W. Livermore (2013). "Jerks abound: An analysis of geomagnetic observatory data from 1957 to 2008." In: *Physics of the Earth and Planetary Interiors* 223, pp. 62–76. DOI: 10.1016/j.pepi.2013.06.001 (cit. on p. 53).

Finch, I. D. and M. Lockwood (2007). "Solar wind-magnetosphere coupling functions on timescales of 1 day to 1 year." In: *Annales Geophysicae* 25.2, pp. 495–506. DOI: 10.5194/angeo-25-495-2007 (cit. on p. 57).

Finch, I. D., M. Lockwood, and A. P. Rouillard (2008). "Effects of solar wind magnetosphere coupling recorded at different geomagnetic latitudes: Separation of directly-driven and storage/release systems." In: *Geophysical Research Letters* 35.21. DOI: 10.1029/2008GL035399 (cit. on p. 67).

Finlay, C. C. et al. (2017). "Challenges Handling Magnetospheric and Ionospheric Signals in Internal Geomagnetic Field Modelling." In: *Space Science Reviews* 206.1, pp. 157–189. DOI: 10.1007/s11214-016-0285-9 (cit. on p. 52).

Finlay, C. C., N. Olsen, S. Kotsiaros, N. Gillet, and L. Tøffner-Clausen (2016). "Recent geomagnetic secular variation from Swarm and ground observatories as estimated in the CHAOS-6 geomagnetic field model." In: *Earth, Planets and Space* 68.112. DOI: 10.1186/s40623-016-0486-1 (cit. on pp. 53, 58).

Gannon, J. L. (2012). "Assessing the validity of station location assumptions made in the calculation of the Geomagnetic Disturbance Index, Dst." In: *Space Weather* 10.S02002. DOI: 10.1029/2011SW000731 (cit. on p. 54).

Ganushkina, N. Y., M. W. Liemohn, and S. Dubyagin (2018). "Current Systems in the Earth's Magnetosphere." In: *Reviews of Geophysics* 56.2, pp. 309–332. DOI: 10.1002/2017RG000590 (cit. on p. 52).

Gillet, N., D. Jault, E. Canet, and A. Fournier (2010). "Fast torsional waves and strong magnetic field within the Earth's core." In: *Nature* 465, pp. 74–77. DOI: 10.1038/nature09010 (cit. on p. 52).

Gillet, N., O. Barrois, and C. C. Finlay (2015). "Stochastic forecasting of the geomagnetic field from the COV-OBS.x1 geomagnetic field model, and candidate models for IGRF-12." In: *Earth, Planets and Space* 67.71. DOI: 10.1186/s40623-015-0225-z (cit. on p. 54).

Grayver, A. V. et al. (2017). "Joint inversion of satellite-detected tidal and magnetospheric signals constrains electrical conductivity and water content of the upper mantle and transition zone." In: *Geophysical Research Letters* 44, pp. 6074–6081. DOI: 10.1002/2017GL073446 (cit. on p. 58).

Harvey, K. L. (1992). "The Cyclic Behavior of Solar Activity." In: *The Solar Cycle, ASP Conference Series*. Ed. by K. L. Harvey. Vol. 27. Astronomical Society of the Pacific Conference Series, p. 335 (cit. on p. 52).

Harvey, K. L. (1994). "Irradiance Models Based on Solar Magnetic Fields." In: *International Astronomical Union Colloquium* 143, pp. 217–225. DOI: 10.1017/S0252921100024714 (cit. on p. 52).

Karinen, A. and K. Mursula (2006). "Correcting the Dst index: Consequences for absolute level and correlations." In: *Journal of Geophysical Research: Space Physics* 111.A8. DOI: 10.1029/2005JA011299 (cit. on p. 64).

Kauristie, K. et al. (2017). "On the Usage of Geomagnetic Indices for Data Selection in Internal Field Modelling." In: *Space Science Reviews* 206.1, pp. 61–90. DOI: 10.1007/s11214-016-0301-0 (cit. on p. 52).

Krivova, N. A. and S. K. Solanki (2004). "Effect of spatial resolution on estimating the Sun's magnetic flux." In: *Astronomy and Astrophysics* 417.3, pp. 1125–1132. DOI: 10.1051/0004-6361:20040022 (cit. on p. 52).

Krivova, N. A., L. E. A. Vieira, and S. K. Solanki (2010). "Reconstruction of solar spectral irradiance since the Maunder minimum." In: *Journal of Geophysical Research: Space Physics* 115.A12112. DOI: 10.1029/2010JA015431 (cit. on p. 57).

Lesur, V. et al. (2015). "Parent magnetic field models for the IGRF-12 GFZ-candidates." In: *Earth, Planets and Space* 67.1, pp. 87–102. DOI: 10.1186/s40623-015-0239-6 (cit. on p. 52).

Lockwood, M. (2013). "Reconstruction and Prediction of Variations in the Open Solar Magnetic Flux and Interplanetary Conditions." In: *Living Reviews in Solar Physics* 10.4. DOI: 10.12942/lrsp-2013-4 (cit. on pp. 52, 57).

Lockwood, M., M. Owens, and A. P. Rouillard (2009a). "Excess open solar magnetic flux from satellite data: 1. Analysis of the third perihelion Ulysses pass." In: *Journal of Geophysical Research: Space Physics* 114.A11. DOI: 10.1029/2009JA014449 (cit. on p. 57).

– (2009b). "Excess open solar magnetic flux from satellite data: 2. A survey of kinematic effects." In: *Journal of Geophysical Research: Space Physics* 114.A11. DOI: 10.1029/2009JA014450 (cit. on p. 57).

Lockwood, M. et al. (2013a). "Reconstruction of geomagnetic activity and near-Earth interplanetary conditions over the past 167 yr – Part 1: A new geomagnetic data composite." In: *Annales Geophysicae* 31.11, pp. 1957–1977. DOI: 10.5194/angeo-31-1957-2013 (cit. on p. 57).

– (2013b). "Reconstruction of geomagnetic activity and near-Earth interplanetary conditions over the past 167 yr – Part 2: A new reconstruction of the interplanetary magnetic field." In: *Annales Geophysicae* 31.11, pp. 1979–1992. DOI: 10.5194/angeo-31-1979-2013 (cit. on p. 57).

Lockwood, M. et al. (2014). "Reconstruction of geomagnetic activity and near-Earth interplanetary conditions over the past 167 yr – Part 4: Near-Earth solar wind speed, IMF, and open solar flux." In: *Annales Geophysicae* 32.4, pp. 383–399. DOI: 10.5194/angeo-32-383-2014 (cit. on pp. 57, 74).

Lühr, H. and S. Maus (2010). "Solar cycle dependence of quiet-time magnetospheric currentsand a model of their near-Earth magnetic fields." In: *Earth, Planets and Space* 62, pp. 843–848. DOI: 10.5047/eps.2010.07.012 (cit. on pp. 52, 63).

Milan, S. E. et al. (2017). "Overview of Solar Wind-Magnetosphere-Ionosphere-Atmosphere Coupling and the Generation of Magnetospheric Currents." In: *Space Science Reviews* 206, pp. 547–573. DOI: 10.1007/s11214-017-0333-0 (cit. on p. 52).

Mursula, K. and A. Karinen (2005). "Explaining and correcting the excessive semi-annual variation in the Dst index." In: *Geophysical Research Letters* 32.14. DOI: 10.1029/2005GL023132 (cit. on p. 52).

Olsen, N. (1993). "The solar cycle variability of lunar and solar daily geomagnetic variations." In: *Annales Geophysicae* 11.4, pp. 254–262 (cit. on p. 55).

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. DOI: 10.1186/BF03351897 (cit. on p. 52).

Olsen, N. et al. (2014). "The CHAOS-4 geomagnetic field model." In: *Geophysical Journal International* 197.2, pp. 815–827. DOI: 10.1093/gji/ggu033 (cit. on p. 52).

Parker, E. N. (1958). "Dynamics of the Interplanetary Gas and Magnetic Fields." In: *Astophysical Journal* 128, p. 664. DOI: 10.1086/146579 (cit. on p. 57).

– (1963). *Interplanetary Dynamical Processes*. Interscience Publishers, New York (cit. on p. 57).

Pick, L. and M. Korte (2017). "An annual proxy for the geomagnetic signal of magnetospheric currents on Earth based on observatory data from 1900–2010." In:

Geophysical Journal International 211.2, pp. 1223–1236. DOI: 10.1093/gji/ggx367 (cit. on pp. 53, 59, 60, 63).
– (2018). "HMC index." In: *GFZ Data Services*. DOI: 10.5880/GFZ.2.3.2018.006 (cit. on p. 68).
Russell, C. T. (1971). "Geophysical Coordinate Transformations." In: *Cosmic Electrodynamics* 2, pp. 184–196 (cit. on p. 54).
Sabaka, T. J., N. Olsen, and M. E. Purucker (2004). "Extending comprehensive models of the Earth's magnetic field with Ørsted and CHAMP data." In: *Geophysical Journal International* 159.2, pp. 521–547. DOI: 10.1111/j.1365-246X.2004.02421.x (cit. on p. 58).
Solanki, S. K., M. Schüssler, and M. Fligge (2000). "Evolution of the Sun's large-scale magnetic field since the Maunder minimum." In: *Nature* 408, pp. 445–447. DOI: 10.1038/35044027 (cit. on p. 52).
– (2002). "Secular variation of the Sun's magnetic flux." In: *Astronomy and Astrophysics* 383.2, pp. 706–712. DOI: 10.1051/0004-6361:20011790 (cit. on p. 52).
Stamper, R., M. Lockwood, M. N. Wild, and T. D. G. Clark (1999). "Solar causes of the long-term increase in geomagnetic activity." In: *Journal of Geophysical Research: Space Physics* 104.A12, pp. 28325–28342. DOI: 10.1029/1999JA900311 (cit. on p. 57).
Sugiura, M. (1964). "Hourly values of equatorial Dst for the IGY." In: *Annales of the International Geophysical Year* 35, pp. 4–45 (cit. on p. 52).
Tapping, K. F. (2013). "The 10.7 cm solar radio flux (F10.7)." In: *Space Weather* 11.7, pp. 394–406. DOI: 10.1002/swe.20064 (cit. on p. 52).
Temerin, M. and X. Li (2015). "The Dst index underestimates the solar cycle variation of geomagnetic activity." In: *Journal of Geophysical Research: Space Physics* 120.7, pp. 5603–5607. DOI: 10.1002/2015JA021467 (cit. on pp. 52, 63).
Thébault, E. et al. (2015). "International Geomagnetic Reference Field: the 12th generation." In: *Earth, Planets and Space* 67.79. DOI: 10.1186/s40623-015-0228-9 (cit. on p. 54).
Thomson, A. W. P. and V. Lesur (2007). "An improved geomagnetic data selection algorithm for global geomagnetic field modelling." In: *Geophysical Journal International* 169.3, pp. 951–963. DOI: 10.1111/j.1365-246X.2007.03354.x (cit. on p. 52).
Wardinski, I. and R. Holme (2011). "Signal from noise in geomagnetic field modelling: denoising data for secular variation studies." In: *Geophysical Journal International* 185.2, pp. 653–662. DOI: 10.1111/j.1365-246X.2011.04988.x (cit. on p. 53).
Wu, C.-J., N. . Krivova, S. K. Solanki, and I. G. Usoskin (2018). "Solar total and spectral irradiance reconstruction over the last 9000 years." In: *Astronomy and Astrophysics* 620.A120. DOI: 10.1051/0004-6361/201832956 (cit. on p. 57).
Yamazaki, Y. and A. Maute (2017). "Sq and EEJ – A review on the daily variation of the geomagnetic field caused by ionospheric dynamo currents." In: *Space Science Reviews* 206.1–4, pp. 299–405. DOI: 10.1007/s11214-016-0282-z (cit. on p. 58).

SUPPORTING INFORMATION

Text S1 and Figure 4.S1 provide additional information on how the iterative inversion (section 4.3.2, Eq. 4.13) is carried out. Text S2 and Figure 4.S2 cover the inverted internal dipole coefficients g_1^0, g_1^1, h_1^1 (section 4.3.2, Eq. 4.10). Text S3 lists all publicly available third partly material and provides information on how to access it. Table 4.S1 lists the HMC geomagnetic observatories together with their locations and estimated crustal biases (section 4.3.1, Eq. 4.4).

Text S1

At each hour we have to know the orientation of the dipole axis (g_1^0, g_1^1, h_1^1) in order to convert the observatory positions into *mag* coordinates needed for the last column of matrix **G** (Eq. 4.12). An iteration is implemented, which starts with the dipole axis orientation given by COV-OBS.x1 (k=0) and subsequently uses the result from our previous iteration to perform the coordinate transformation. HMC takes one iteration longer (4 iterations) than the internal coefficients to converge, which is why we terminate the calculation after iteration four (Fig. 4.S1).

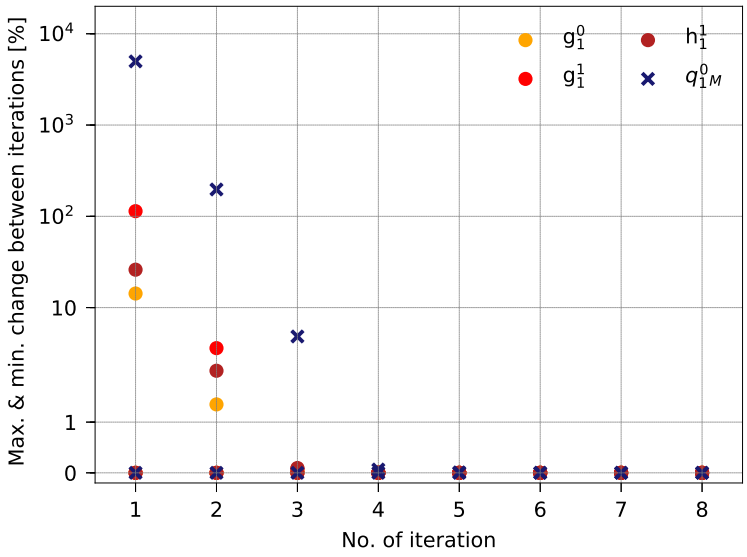

Figure 4.S1: Relative change in internal dipole coefficients (*seu* system, dots) and the external axial dipole coefficient (*mag* system, crosses) between consecutive iterations over the dipole axis orientation. For each iteration the largest and the smallest change from all hours within the period 1900–2015 are shown. The COV-OBS.x1 dipole axis orientation is used for iteration $k = 0$. The normalization is done with respect to the largest change in the respective coefficients within 1900–2015 as given by COV-OBS.x1.

Text S2

Our internal coefficients (g_1^0, g_1^1, h_1^1) are not directly related to questions **Q1** & **Q2**, yet they are an integral part of our inversion and interact with HMC (Eq. 4.10). We

show the difference between our coefficients and those from COV-OBS.x1 in Figure 4.S2. The spectral coherence (Fig. 4.S2a) is highest for g_1^0 followed by h_1^1 and g_1^1. Explaining the prominent notches at periods of ∼2 and ∼3 years requires further work beyond the scope of this study. The time series differences (Fig. 4.S2b) adhere to the three-part structure observed for the external dipole coefficient (Fig. 4.5d), this time inverted: Deviations are largest (up to ∼20 nT) prior to 1935, they reduce during 1935–1965 and increase again afterwards.

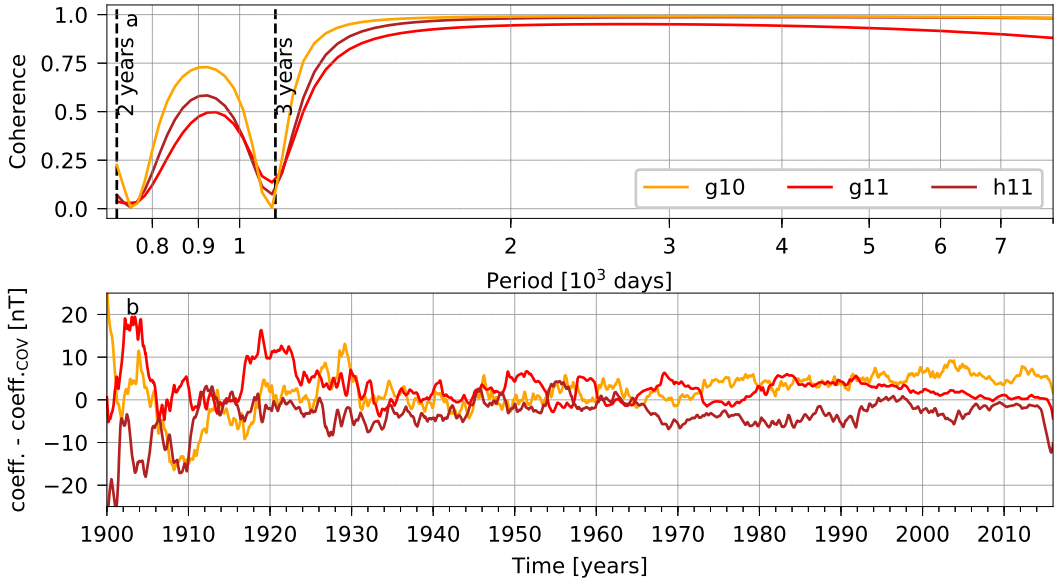

Figure 4.S2: **a:** Magnitude squared coherence of our internal dipole coefficients (g_1^0, g_1^1, h_1^1) and those from COV-OBS.x1 on semi-log scale. **b:** Differences between our internal dipole coefficients and those from COV-OBS.x1.

Text S3

All websites were last accessed on December 31, 2018.

Data:

1. Observatory data (hourly means): WDC Edinburgh.
   `ftp://ftp.nmh.ac.uk/wdc/obsdata/hourval/`

2. Observatory locations: IAGA Division 5 fusion table.
   `https://www.bgs.ac.uk/iaga/vobs/`

3. Relocation times & offsets: British Geological Survey.
   `http://www.geomag.bgs.ac.uk/data_service/data/annual_means.shtml`

Activity indicators:

1. Dcx (hourly means): University of Oulu, Finland.
   `http://dcx.oulu.fi/?link=queryDefinite`

2. RC (hourly means): DTU Space, version 07/02/2018.
   http://www.spacecenter.dk/files/magnetic-models/RC
3. F10.7 (rotational means): Space Weather Canada.
   https://www.spaceweather.gc.ca/solarflux/sx-5-en.php
4. Total sunspot number SN (annual means): SILSO.
   http://www.sidc.be/silso/datafiles
5. OSF recon. (annual means): Lockwood et al., (2014), "OSFg1".

Model forward codes:

1. COV-OBS.x1: DTU Space.
   http://www.spacecenter.dk/files/magnetic-models/COV-OBSx1/
2. CHAOS-6: DTU Space.
   http://www.spacecenter.dk/files/magnetic-models/CHAOS-6/
3. CM4: NASA (not available anymore).
   https://denali.gsfc.nasa.gov/geomag.html

Table 4.S1: Observatory properties: Locations in *ned* system are taken from IAGA Division 5 fusion table. For observatories having a * following their IAGA code we calculated the biases according to Equation 4.4. Otherwise, we used successor observatories to determine the bias (Eq. 4.5). MAD stands for median absolute deviation.

IAGA code	Lat. θ'_d [°]	Lon. ϕ [°]	Height h [m]	Bias X [nT]	Bias Y [nT]	Bias Z [nT]	MAD X [nT]	MAD Y [nT]	MAD Z [nT]
POT	52.383	13.0625	–	-1.769	-10.918	-117.393			
SED	52.278	13.010	–	-8.641	-13.999	-81.094			
NGK*	52.072	12.675	78	-14.209	-1.481	-80.784	2.551	2.811	1.291
PSM	48.817	2.483	49	63.289	-49.114	-29.570			
VLJ	48.821	2.014	–	-18.964	-134.519	19.797			
CLF*	48.025	2.260	145	-60.948	-14.921	103.195	2.598	2.588	1.295
EKT	56.827	60.632	278	-255.917	-679.596	181.049			
SVD	56.733	61.067	–	343.299	-331.398	-176.457			
ARS*	56.433	58.567	290	122.799	-266.249	455.533	3.256	3.287	1.633
CLH	38.733	-76.833	–	-35.042	-11.417	272.930			
FRD*	38.2047	-77.3729	69	79.378	-52.297	119.023	2.843	2.996	1.402
WIA	48.2	16.233	–	36.292	9.436	-20.095			
HRB*	47.873	18.190	112	28.032	-21.072	-83.169	3.666	3.507	2.947
IRT*	52.167	104.450	465	45.217	40.778	-118.210	3.318	3.084	4.183
TUC*	32.1745	-110.7337	946	-72.999	-268.188	423.825	2.814	2.980	1.894
AAA*	43.250	76.917	1300	157.892	29.966	-217.033	4.321	8.273	3.732
TOK	35.683	139.750	–	46.736	-1.595	-21.583			
KAK*	36.232	140.186	36	10.444	12.438	-94.422	2.040	2.624	1.572
VQS	18.150	-65.450	20	-198.155	111.590	-285.430			
SJG*	18.111	-66.1498	424	-50.052	129.196	126.491	2.944	6.747	2.371
HON*	21.317	-157.9996	4	-162.997	84.407	-322.935	2.283	2.693	1.674
SSH*	31.100	121.183	100	-221.906	79.907	214.335	3.053	9.570	2.680

. . . continued

IAGA code	Lat. θ'_d [°]	Lon. ϕ [°]	Height h [m]	Bias X [nT]	Bias Y [nT]	Bias Z [nT]	MAD X [nT]	MAD Y [nT]	MAD Z [nT]
KNY*	31.424	130.880	107	8.145	52.840	-41.295	2.082	2.790	1.477
MBO*	14.392	-16.958	7	113.895	35.333	58.108	2.673	2.789	2.268
ABG*	18.64	72.87	6	-76.002	440.596	717.086	4.450	11.873	4.298
VSS*	-22.400	-43.650	457	67.230	-54.079	-39.933	3.137	2.341	3.527
PIL*	-31.667	-63.883	336	17.057	-3.649	-12.124	4.000	2.681	1.707
TRW*	-43.267	-65.383	15	-6.615	-31.276	107.186	3.997	3.927	1.923
API*	-13.807	-171.775	4	-12.581	253.203	-874.726	4.280	9.124	10.355
CTO	-33.950	18.467	–	-0.408	-50.830	90.668			
HER*	-34.425	19.225	26	30.116	8.170	27.793	2.777	2.598	2.117
WAT	-30.317	115.883	–	-97.115	-286.820	58.676			
GNA*	-31.356	115.950	60	-15.132	-97.981	114.065	2.435	2.534	2.575
GNG*	-31.356	115.715	50	0.009	-39.710	164.200	2.376	2.586	1.956

5

A STATISTICAL CLASSIFIER FOR HISTORICAL GEOMAGNETIC STORM DRIVERS DERIVED SOLELY FROM GROUND-BASED MAGNETIC FIELD MEASUREMENTS

Leonie Pick[1,2,3], Frederic Effenberger[1], Irina Zhelvaskaya[1,2], Monika Korte[1]

[1] Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany
[2] University of Potsdam, Am Neuen Palais 10, House 9, 14469 Potsdam, Germany
[3] Jacobs University, Campus Ring 1, 28759 Bremen, Germany

 doi: 10.1029/2019EA000726

abstract Solar wind observations show that geomagnetic storms are mainly driven by Interplanetary Coronal Mass Ejections (ICMEs) and Co-rotating or Stream Interaction Regions (C/SIRs). We present a binary classifier that assigns one of these drivers to 7546 storms between 1930–2015 using ground-based geomagnetic field observations only. The input data consists of the long-term stable Hourly Magnetospheric Currents index alongside the corresponding mid latitude geomagnetic observatory time series. This data set provides comprehensive information on the global storm time magnetic disturbance field, particularly its spatial variability, over eight solar cycles. For the first time, we use this information statistically with regard to an automated storm driver identification. Our supervised classification model significantly outperforms unskilled baseline models (78 % accuracy with 26[19] % misidentified ICMEs[C/SIRs]) and delivers plausible driver occurrences with regard to storm intensity and solar cycle phase. Our results can readily be used to advance related studies fundamental to space weather research, e.g., studies connecting galactic cosmic ray modulation and geomagnetic disturbances. They are fully reproducible by means of the underlying open-source software (Pick, 2019, `https://doi.org/10.5880/GFZ.2.3.2019.003`).

plain language summary The Earth's magnetic field reaches out into space where it constantly interacts with the solar wind, a stream of charged particles from the Sun. Geomagnetic storms occur when discontinuities in the solar wind disturb the geomagnetic field, possibly causing failures of, e.g., electricity transmission and satellite communications. In order to mitigate the socioeconomic risk, a better understanding of these processes is necessary. So far, the two main storm drivers have been identified from satellite observations: Interplanetary Coronal Mass Ejections and Co-rotating or Stream Interaction Regions. However, quantitative investigations require as many storm events as possible from multiple decades with varying solar conditions. We present an innovative method that classifies the drivers of 7546 storms since 1930, using ground-based magnetic field measurements only. This effectively increases the sample size, because geomagnetic observatories were operational long before the space era. Our results are directly applicable to other current space weather studies.

5.1 INTRODUCTION

Understanding cause and effect of solar processes and geomagnetic disturbances is crucial in order to mitigate our socioeconomic vulnerability to space weather (e.g., Oughton et al., 2018). Measurements of the magnetic field disturbances at ground level ($\boldsymbol{\delta B}$) lead to the discovery that geomagnetic storms tend to be either weak and periodic, recurring after ≈27 days (one synodic solar rotation), or strong and sporadic (Greaves and Newton, 1929; Maunder, 1904). However, it was not until the beginning of the space age that Stream Interaction Regions (SIRs) and Interplanetary Coronal Mass Ejections (ICMEs) were identified as major storm drivers (both reviewed by Kilpua et al., 2017). Their geoeffectiveness depends primarily on a sustained southward directed Z-component ($B_{sw,z}$) of the interplanetary magnetic field (IMF, $\boldsymbol{B_{sw}}$) with regard to the ecliptic pole (Rostoker and Fälthammar, 1967). SIRs are regions of compressed plasma that typically form when tenuous, fast solar wind streams from coronal holes (Krieger et al., 1973) collide with preceding dense, slow streams from the streamer belt (Feldman et al., 1981). If the coronal hole survives long enough, the SIR co-rotates with the Sun and is called a Co-rotating Interaction Region (CIR), causing recurrent geomagnetic disturbances (Snyder et al., 1963; Wilcox and Ness, 1965). On the other hand, ICMEs are the interplanetary manifestations of transient magnetized plasma ejections from the Sun, causing sporadic geomagnetic disturbances (Gosling, 1993; Gosling et al., 1975).
Today, SIRs and ICMEs are identified based on characteristic signatures in solar wind plasma and IMF observations from the Wind, ACE (e.g., Jian et al., 2006a,b), and, since 2015, DISCOVR spacecraft near the Sun-Earth L1 Lagrangian Point. L1 is located at ≈1.5 million kilometer from the Earth towards the Sun. There, the slow–fast stream interfaces of SIRs are marked by a proton density drop and a rise of solar wind speed ($\boldsymbol{v_{sw}}$), proton temperature (T_p), and plasma beta (β), accompanied by large $|\boldsymbol{B_{sw}}|$ fluctuations. ICMEs are often associated with a preceding shock and a sheath region between the shock and the ejecta, which can be “magnetic clouds". These are marked by a slow rotation of $\boldsymbol{B_{sw}}$ with enhanced field magnitude and reduced variability, as well as low T_p and β. Due to their ambiguity a combination of these signatures is required for a reliable classification.
The type of geomagnetic disturbance provoked by SIRs and ICMEs differs according to the response system. In this study, we focus on hourly $\boldsymbol{\delta B}$ measurements from geomagnetic observatories at low and mid latitudes. It is well known that the Dst index (Sugiura, 1964), i.e., the symmetric part of the horizontal disturbance field ($\boldsymbol{\delta B_{H,sym}}$), responds more intensely to ICMEs than to SIRs, which is commonly attributed to differences in ring current (RC) intensity (e.g., Borovsky and Denton, 2006). At the same time, it is known that $\boldsymbol{\delta B_H}$ exhibits a pronounced asymmetry ($\boldsymbol{\delta B_{H,asy}}$), particularly between dawn and dusk (e.g., Akasofu and Chapman, 1964), which is quantified by the ASY-H index (Iyemori, 1990). This is associated with the partial ring current (PRC) system superposing the symmetric RC on the night side (e.g., Kamide and Fukushima, 1971). While the contributions of different magnetospheric current systems, i.e., cross-tail current, RC, and PRC, to $\boldsymbol{\delta B_{H,asy}}$ are studied extensively (e.g., Dubyagin et al., 2014), $\boldsymbol{\delta B_{H,asy}}$ has not been tested for its statistical C/SIR vs. ICME discrimination capability.

To this end, we demonstrate that the long-term record of $\boldsymbol{\delta B}$ can be exploited beyond the current state of the art. This is done by means of a binary classifier that assigns driver classes $Y \in \{0, 1\}$, 0: CIRs or SIRs, 1: ICMEs, to geomagnetic storms from 1930–2015. Although the vast majority of strong events falls into one of these two categories (cf. Table 1 of Richardson and Cane, 2012), the binary nature of the classifier is a simplification that does not appropriately account for mixed, slow solar wind or unclear drivers. This is particularly relevant for the relatively weak events we incorporate in our analysis (section 5.2.2). However, the fact that we can predict whether a storm event is clearly ICME driven or not (mostly C/SIRs) from ground-level data alone is already a noteworthy result.
Covering 85 years, the event detection method should consider the ≈11-year solar cycle variation of geomagnetic activity. A Dst-based selection is inappropriate, because the index suffers from a baseline error that causes an underestimation of solar cycle variability (e.g., Lühr and Maus, 2010; Temerin and Li, 2015). Therefore, we use a different $\boldsymbol{\delta B}$-based index, which is morphologically similar to Dst, but represents the low-frequency variations more robustly: the Hourly Magnetospheric Currents index (HMC, Pick et al., 2019).
Our results advance the understanding of how SIR and ICME structures are typically imprinted at ground level and how their occurrence varies over eight solar cycles. They can contribute to other studies linking long-term geomagnetic activity measurements to solar wind evolution (e.g., Martini et al., 2015; Mursula et al., 2004), but can also assist studies in related areas, e.g., the long-term correlation between galactic cosmic ray modulation and geomagnetic activity (e.g., Dumbović et al., 2012).

5.2 DATA

5.2.1 *Geomagnetic data and HMC index*

We use the HMC index (H_t, $t = 1, \ldots, N_h$) and the underlying processed geomagnetic observatory data, both available from Pick and Korte, (2018). The basis are the time series of hourly vector magnetic field measurements, $\boldsymbol{B_{it}}$, taken by $i = 1, \ldots, 28$ geomagnetic observatories (Figure 5.1a, IAGA codes from `http://www.bgs.ac.uk/iaga/vobs/`) and distributed by the World Data Centre for Geomagnetism, Edinburgh (`http://www.wdc.bgs.ac.uk/data.html`). Throughout 1930–2015, they cover geomagnetic latitudes 52° (SED) $\leq \theta'_{M,it} \leq$ 9° (ABG) on the northern and −43° (GNA) $\leq \theta'_{M,it} \leq$ −12° (VSS) on the southern hemispheres (for a definition of geomagnetic coordinates see Laundal and Richmond, 2017). As described in detail by (Pick et al., 2019), a quality check of the observatory time series led to a modified data set. Core, ionospheric (both time varying) and crustal field (static) contributions were calculated using third-party geomagnetic field models and subtracted from the data matrix. The residuals were transformed into the geomagnetic frame, $\boldsymbol{\delta B_{M,it}}$, and fit hour-by-hour to the gradient of a scalar potential, V, which was modeled by spherical harmonics of degree one (dipole field). With regard to the Earth's surface, internal and external contributions to V were separated resulting in three Gauss coefficients describing the core field

and one, the HMC index, describing the external magnetic field directed along the dipole axis. Similar to Dst, HMC is expected to measure the diamagnetic effect of the magnetospheric currents, especially the ring current. Throughout the study, we use a three-hour running mean with weights $[0.25, 0.5, 0.25]$ of the original HMC. This prevents a possible impairment of the auto-correlation of HMC values on the time scale of geomagnetic storms due to, e.g., sporadic observatory dropouts and thus supports the identification of storm peaks.

Our storm driver classification (section 5.3) exploits not only the symmetric disturbance field, measured by HMC, but also the asymmetric one. We calculate the field residuals as before, but we now remove our co-estimated core field (see above). Consistent with the definition of HMC, we transform the residuals into a Cartesian system, $\boldsymbol{\delta B_{M,it}} = (X_M, Y_M, Z_M)_{it}$ and consider only the Z_M-components, which are aligned with the dipole axis. This data set, $Z_{M,it}$, corresponds to different distributions in magnetic local time (MLT), which change throughout universal time (UT) according to the availability of observatory measurements. We follow the definition $\text{MLT} = (\phi_M - \phi_{M,S})\,/15 + 12$, where ϕ_M and $\phi_{M,S}$ are the geomagnetic longitudes of the observatory and the subsolar point, respectively. The largest gap in MLT ranges between 15 (March, 1936) and three hours (January, 1949) with a mean of $\approx$4.3 hours. We average $Z_{M,it}$ over 1-hour MLT bins, centered on half hours, and get Z_{jt}, $j = 0, \ldots, 23$ per UT hour. In recognition of the prominent dawn-dusk asymmetry in the magnetic disturbance field (e.g., Newell and Gjerloev, 2012), and the ASY-H index, we define

$$ASY_{dd,t} = \frac{Z_{6t} - Z_{18t}}{2}\ [\text{nT}], \qquad ASY_{max,t} = \frac{max\left(Z_{jt}\right) - min\left(Z_{jt}\right)}{2}\ [\text{nT}] \tag{5.1}$$

as measures of the asymmetry size. To mitigate the problem of empty MLT bins, $\boldsymbol{Z}$ is again averaged over six 1-hour bins, giving a total of four new bins: $\hat{Z}_{jt}$, $j = 0, 6, 12, 18$, centered on 00:30 (midnight), 06:30 (dawn), 12:30 (noon) and 18:30 (dusk).

Figure 5.1b shows HMC and $\hat{\boldsymbol{Z}}$ for selected "training events" (section 5.2.2) driven by either C/SIRs or ICMEs, superposed at the event peaks. Panel c shows $\boldsymbol{Z}$ averaged over all peak hours, while two examples of individual events, for which some MLT bins are not sampled by observatory measurements, are given in panel d. Several previously discovered storm characteristics are evident, namely, that (1) ICME-driven storms trigger a stronger response, i.e., larger absolute HMC peak values (panel b, see section 5.1), (2) ASY_{dd} and ASY_{max} are larger for stronger than for weaker storms (panels b, c; see also Love and Gannon, 2009; Siscoe et al., 2012; Yakovchouk et al., 2012) and (3) departures from the typical behaviour can be large at any time (panel d).

5.2.2 *Training and target events*

Fundamental to our supervised classification method (section 5.3) is the definition of a training set (d_{tr}), from which the model parameters are "learned". This data set is defined for geomagnetic storm events $t_{tr,i}$, $i = 1, \ldots, N_{tr}$ with known input data and driver class labels as output ($\boldsymbol{Y_{tr}}$). The events are referred to by the UT hours at which HMC reaches local minima, i.e., the event peak times. The zero-entries

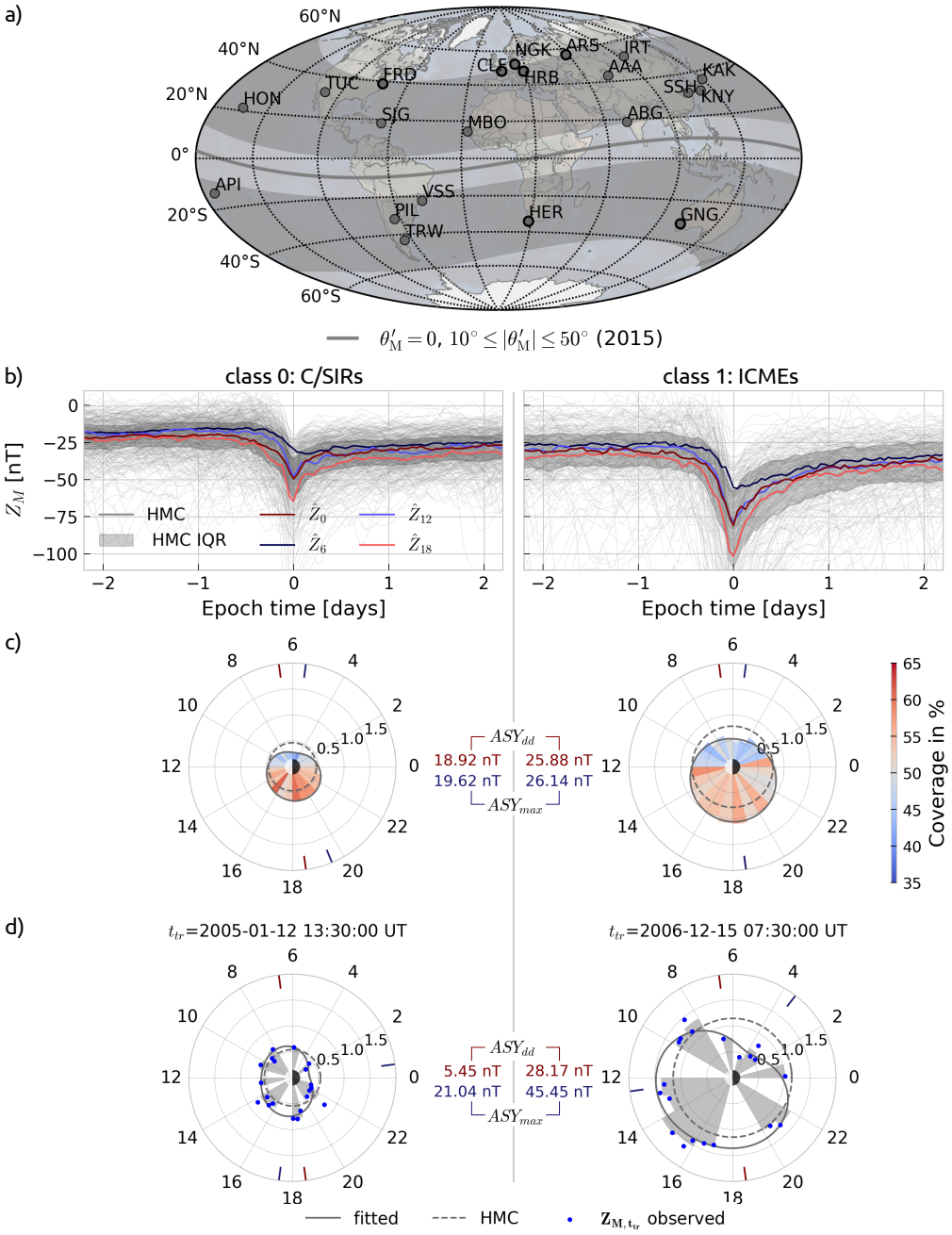

Figure 5.1: *Input data.* **a)** HMC observatories with respect to the geomagnetic equator in 2015 from IGRF-12 (gray line, Thébault et al., 2015). Thick circles mark stations with predecessors, gray bands mark [10°, 50°] latitude ranges. **b)** HMC (gray) and MLT-resolved disturbances ($\hat{\boldsymbol{Z}}$, colored) of training events superposed at peak times $\boldsymbol{t_{tr}}$. **c)** Polar plots showing $-\boldsymbol{Z}/100$ [nT] (radius) averaged over $\boldsymbol{t_{tr}}$ in dependence on MLT (angle) with a second order Fourier fit to the data (gray solid) and MLT-symmetric HMC (dashed). The color bar indicates the sampling degree per MLT bin. Red and blue dashes mark the MLTs for ASY_{dd} and ASY_{max}. **d)** Same as c) for two events with particular observatory data distributions (dots). **Abbr.** HMC: Hourly Magnetospheric Currents index; MLT: magnetic local time; IQR: interquartile range.

of $\boldsymbol{Y_{tr}}$ ($i_{tr0,j}, j = 1, \ldots, N_{tr0}$) refer to C/SIR driven events, while the remaining ones ($i_{tr1,j}, j = 1, \ldots, N_{tr1}$) have class label one and refer to ICME driven events.
In order to compile the training set, we collected a reference set from published catalogues (d_{re}) with N_{re}=868 already classified events, dated between 1995–2015. Of these, 571 are C/SIR and 297 ICME driven, giving a class ratio of $\gamma_{re} = N_{re0}/N_{re1} \approx 1.9$ in favor of the C/SIR drivers. Specifically, 745 events ($N_{re0} = 522$, $N_{re1} = 223$) are taken from Jian et al., (2011, 2006a,b) during 1995–2009, 77 additional events ($N_{re0} = 29$, $N_{re1} = 48$) from Turner et al., (2009) during 1995–2004 (classification following Richardson et al., 2000) and 46 events ($N_{re0} = 20$, $N_{re1} = 26$) from Shen et al., (2017) during 2013–2015. The underlying classification methods are based on a manual identification of characteristic signatures in plasma and IMF observations from ACE and Wind as outlined in section 5.1.
Once the model is trained, it can predict the driver classes of storm events from the target set (d_{ta}) between 1930 and 2015. The target event peak hours are identified with regard to HMC ($\boldsymbol{H}$) in three steps, which are visualized for clarity in Appendix A.

1. All hours, t, for which $H_t < Hl_t$ are marked:

$$H_t < Hl_t = P_n \left(\frac{H_{11y,t}}{min\left(H_{11y,t}\right)} \right)^{p_{sc}} \tag{5.2}$$

   P_n is the n-th HMC percentile during 1930–2015 and $\boldsymbol{H_{11y}}$ is a low-pass filtered HMC with a cutoff period of 11 years. With $p_{sc} = 1$, the term in brackets acts as a scaling factor in dependence of the solar cycle, ranging between 0.39 (October, 2009) and one (May, 1951). Small factors correspond to solar minimum years when disturbed times have smaller absolute HMC values compared to solar maximum years.

2. The local HMC minima of consecutively marked hours (Step 1) are identified by changes in the sign of dH_t/dt. Should two successive local minima be separated by less than Δt hours only the deeper one is kept, so that:

$$t_{i+1} - t_i \geq \Delta t \tag{5.3}$$

3. Only those local minima hours (Step 2) associated with a HMC drop of at least Hs nT are finally selected as target event peak hours $\boldsymbol{t_{ta}}$:

$$|H_t - H_{5d,t}| \geq Hs \tag{5.4}$$

   Here, $\boldsymbol{H_{5d}}$ is the five-day low-pass filtered HMC and a measure of the disturbance background level.

We define the training set as the intersection between the selected target events and the reference events, i.e., $d_{tr} = d_{re} \cap d_{ta}$. The parameters of the selection scheme, $P_{29} = -37.6$ nT, $\Delta t = 27$ hours, and $Hs = 7$ nT, are chosen so that d_{ta} is as small while d_{tr} is as large as possible, i.e., at least 60 % of d_{re}. This guarantees that the target event selection scheme recovers most of the reference events (Figure 5.2) and that the training events are geoeffective at low and mid latitudes, i.e., they produce

a storm-like geomagnetic disturbance with regard to the prevailing activity level on both annual (Equation 5.2) and daily (Equation 5.4) time scales (see also Figure 5.1b). Given the variability in the events' recovery times, there is no perfect choice of Δt (Equation 5.3). We justify our choice of 27 hours by the fact that it is not chosen subjectively, but optimized with regard to the reference set.
The procedure gives $N_{ta} = 7547$ target events, of which $N_{tr} = 538$ overlap with the reference events ($N_{tr0} = 342$ C/SIRs, $N_{tr1} = 196$ ICMEs). The distribution of target events is skewed towards smaller absolute HMC values (median at -50.4 nT). It is similar to the corresponding Dst distribution, but notably different from the relatively symmetric Kp index (Bartels et al., 1939) distribution (Figure 5.2). This is expected given the different latitudinal sensitivity ranges of HMC and Dst on the one hand and Kp on the other hand. The comparison shows that the solar cycle variability is indeed more pronounced for HMC than for Dst. We note that some of the target events are not considered to be geomagnetic storms according to a frequently used minimum-Dst categorization proposed by Loewe and Prölss, (1997, ≈25 %) or the Kp-based NOAA G-scale (`https://www.swpc.noaa.gov/noaa-scales-explanation`, >50 %). However, since our target events produce storm-like geomagnetic disturbances and the majority (≈ 75 %) meets the above Dst-definition (Figure 5.2b) we refer to our target events as storms.
Although ≈38 % of the reference events are not taken into account, the training class ratio $\gamma_{tr} \approx 1.7$ is comparable to that of the reference events ($\gamma_{re} \approx 1.9$). We have thoroughly tested the sensitivity of the classification result to the parameters of the event selection method (Supporting Information, Text S1) and conclude that the presented selection of training and target events is optimal given the available reference events.
In summary, we have $N_{tr} = 538$ training storm events with known inputs and class labels. The inputs are HMC and the MLT-resolved disturbance field ($\boldsymbol{Z}$, $\hat{\boldsymbol{Z}}$). Furthermore, we have $N_{ta} = 7547$ target events with known inputs, but unknown class labels (Supporting Information, Figure 5.S1).

5.3 BINARY LOGISTIC REGRESSION MODEL

5.3.1 *Feature definition*

The core of this work is the identification of driver-characteristic features from the input data. Among the considered possibilities, eleven features are listed according to their relative importance (section 5.3.3) in Table 5.3.1. Their derivation consists of a physically motivated initial feature definition (column 1), followed by a refinement aimed at maximizing the separation of the standardized feature distributions for C/SIR and ICME driven training events (Figure 5.3). Standardization is a common pre-processing step in machine learning and means that the features are centered, i.e., the means are removed, and scaled, i.e., through a division by the standard deviation. The further the class medians are separated from one another and the narrower their interquartile ranges (75th−25th percentiles; Figure 5.3, colored bars) are, the better the feature is. For comparison, we synthetically generated an ideal feature (k=12) by randomly drawing samples from two normal distributions

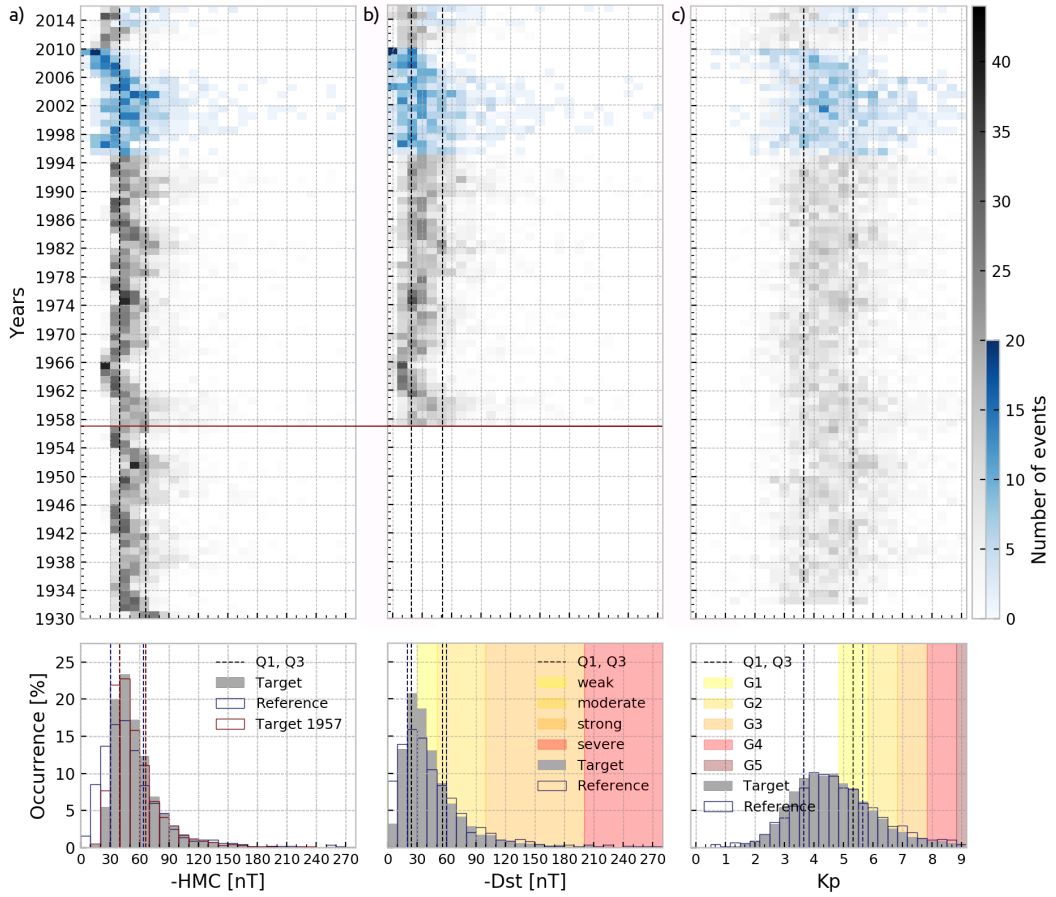

Figure 5.2: *Event histograms.* Distribution of HMC (**a**), Dst (**b**) and Kp (**c**) for 7547 target (gray) and 868 reference (blue) event peaks. Top panels show the number of events per year and index bin. Bottom panels show the occurrence per index bin, summed over all years. The red bar plot in a) refers only to years since 1957 (red horizontal line). For Kp, the largest value in 26-hour intervals centered on the HMC peaks is considered (t_{ta}). Dashed lines mark the 25th (Q1) and 75th (Q3) percentiles. Colored bars indicate commonly used Dst- and Kp-based storm intensity scales.

centered on 0.25 (C/SIRs) and 0.75 (ICMEs), each with a standard deviation of 0.01 (Figure 5.3, bottom right).

Features $k = 1, 4, 5$ describe the well-known ring current enhancement, solar cycle phase, and the recurrence pattern associated with the storms (e.g., Borovsky and Denton, 2006). As expected, ICMEs are statistically more intense and occur closer to solar maximum (Figures 5.1b, 5.3). The deviation of the data-based features from the ideal case is particularly striking for the "Recurrence" feature, for which the medians coincide at 0.5. Nevertheless, the feature still has skill by contributing the information that 75 % of C/SIRs recur at least once, while this is true for just 25 % of ICMEs (Figure 5.3, $k = 5$). The latter poses a problem which is discussed in section 5.5. Of particular interest are the innovative, high-priority features $k = 2, 3$, which are designed to capture PRC-related characteristics of the disturbances at dawn and dusk. Apparently, the disturbances in these MLT-sectors are better correlated

($k = 2$) and the temporal change at dawn is more variable ($k = 3$) for ICMEs than for C/SIRs over a period of 24 hours after the HMC-peak. The former seems to be compatible with the finding of Yakovchouk et al., (2012) that strong storms (mostly CME driven) are relatively more symmetric than weak ones (mostly CIR related). The related feature $k = 10$ describes to what extent feature $k = 2$ changes if the disturbances are time-lagged and is found to be of limited use. Feature $k = 6$ quantifies whether the disturbance peak is located “abnormally", i.e., not in the dusk sector, which is more prevalent for ICMEs. Features $k = 7, 8$ are related to the asymmetry measures (Equation 5.1), while features $k = 9, 11$ describe the rate of change in HMC for the storm main and recovery phases.
Features $k = 1, \ldots, N_f$ form the $N_{tr \vee ta} \times N_f$ matrices $\boldsymbol{X_{tr \vee ta}}$ for the training and target events. Because they require data prior to the storm peaks (e.g., $k = 6$), they are defined for just N_{ta}=7546 out of the 7547 selected target events.

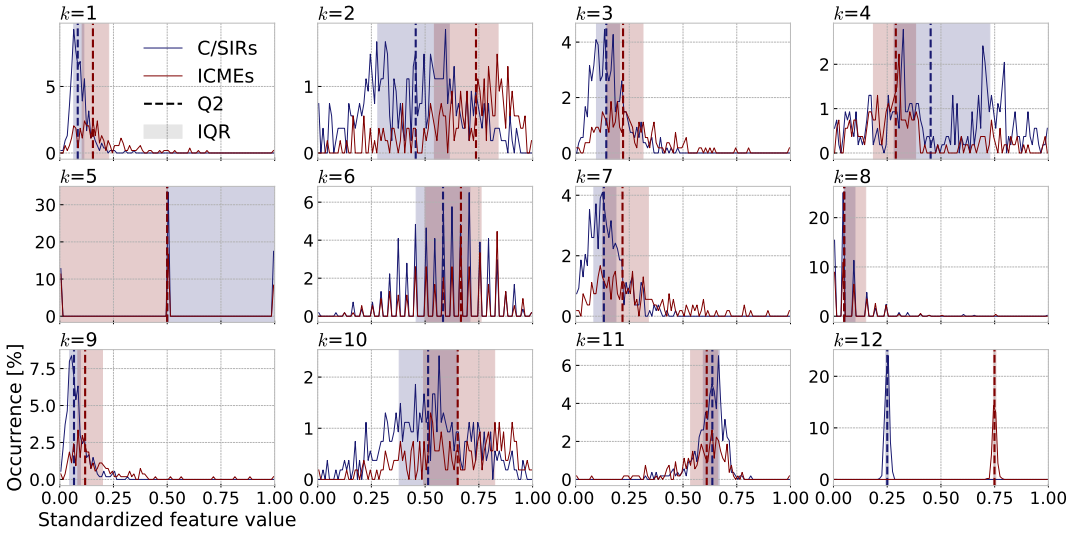

Figure 5.3: *Feature distributions.* Histograms of standardized features for C/SIR- (blue) and ICME-driven (red) training events with interquartile ranges (IQR) and medians (Q2, Table 5.1). Feature k=12 is an ideal synthetic feature. Features are scaled to $[0, 1]$ for visual comparability.

Table 5.1: *Feature definitions.* Features k=1,...,11 sorted according to their importance from top to bottom with medians for C/SIR- and ICME-driven training events ($\tilde{\boldsymbol{X}}_{tr}[\boldsymbol{i}_{tr0}, k]$, $\tilde{\boldsymbol{X}}_{tr}[\boldsymbol{i}_{tr1}, k]$; Figure 5.3) and corresponding model coefficients (α_k; section 5.3.2). **Abbr.** $\boldsymbol{t_{ta}}$: Storm peak times of target events; $\boldsymbol{H_{ny}}$: low-pass filtered HMC with cutoff period of n years; r: Pearson's correlation coefficient; MAD: Median Absolute Deviation; $FDHM$: Full Duration at Half Maximum.

ID (k)	Description	Formal definition	$\tilde{\boldsymbol{X}}_{tr}[\boldsymbol{i}_{tr0}, k]$, $\tilde{\boldsymbol{X}}_{tr}[\boldsymbol{i}_{tr1}, k]$	α_k
1	Event intensity.	$\boldsymbol{\Delta H} = \boldsymbol{H}_{t_{ta}} - \boldsymbol{H}_{\boldsymbol{1y},t_{ta}}$ in nT.	0.08, 0.15	0.63
2	Correlation of disturbances at dawn & dusk.	$r\left(\hat{\boldsymbol{Z}}_{\boldsymbol{6t}}, \hat{\boldsymbol{Z}}_{\boldsymbol{18t}}\right)$ for $\boldsymbol{t} = [\boldsymbol{t_{ta}}, \boldsymbol{t_{ta}} + 24]$.	0.46, 0.74	0.77
3	Variance of disturbance change at dawn.	$MAD\left(d\hat{\boldsymbol{Z}}_{\boldsymbol{6t}}/d\boldsymbol{t}\right)$ for $\boldsymbol{t} = [\boldsymbol{t_{ta}}, \boldsymbol{t_{ta}} + 24]$ in nT/h.	0.14, 0.22	0.63
4	Solar cycle phase.	$\boldsymbol{\phi_{sc}} = d\boldsymbol{H}_{\boldsymbol{11y},t_{ta}}/d\boldsymbol{t_{ta}}$ in nT/h.	0.45, 0.29	-0.42
5	Recurrence.	$\boldsymbol{R}_{t_{ta}} \in [0, 1, 2]$. Either the event 0: does not recur, 1: recurs once with period 27.28 days $\pm$ 26 hours, or 2: recurs twice.	0.5, 0.5	-0.32
6	Robustness of disturbance peak location.	Number of instances $\underset{j}{argmax}\ \lvert Z_{jt}\rvert \leq 17$ for $\boldsymbol{t} = [\boldsymbol{t_{ta}} - 48, \boldsymbol{t_{ta}} - 24]$.	0.58, 0.67	0.28
7	Largest spatial disturbance asymmetry.	$max\left(\boldsymbol{ASY}_{\boldsymbol{max},t}\right)$ for $\boldsymbol{t} = [\boldsymbol{t}_{t_a} - 24, \boldsymbol{t}_{t_a}]$ in nT.	0.13, 0.22	0.31
8	Speed of dawn-dusk asymmetry decrease.	FDHM $\boldsymbol{ASY}_{\boldsymbol{dd},t}$ with respect to $\boldsymbol{t_{ta}}$ in hours.	0.05, 0.05	0.16
9	Mean intensity change for recovery phase.	$\overline{d\boldsymbol{H}_{t_{rec}}/d\boldsymbol{t_{rec}}}$ for $\boldsymbol{t_{rec}} = [argmax\ \lvert d\boldsymbol{H}_{t_p}/d\boldsymbol{t_p}\rvert, \boldsymbol{t_{ta}} + 12]$, $\boldsymbol{t_p} = [\boldsymbol{t_{ta}}, \boldsymbol{t_{ta}} + 12]$ in nT/h.	0.07, 0.12	
10	Range in $k = 2$ for time-shifted disturbances.	Range in $r\left(\hat{\boldsymbol{Z}}_{\boldsymbol{6t_l}}, \hat{\boldsymbol{Z}}_{\boldsymbol{18t}}\right)$ for $\boldsymbol{t} = [\boldsymbol{t_{ta}}, \boldsymbol{t_{ta}} + 24]$, $\boldsymbol{t_l} = \boldsymbol{t} + [-12, \ldots, 12]$.	0.51, 0.65	
11	Difference in intensity changes for main & recovery phases.	$\overline{d\boldsymbol{H}_{t_{rec}}/d\boldsymbol{t_{rec}}} - \overline{d\boldsymbol{H}_{t_{main}}/d\boldsymbol{t_{main}}}$ for $\boldsymbol{t_{main}} = [\boldsymbol{t_{ta}} - 12, argmax\ \lvert d\boldsymbol{H}_{t_p}/d\boldsymbol{t_p}\rvert]$, $\boldsymbol{t_p} = [\boldsymbol{t_{ta}} - 12, \boldsymbol{t_{ta}}]$ in nT/h.	0.64, 0.61	

5.3.2 *Model setup*

The binary logistic regression model takes a linear combination of the features, $\boldsymbol{X_i}$, and delivers the logarithm of the odds ("Logit") that the storm event i is driven by an ICME (Equation 5.5, default class 1). The odds are defined as the probability $P(Y_i = 1|\boldsymbol{X_i}) \in [0,1]$, divided by the complementary probability $1 - P(Y_i = 1|\boldsymbol{X_i})$.

$$\text{Logit}\left[P(Y_i = 1|\boldsymbol{X_i})\right] = \ln\left(\frac{P(Y_i = 1|\boldsymbol{X_i})}{1 - P(Y_i = 1|\boldsymbol{X_i})}\right) = \alpha_0 + \boldsymbol{X_i^T}\boldsymbol{\alpha} \tag{5.5}$$

$$P(Y_i = 1|\boldsymbol{X_i}) = \frac{1}{1 + \exp\left[-\left(\alpha_0 + \boldsymbol{X_i^T}\boldsymbol{\alpha}\right)\right]} \tag{5.6}$$

The model is trained by using $\boldsymbol{X_{tr}}$ and known class labels $\boldsymbol{Y_{tr}}$ to solve Equation 5.5 for the intercept $\alpha_0 = -0.17$ and the coefficients $\alpha_k, k = 1\ldots, N_f$ (Table 5.1). The class ratio γ_{tr} is balanced by applying weights inversely proportional to the class frequencies, i.e., $N_{tr}/\left(2 \cdot N_{tr0}\right) = 0.79$ for C/SIRs and $N_{tr}/\left(2 \cdot N_{tr1}\right) = 1.37$ for ICMEs. We choose the incremental gradient algorithm "SAGA" (Defazio et al., 2014), because it is the most versatile solver available in Python's scikit-learn library (`https://scikit-learn.org/`) that supports sparse regression and a potential future extension to multinomial cases.

5.3.3 *Model assessment*

The most important features and optimum model hyperparameters are identified with respect to Matthew's correlation coefficient (MCC):

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP) \cdot (TP + FN) \cdot (TN + FP) \cdot (TN + FN)}} \in [-1, 1] \tag{5.7}$$

MCC, also known as the ϕ-coefficient (e.g., Parker, 2013), is particularly useful for data sets with imbalanced classes as it takes into account both True (T) and False (F) predictions of Negatives (N, $\boldsymbol{Y_{tr}}[\boldsymbol{i_{tr0}}]$: C/SIRs) and Positives (P, $\boldsymbol{Y_{tr}}[\boldsymbol{i_{tr1}}]$: ICMEs). A stratified K-fold cross-validation scheme is used (Figure 5.4a), in which the training set is split into K_o=4 folds, so that each consists of 134–135 data points with preserved driver class percentages (outer loop). The hyperparameters of the model are initialized with default values. The classifier is recursively trained on three of the folds, so that the feature with the lowest absolute α_k is pruned from $\boldsymbol{X_{tr}}$ at each iteration, until $N_f = 1, \ldots, 11$ features are left. For each of these 11 runs, *MCC* is calculated four times on the different validation and training sets and then averaged over the sets ($MCC_{o,l}$) and the $l = 1, \ldots, N_o$=25 splits. Figure 5.4b shows that $N_f = 8$ features, namely features $k = 1, \ldots, 8$ (Table 5.1), give the highest *MCC*, so that the feature matrices are fixed accordingly. We note that the training scores confirm this result, but that the absolute *MCC* values have no meaning yet, since they refer to an interim model.

We optimize two of the hyperparameters, which are the inverse regularization strength C (default: $C = 1$) and the norm used in the penalization L (default: L = L2 norm). An inner loop is introduced to the cross-validation scheme with K_i=3 and $N_i = 25$ (Figure 5.4a) in order to prevent an artificial overestimation

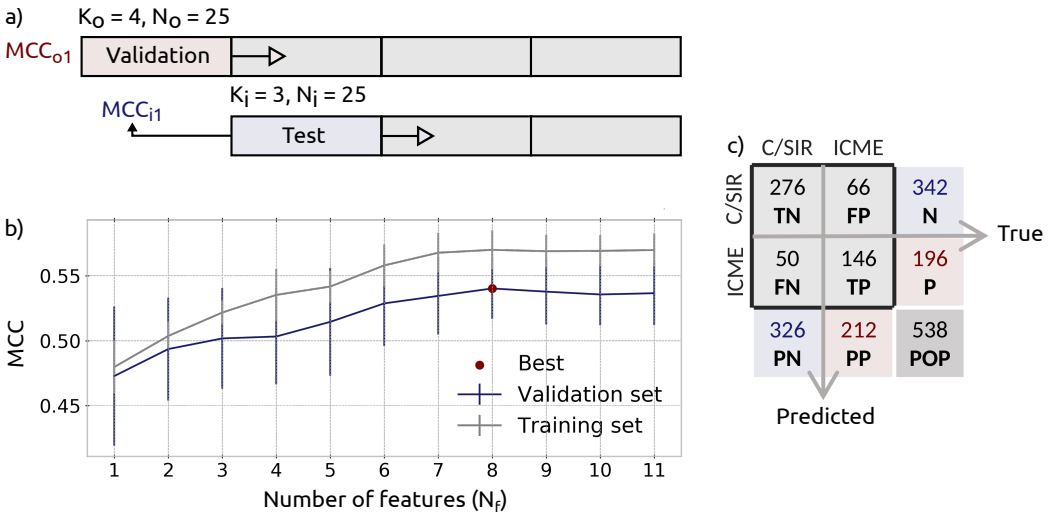

Figure 5.4: *Model assessment*. **a)** K-fold cross-validation scheme on the training set (gray) with four folds in the outer and three folds in the inner loop, each repeated 25 times. **b)** Cross-validated *MCC* with uncertainty (two standard deviations) from models using the N_f best features. **c)** Confusion matrix from final model. **Abbr.** *MCC*: Matthew's correlation coefficient; *P*: Positives; *N*: Negatives; *POP*=*N*+*P*: Population; *T*: True; *F*: False; *P*[*P*/*N*] = Predicted *P*/*N*.

of *MCC* (e.g., Krstajic et al., 2014). *MCC* is calculated for ten combinations of hyperparameters $C \in \{1, 0.1, 0.01, 1e-3, 1e-4\}$ and $L \in \{\text{L1-norm}, \text{L2-norm}\}$ on the test sets of the inner loop (7500 scores each) and averaged over the inner sets/splits (MCC_i) and the outer sets/splits. The combination of $C = 0.1$ and $L =$ L2-norm gives the highest averaged *MCC* (Supporting Information, Figure 5.S2) and is thus fixed in the final model, which is assessed on the validation sets of the outer loop.

From the confusion matrix of the final model (Figure 5.4c) a cross-validated *MCC* score of 0.54±0.013 is calculated, which means that the classifier is skillful and performs significantly better than chance ($MCC = 0$). For completeness and in order to facilitate comparisons, other frequently cited scores are listed in Appendix B.

Apart from the classifier's skill, we can also assess its probabilistic predictions (Equation 5.6), i.e., its reliability. By calculating that fraction of events with a certain predicted ICME-driver probability that are truly ICME-driven, we find that the classifier's probabilistic predictions are slightly overestimated, but reliable within a tolerance of 15 % (Supporting Information, Figure 5.S3b,c).

5.4 CLASSIFICATION RESULT

In the prediction step, the model coefficients α_0, $\boldsymbol{\alpha}$ (section 5.3.2, Table 5.1), and the feature matrix $\boldsymbol{X_{ta}}$ are inserted into Equation 5.6 to give the ICME-driver probabilities $P(Y_{ta,i} = 1)$ for the target events. The class labels, $\boldsymbol{Y_{ta}}$, are set according to the decision rule

$$Y_{ta,i} = \begin{cases} 0 & \text{if } P(Y_{ta,i} = 1) \leq 0.5, \\ 1 & \text{otherwise.} \end{cases} \tag{5.8}$$

Applying the default decision boundary $P(Y_{ta,i} = 1) = 0.5$ is appropriate, because it delivers the best classification result as judged by the mean of several performance scores (Supporting Information, Figure 5.S3a).
The ICME driver probabilities are color coded in Figure 5.5a (left), which groups the target events in a “time plane" spanned by the decimal year and the solar rotation phase. In this view, CIR-driven events are vertically aligned. Of all target events, ≈63 % are driven by C/SIRs, giving a class ratio $\gamma_{ta} \approx 1.7$. On average, the uncertainties of the predictions are similar for both drivers with median class probabilities $\tilde{P}_0 = 1 - \tilde{P}(\boldsymbol{Y_{ta}}[\boldsymbol{i_{ta0}}] = 1) = 0.74$ for C/SIRS and $\tilde{P}_1 = \tilde{P}(\boldsymbol{Y_{ta}}[\boldsymbol{i_{ta1}}] = 1) = 0.73$ for ICMEs.
Throughout time, ICME-dominated periods (red) alternate with C/SIRs-dominated ones (blue), while γ_{ta} seems to increase from the past to the present (cf. Supporting Information, Figure 5.S4). The stacked histograms of ten probability intervals $P(Y_{ta,i} = 1) \in [0.0, 0.1), \ldots, [0.9, 1.0]$ per year (Figure 5.5a, right) accentuate that the occurrence pattern is tightly linked to the solar cycle (SC). The latter is manifest in either the 11-year low-pass filtered HMC (dashed line) or in the independent yearly mean total sunspot number SN (solid line). Both these time series are normalized, i.e., divided by their largest absolute value. On average, ICME drivers are most frequent at SC maximum years, while C/SIRs prevail during the declining phase and SC minimum years. The dependency of the driver occurrence statistics on the intensity of the storm events and the solar cycle phase is shown in Figure 5.5b. Five intervals of HMC values between −16 nT (weak) and −478 nT (strong) are defined, such that they include a roughly comparable number of events. Since the HMC distribution of target events is strongly skewed towards smaller absolute values (Figure 5.2a), the intervals get wider as the event intensity increases. Nevertheless, the number of events per class still decreases as intensity increases, especially for solar minimum years (upper row). We define “solar minimum" years as years for which SN≤70 and all other years as “solar maximum years". This threshold was chosen, so that solar minima (40 years) and maxima (46 years) are roughly balanced and the transitions visually coincide with ascending or decreasing flanks of SN (Figure 5.5a, black horizontal lines). As can be seen in Figure 5.5b, ICMEs predominantly drive intense geomagnetic storms and their fraction is generally larger at solar maximum years.

5.5 DISCUSSION

The binary classifier we developed can be seen as a proof of concept, demonstrating that it is possible to identify storm drivers from ground-level data without the need for satellite observations. However, the performance of the classifier could probably be improved technically, i.e., without changing the fundamental type of input information. First, there is much more geomagnetic data available from mid latitudes than we use, especially for recent years, which can improve the coverage of MLTs (blue dots in Figure 5.1d). Our choice is motivated by the fact

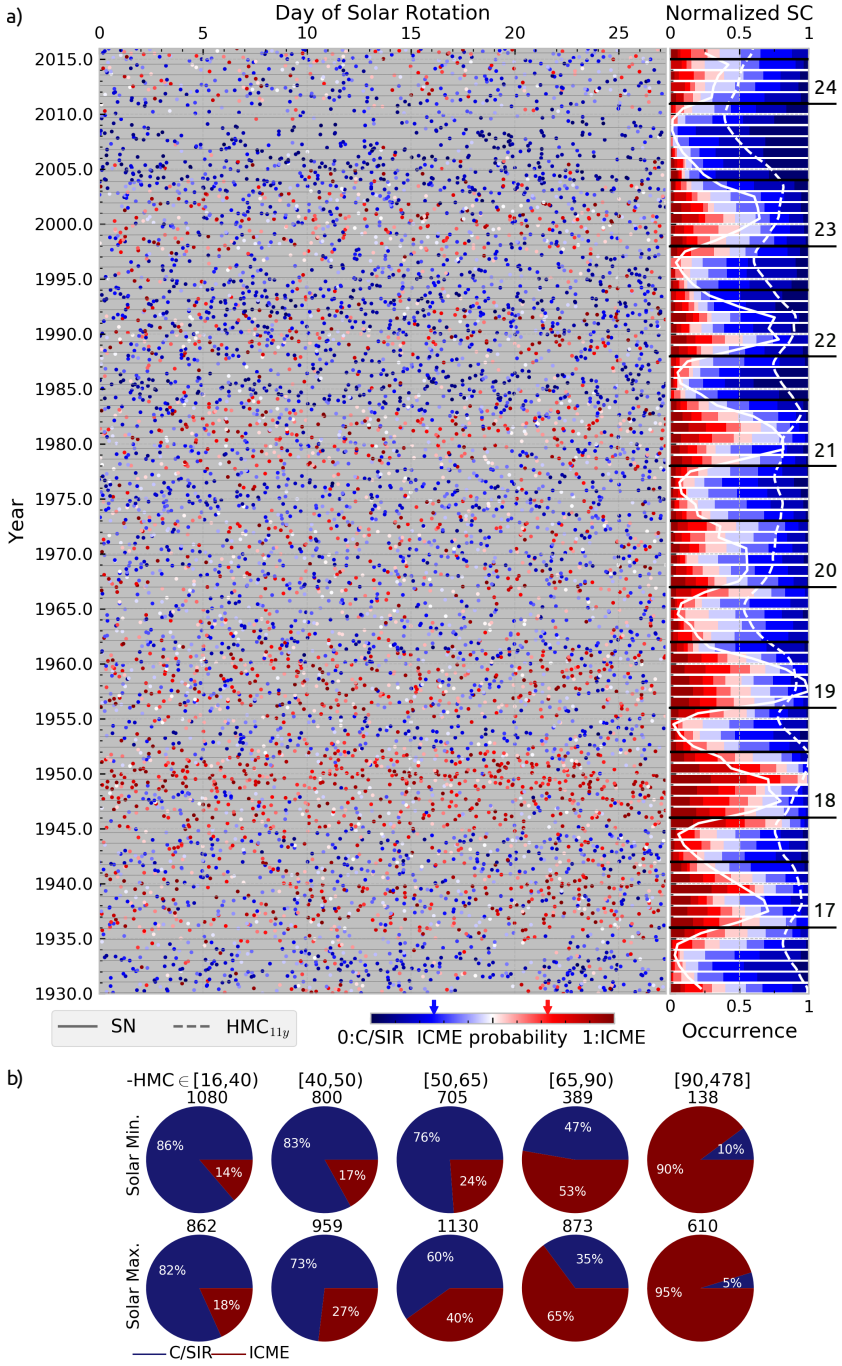

Figure 5.5: *Classification result.* **a)** Left: Probabilities $P(Y_{ta,i}{=}1)$ for target events in a year vs. day of solar rotation plot. Median class probabilities $\tilde{P}_0{=}0.74$, $\tilde{P}_1{=}0.73$ are marked on the color bar. Right: Stacked histograms for 0.1 probability intervals per year with normalized sunspot number (SN, white solid) and 11-year low-pass filtered HMC (dashed) representing the solar cycle (SC; numbers in black). **b)** Driver fractions during solar min. (top) and max. years (bottom), as separated by horizontal lines in the histograms, grouped into five intensity levels defined by HMC [nT]. Group definitions and event numbers are given above the pie plots.

that it is fairly stable throughout time, we have assured a decent quality of the observatory time series during early years and it is consistent with the input data used for HMC. Secondly, the definition of features could be avoided altogether by using data-interpolating empirical orthogonal functions to decompose the data set into spatiotemporal patterns, that replace the features (e.g., Holappa et al., 2015, 2014; Shore et al., 2018). Lastly, more sophisticated machine learning techniques could be applied, e.g., Support Vector Machines, Random Forest Classifiers or Neural Networks, which would however preclude a simple ranking of features. As opposed to the aforementioned points, the following issues are relevant to the scientific interpretability of our results.

5.5.1 *Driver class statistics*

Although the time period from which we derive the geomagnetic driver statistics exceeds that of previous studies by roughly three solar cycles (e.g., Echer et al., 2011; Richardson and Cane, 2012), the C/SIR vs. ICME occurrence statistics are in general agreement (see section 5.4). This supports the validity of our model beyond the mathematical perspective taken in section 5.3.3. However, an analysis of the confusion matrix (Figure 5.4c) reveals that the model suffers from False Negatives (FNs), i.e., the misidentification of ICMEs as C/SIRs, more than from False Positives (FPs). A decrease of FNs would lower the rate of misses ($\sim$26 %) and false discoveries ($\sim$31 %), and thereby improve MCC $\approx$1.3 times more than the same decrease of FPs.

The validity of the classifier is probably challenged most by the weakest events of the target distribution. Owing to feature $k = 1$ (Table 5.1), it is likely that for these events (1) slow solar wind or unclear drivers (as considered in Richardson and Cane, 2012) are erroneously forced into the C/SIR class and (2) ICME drivers are misidentified (FNs). There are several instances of decreases in ICME-driven storm activity near SC maximum (e.g., cycle 20), which could either be due to such FNs, or to an actual phenomenon related to the “Gnevyshev gap" (Feminella and Storini, 1997), or a mixture of both.

Additionally, weaknesses in the features’ definitions mitigate the classifier’s skill. Feature $k = 4$ (“SC phase") is defined as the derivative of the low-passed filtered HMC at $\boldsymbol{t_{ta}}$, which lags behind SN (HMC_{11y}, Figure 5.5a). Therefore, driver preferences for solar cycle phases are “blurred", such that negative values represent both SN maxima, favorable for ICMEs, and part of the declining phase, favorable for C/SIRs. Feature $k = 5$ (“Recurrence") gives information on whether or not an event recurs in the target set. There is no mechanism to prevent a positive response for ICME driven events, in which case they are “unphysically" labeled as recurring. On the other hand, SIRs are likely to be confused as ICMEs since both are non-recurring. Although overall less effective, splitting the C/SIR class into two would likely reduce FPs.

Finally, the solar cycle dependence of the drivers is reproduced even if the classification is done without features $k = 4, 5$ (Supporting Information, Figure 5.S5). However, these results show significantly reduced median class probabilities

(≈ 0.6), underlining that, despite their weaknesses, both features contribute to the classifier's skill.

5.5.2 *Geoeffectiveness of training events*

One aspect not explicitly taken into account is the mechanism and efficiency with which the interplanetary drivers of the training set couple to the magnetosphere. Thus, there is no guarantee that each solar wind structure identified at L1 is geoeffective at mid latitudes, i.e., the region to which our analysis is sensitive. This explains, why $\approx 38\%$ of the reference events could not be included in the training set (see section 5.2.2). Amongst other criteria involving, e.g., $\boldsymbol{v_{sw}}$, a southward directed $B_{sw,z}$ is decisive for the storm generation (see section 5.1). ICMEs with clear signatures of MCs often comply with this favorable condition, making them the most geoeffective subset of ICMEs (e.g., Nikolaeva et al., 2011). Other cases include instances where the spacecraft misses the centre of the flux rope or there are successive ICMEs interacting with each other (complex ejecta). Similarly, C/SIRs driving shock waves at their leading edges and followed by Alfvénic high speed streams are more geoeffective than others (Snekvik et al., 2013; Zhang et al., 2008).

Since HMC is particularly sensitive to the ring current, the degree of HMC disturbance depends on that of ring current build-up, resulting from the above mentioned geoeffective drivers. The latter is controlled by the plasma sheet density (e.g., Jordanova et al., 2009) and ring current composition, particularly the concentration of oxygen ions (e.g., Denton et al., 2006).
Consequently, one could enhance the physical expressiveness of the study by selecting a subset of training events showing unambiguous C/SIR and ICME signatures, that promote ring current build-up. However, the price of such action is a reduced training set, less appropriate for machine learning methods. On the other hand, one could abandon the attempt to link ground-level disturbances to specific magnetospheric current systems and open the study to input data from high-latitude magnetometers sensitive to different kinds of current systems and disturbances (e.g., substorms). Here, we have steered a middle course on the basis of which both approaches can be further developed.

5.6 CONCLUSIONS

Our binary classifier reliably determines how likely a given geomagnetic storm is driven by an ICME as opposed to a C/SIR. The inputs (features) for the classifier are derived solely from the MLT-resolved geomagnetic disturbance field at ground level. Among them are two high-priority features, that have not been formulated previously in this context. We provide the driver probabilities for 7546 storms between 1930–2015, whose statistics in dependence of event intensity and solar activity are in general agreement with previous satellite-based studies covering about four solar cycles. Our results can be taken as a proof of concept for a further exploitation of geomagnetic observatory data with regard to an interplanetary

driver identification. To this end, several ideas for future refinement or modification of the presented method are discussed.

ACKNOWLEDGMENTS

The authors thank Yuri Shprits, Joachim Vogt and Robert Shore for useful discussions and two anonymous reviewers for helping to improve the manuscript. LP and MK acknowledge funding by the Deutsche Forschungsgemeinschaft under project SPP1788 "Dynamic Earth" (DFG KO 2870/6-1). FE acknowledges support from a group linkage program with South Africa, funded by the Alexander von Humboldt Foundation. IZ was supported by Geo.X, the Research Network for Geosciences in Berlin and Potsdam, under grant no. SO_087_GeoX. The sunspot number (`http://www.sidc.be/silso/datafiles`) and Dst/Kp/SYM-H indices (`http://wdc.kugi.kyoto-u.ac.jp/`) are publicly available. All input data, the results, and the corresponding software can be accessed via GFZ Data Services (Pick, 2019, `http://doi.org/10.5880/GFZ.2.3.2019.003`).

REFERENCES

Akasofu, S.-I. and S. Chapman (1964). "On the asymmetric development of magnetic storm fields in low and middle latitudes." In: *Planetary and Space Science* 12.6, pp. 607–626. DOI: 10.1016/0032-0633(64)90008-X (cit. on p. 78).

Bartels, J., N. H. Heck, and H. F. Johnston (1939). "The three-hour-range index measuring geomagnetic activity." In: *Terrestrial Magnetism and Atmospheric Electricity* 44.4, pp. 411–454. DOI: 10.1029/TE044i004p00411 (cit. on p. 83).

Borovsky, J. E. and M. H. Denton (2006). "Differences between CME-driven storms and CIR-driven storms." In: *Journal of Geophysical Research: Space Physics* 111.A07S08. DOI: 10.1029/2005JA011447 (cit. on pp. 78, 84).

Defazio, A., F. Bach, and S. Lacoste-Julien (2014). "SAGA: A Fast Incremental Gradient Method With Support for Non-Strongly Convex Composite Objectives." In: *Advances in Neural Information Processing Systems*. Ed. by Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger. Vol. 27. Montreal, Quebec: Curran Asociates, Inc., pp. 1646–1654 (cit. on p. 87).

Denton, M. H. et al. (2006). "Geomagnetic storms driven by ICME- and CIR-dominated solar wind." In: *Journal of Geophysical Research: Space Physics* 111, A07S07. DOI: 10.1029/2005JA011436 (cit. on p. 92).

Dubyagin, S., N. Ganushkina, M. Kubyshkina, and M. Liemohn (2014). "Contribution from different current systems to SYM and ASY midlatitude indices." In: *Journal of Geophysical Research: Space Physics* 119.9, pp. 7243–7263. DOI: 10.1002/2014JA020122 (cit. on p. 78).

Dumbović, M., B. Vršnak, J. Čalogović, and R. Župan (2012). "Cosmic ray modulation by different types of solar wind disturbances." In: *Astronomy & Astrophysics* 538, A28. DOI: 10.1051/0004-6361/201117710 (cit. on p. 79).

Echer, E., W. Gonzalez, and B. Tsurutani (2011). "Statistical studies of geomagnetic storms with peak Dst $\leq$ -50 nT from 1957 to 2008." In: *Journal of Atmospheric and*

Solar-Terrestrial Physics 73.11, pp. 1454–1459. DOI: 10.1016/j.jastp.2011.04.021 (cit. on p. 91).

Feldman, W. C., J. R. Asbridge, S. J. Bame, E. E. Fenimore, and J. T. Gosling (1981). "The solar origins of solar wind interstream flows: Near-equatorial coronal streamers." In: *Journal of Geophysical Research: Space Physics* 86.A7, pp. 5408–5416. DOI: 10.1029/JA086iA07p05408 (cit. on p. 78).

Feminella, F. and M. Storini (1997). "Large-scale dynamical phenomena during solar activity cycles." In: *Astronomy & Astrophysics* 322, pp. 311–319 (cit. on p. 91).

Gosling, J. T. (1993). "The solar flare myth." In: *Journal of Geophysical Research: Space Physics* 98.A11, pp. 18937–18949. DOI: 10.1029/93JA01896 (cit. on p. 78).

Gosling, J. T. et al. (1975). "Direct observations of a flare related coronal and solar wind disturbance." In: *Solar Physics* 40.2, pp. 439–448. DOI: 10.1007/BF00162390 (cit. on p. 78).

Greaves, W. M. H. and H. W. Newton (1929). "On the Recurrence of Magnetic Storms. (Plate 9.)" In: *Monthly Notices of the Royal Astronomical Society* 89.7, pp. 641–646. DOI: 10.1093/mnras/89.7.641 (cit. on p. 78).

Holappa, L., K. Mursula, and T. Asikainen (2015). "A new method to estimate annual solar wind parameters and contributions of different solar wind structures to geomagnetic activity." In: *Journal of Geophysical Research: Space Physics* 119.12, pp. 9407–9418. DOI: 10.1002/2014JA020599 (cit. on p. 91).

Holappa, L., K. Mursula, T. Asikainen, and I. G. Richardson (2014). "Annual fractions of high-speed streams from principal component analysis of local geomagnetic activity." In: *Journal of Geophysical Research: Space Physics* 119.6, pp. 4544–4555. DOI: 10.1002/2014JA019958 (cit. on p. 91).

Iyemori, T. (1990). "Storm-Time Magnetospheric Currents Inferred from Mid-Latitude Geomagnetic Field Variations." In: *Journal of geomagnetism and geoelectricity* 42.11, pp. 1249–1265. DOI: 10.5636/jgg.42.1249 (cit. on pp. 78, 98).

Jian, L., C. T. Russell, and J. G. Luhmann (2011). "Comparing Solar Minimum 23/24 with Historical Solar Wind Records at 1 AU." In: *Solar Physics* 274.1, pp. 321–344. DOI: 10.1007/s11207-011-9737-2 (cit. on pp. 82, 103).

Jian, L., C. T. Russell, J. G. Luhmann, and R. M. Skoug (2006a). "Properties of Interplanetary Coronal Mass Ejections at One AU During 1995 – 2004." In: *Solar Physics* 239.1, pp. 393–436. DOI: 10.1007/s11207-006-0133-2 (cit. on pp. 78, 82, 103).

– (2006b). "Properties of Stream Interactions at One AU During 1995 – 2004." In: *Solar Physics* 239.1, pp. 337–392. DOI: 10.1007/s11207-006-0132-3 (cit. on pp. 78, 82, 103).

Jordanova, V. K. et al. (2009). "Ring current development during high speed streams." In: *Journal of Atmospheric and Solar-Terrestrial Physics* 71.10, pp. 1093–1102. DOI: 10.1016/j.jastp.2008.09.043 (cit. on p. 92).

Kamide, Y. and N. Fukushima (1971). "Analysis of Magnetic Storms with DR Indices for Equatorial Ring-Current Field." In: *Radio Science* 6.2, pp. 277–278. DOI: 10.1029/RS006i002p00277 (cit. on p. 78).

Kilpua, E. K. J., A. Balogh, R. von Steiger, and Y. D. Liu (2017). "Geoeffective Properties of Solar Transients and Stream Interaction Regions." In: *Space Science Reviews* 212.3, pp. 1271–1314. DOI: 10.1007/s11214-017-0411-3 (cit. on p. 78).

Krieger, A. S., A. F. Timothy, and E. C. Roelof (1973). "A coronal hole and its identification as the source of a high velocity solar wind stream." In: *Solar Physics* 29.2, pp. 505–525. DOI: 10.1007/BF00150828 (cit. on p. 78).

Krstajic, D., L. J. Buturovic, D. E. Leahy, and S. Thomas (2014). "Cross-validation pitfalls when selecting and assessing regression and classification models." In: *Journal of Cheminformatics* 6.10. DOI: 10.1186/1758-2946-6-10 (cit. on p. 88).

Laundal, K. M. and A. D. Richmond (2017). "Magnetic Coordinate Systems." In: *Space Science Reviews* 206.1, pp. 27–59. DOI: 10.1007/s11214-016-0275-y (cit. on p. 79).

Liemohn, M. W. et al. (2018). "Model Evaluation Guidelines for Geomagnetic Index Predictions." In: *Space Weather* 16.12, pp. 2079–2102. DOI: 10.1029/2018SW002067 (cit. on p. 99).

Loewe, C. A. and G. W. Prölss (1997). "Classification and mean behavior of magnetic storms." In: *Journal of Geophysical Research: Space Physics* 102.A7, pp. 14209–14213. DOI: 10.1029/96JA04020 (cit. on p. 83).

Love, J. J. and J. L. Gannon (2009). "Revised D_{st} and the epicycles of magnetic disturbance: 1958–2007." In: *Annales Geophysicae* 27.8, pp. 3101–3131. DOI: 10.5194/angeo-27-3101-2009 (cit. on p. 80).

Lühr, H. and S. Maus (2010). "Solar cycle dependence of quiet-time magnetospheric currentsand a model of their near-Earth magnetic fields." In: *Earth, Planets and Space* 62, pp. 843–848. DOI: 10.5047/eps.2010.07.012 (cit. on p. 79).

Martini, D., K. Mursula, M. Orispää, and H.-J. Linthe (2015). "Long-term decrease in the response of midlatitude stations to high-speed solar wind streams in 1914–2000." In: *Journal of Geophysical Research: Space Physics* 120.4, pp. 2662–2674. DOI: 10.1002/2014JA020813 (cit. on p. 79).

Maunder, E. W. (1904). "Magnetic Disturbances, 1882 to 1903, as recorded at the Royal Observatory, Greenwich, and their Association with Sun-spots." In: *Monthly Notices of the Royal Astronomical Society* 65.1, pp. 2–18. DOI: 10.1093/mnras/65.1.2 (cit. on p. 78).

Mursula, K., D. Martini, and A. Karinen (2004). "Did Open Solar Magnetic Field Increase During The Last 100 Years? A Reanalysis of Geomagnetic Activity." In: *Solar Physics* 224.1, pp. 85–94. DOI: 10.1007/s11207-005-4981-y (cit. on p. 79).

Newell, P. T. and J. W. Gjerloev (2012). "SuperMAG-based partial ring current indices." In: *Journal of Geophysical Research: Space Physics* 117.A5. DOI: 10.1029/2012JA017586 (cit. on p. 80).

Nikolaeva, N. S., Y. I. Yermolaev, and I. G. Lodkina (2011). "Dependence of geomagnetic activity during magnetic storms on the solar wind parameters for different types of streams." In: *Geomagnetism and Aeronomy* 51.1, pp. 49–65. DOI: 10.1134/S0016793211010099 (cit. on p. 92).

Oughton, E. J. et al. (2018). "A Risk Assessment Framework for the Socioeconomic Impacts of Electricity Transmission Infrastructure Failure Due to Space Weather: An Application to the United Kingdom." In: *Risk Analysis* 39.5, pp. 1022–1043. DOI: 10.1111/risa.13229 (cit. on p. 78).

Parker, C. (2013). "On measuring the performance of binary classifiers." In: *Knowledge and Information Systems* 35.1, pp. 131–152. DOI: 10.1007/s10115-012-0558-x (cit. on p. 87).

Pick, L. (2019). "ClassifyStorms - an automatic classifier for geomagnetic storm drivers based on machine learning techniques. V. 1.0.1." In: *GFZ Data Services*. DOI: 10.5880/GFZ.2.3.2019.003 (cit. on p. 93).

Pick, L. and M. Korte (2018). "HMC index." In: *GFZ Data Services*. DOI: 10.5880/GFZ.2.3.2018.006 (cit. on p. 79).

Pick, L., M. Korte, Y. Thomas, N. Krivova, and C.-J. Wu (2019). "Evolution of large-scale magnetic fields from near-Earth space during the last 11 solar cycles." In: *Journal of Geophysical Research: Space Physics* 124.4, pp. 2527–2540. DOI: 10.1029/2018JA026185 (cit. on p. 79).

Richardson, I. G., E. W. Cliver, and H. V. Cane (2000). "Sources of geomagnetic activity over the solar cycle: Relative importance of coronal mass ejections, high-speed streams, and slow solar wind." In: *Journal of Geophysical Research: Space Physics* 105.A8, pp. 18203–18213. DOI: 10.1029/1999JA000400 (cit. on p. 82).

Richardson, I. G. and H. V. Cane (2012). "Solar wind drivers of geomagnetic storms during more than four solar cycles." In: *Journal of Space Weather and Space Climate* 2, A01. DOI: 10.1051/swsc/2012001 (cit. on pp. 79, 91, 101).

Rostoker, G. and C.-G. Fälthammar (1967). "Relationship between changes in the interplanetary magnetic field and variations in the magnetic field at the Earth's surface." In: *Journal of Geophysical Research* 72.23, pp. 5853–5863. DOI: 10.1029/JZ072i023p05853 (cit. on p. 78).

Shen, X.-C. et al. (2017). "Statistical study of the storm time radiation belt evolution during Van Allen Probes era: CME- versus CIR-driven storms." In: *Journal of Geophysical Research: Space Physics* 122.8, pp. 8327–8339. DOI: 10.1002/2017JA024100 (cit. on pp. 82, 103).

Shore, R. M., M. P. Freeman, and J. W. Gjerloev (2018). "An Empirical Orthogonal Function Reanalysis of the Northern Polar External and Induced Magnetic Field During Solar Cycle 23." In: *Journal of Geophysical Research: Space Physics* 123.1, pp. 781–795. DOI: 10.1002/2017JA024420 (cit. on p. 91).

Siscoe, G. L., J. J. Love, and J. L. Gannon (2012). "Problem of the Love-Gannon relation between the asymmetric disturbance field and Dst." In: *Journal of Geophysical Research: Space Physics* 117, A09216. DOI: 10.1029/2012JA017879 (cit. on p. 80).

Snekvik, K., E. I. Tanskanen, and E. K. J. Kilpua (2013). "An automated identification method for Alfvénic streams and their geoeffectiveness." In: *Journal of Geophysical Research: Space Physics* 118.10, pp. 5986–5998. DOI: 10.1002/jgra.50588 (cit. on p. 92).

Snyder, C. W., M. Neugebauer, and U. R. Rao (1963). "The solar wind velocity and its correlation with cosmic-ray variations and with solar and geomagnetic activity." In: *Journal of Geophysical Research* 68.24, pp. 6361–6370. DOI: 10.1029/JZ068i024p06361 (cit. on p. 78).

Sugiura, M. (1964). "Hourly values of equatorial Dst for the IGY." In: *Annales of the International Geophysical Year* 35, pp. 4–45 (cit. on p. 78).

Temerin, M. and X. Li (2015). "The Dst index underestimates the solar cycle variation of geomagnetic activity." In: *Journal of Geophysical Research: Space Physics* 120.7, pp. 5603–5607. DOI: 10.1002/2015JA021467 (cit. on p. 79).

Thébault, E. et al. (2015). "International Geomagnetic Reference Field: the 12th generation." In: *Earth, Planets and Space* 67.79. DOI: 10.1186/s40623-015-0228-9 (cit. on p. 81).

Turner, N. E., W. D. Cramer, S. K. Earles, and B. A. Emery (2009). "Geoefficiency and energy partitioning in CIR-driven and CME-driven storms." In: *Journal of Atmospheric and Solar-Terrestrial Physics* 71.10–11, pp. 1023–1031. DOI: 10.1016/j.jastp.2009.02.005 (cit. on pp. 82, 103).

Wilcox, J. M. and N. F. Ness (1965). "Quasi-stationary corotating structure in the interplanetary medium." In: *Journal of Geophysical Research* 70.23, pp. 5793–5805. DOI: 10.1029/JZ070i023p05793 (cit. on p. 78).

Woodcock, F. (1976). "The Evaluation of Yes/No Forecasts for Scientific and Administrative Purposes." In: *Monthly Weather Review* 104.10, pp. 1209–1214. DOI: 10.1175/1520-0493(1976)104<1209:TEOYFF>2.0.CO;2 (cit. on p. 99).

Yakovchouk, O. S., K. Mursula, L. Holappa, I. S. Veselovsky, and A. Karinen (2012). "Average properties of geomagnetic storms in 1932–2009." In: *Journal of Geophysical Research: Space Physics* 117.A3. DOI: 10.1029/2011JA017093 (cit. on pp. 80, 85).

Zhang, Y. et al. (2008). "Statistical analysis of corotating interaction regions and their geoeffectiveness during solar cycle 23." In: *Journal of Geophysical Research: Space Physics* 113.A8. DOI: 10.1029/2008JA013095 (cit. on p. 92).

APPENDIX A: ILLUSTRATION OF EVENT SELECTION SCHEME

Figure 5.A1 illustrates the event selection scheme by the example of September 1981, when Dst, SYM-H (Iyemori, 1990) and Kp indices are available for comparison. The offset between HMC and Dst results from Dst's erroneous baseline (see section 5.1).

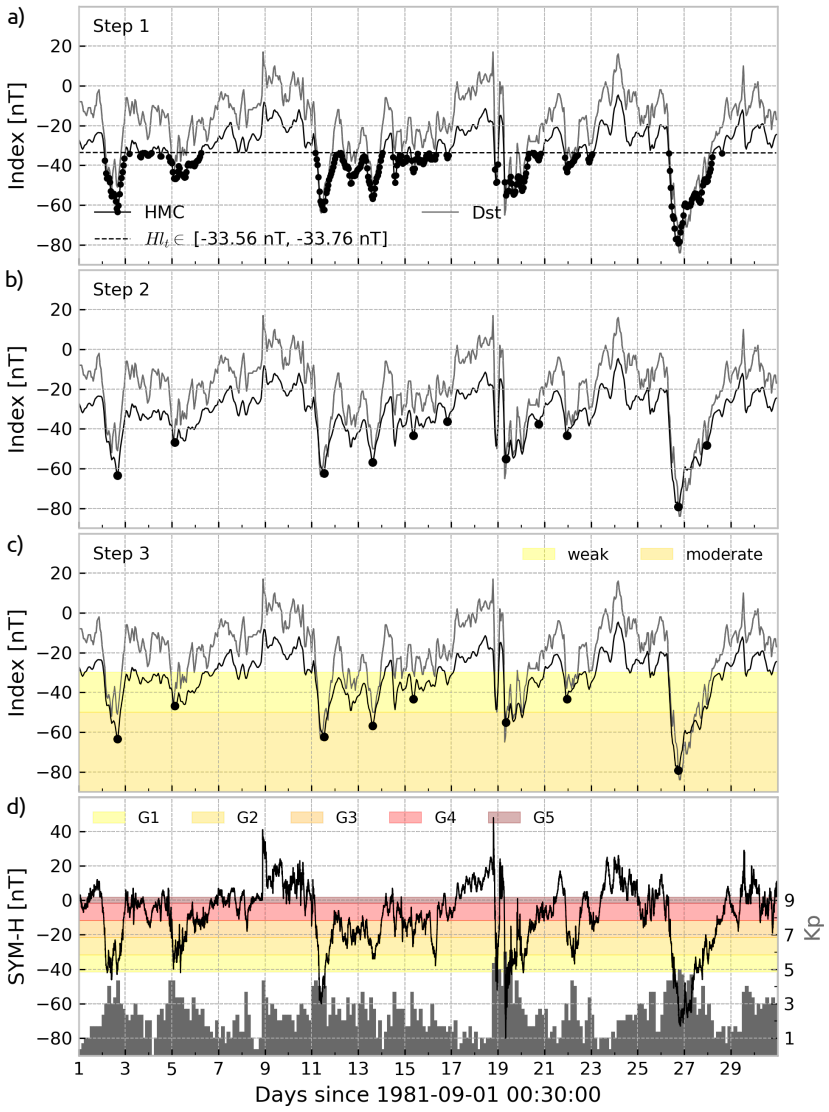

Figure 5.A1: *Event selection.* Black dots mark **a)** hours at which HMC falls below the detection threshold (dashed), which varies by 0.2 nT during this interval; **b)** local HMC minima, separated by at least 27 hours; **c)** final event peaks, at which HMC lies at least 7 nT below the background. **d)** SYM-H (left) and Kp (right) indices. Colored bands in panels c) and d) indicate commonly used Dst- and Kp-based geomagnetic storm intensity scales (see section 5.2.2).

APPENDIX B: PERFORMANCE SCORES

Figure 5.B1 shows the row- and column-normalized confusion matrices (Figure 5.4c), whose entries are used in Table 5.B1 to calculate several performance scores (see also Supporting Information, Figure 5.S6). Note that the terms used here have long-standing equivalents in the field of weather forecasting (e.g., Woodcock, 1976), which have been adopted for space weather research (e.g., Liemohn et al., 2018).

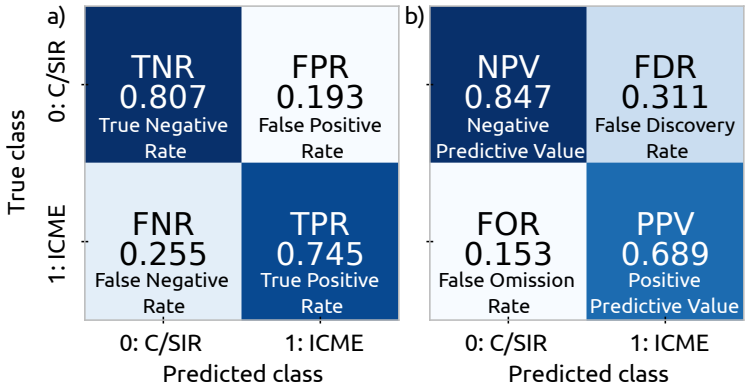

Figure 5.B1: *Normalized confusion matrices.* Confusion matrices of **a)** row-normalized type using N=342 Negatives and P=196 Positives and **b)** column-normalized type using PN=326 Predicted Negatives and PP=212 Predicted Positives.

Table 5.B1: *Performance scores.*Cross-validated performance scores for the validation (columns 4 & 5, top row) and training sets (bottom row). **Abbr.** P: Positives; N: Negatives; POP=N+P: Population; T: True; F: False; R: Rate; $[P/N]PV$: $[P/N]$ Predictive Value; see Figures 5.4c & 5.B1.

Metric	Definition	Range	Scores validation set training set	Std validation set training set
Matthew's Correlation Coefficient	$\text{MCC} = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP) \cdot (TP+FN) \cdot (TN+FP) \cdot (TN+FN)}}$	[-1,1]	0.54 0.57	0.013 0.005
Heidke Skill Score	$\text{HSS} = \frac{2(TP \cdot TN - FN \cdot FP)}{(FN+TN) \cdot (TP+FN) + (TP+FP) \cdot (FP+TN)}$	[-1,1]	0.54 0.57	0.012 0.005
Accuracy	$\text{ACC} = \frac{TP+TN}{POP}$	[0,1]	0.78 0.80	0.005 0.003
Diagnostic odds ratio	$\text{DOR} = \frac{TPR \cdot TNR}{FPR \cdot FNR}$	[0,∞]	13.29 14.57	1.296 0.503
F1 score	$\text{F1} = 2\frac{PPV \cdot TPR}{PPV+TPR}$	[0,1]	0.71 0.73	0.009 0.003
Frequency bias	$\text{FB} = \frac{TP+FP}{TP+FN}$	[0,∞]	1.08 1.09	0.018 0.007
Youden's J statistic (Informedness)	$\text{J} = TPR + TNR - 1$	[-1,1]	0.55 0.58	0.014 0.005
Deltap (Markedness)	$\text{Deltap} = PPV + NPV - 1$	[-1,1]	0.53 0.56	0.011 0.006

SUPPORTING INFORMATION

Text S1

We have executed nine test runs of the binary logistic regression model in order to quantify the sensitivity of the classification result to the parameters of our event selection scheme (Table 5.S1). In each of the runs 1–6, we have set one of the three selection parameters to an extreme value (either much lower or much higher than the default), while keeping everything else unchanged. In run 7, we have "switched off" the solar cycle scaling of $\boldsymbol{Hl}$ (see Equation 5.2) and in run 8, we have trained the classifier on the whole reference set ($d_{tr} = d_{re}$). The following conclusions can be drawn:

1. A formally better classification result compared to that of the default model is reached only by runs 2 and 4, corresponding to decreases of Hs or n, i.e., a more negative Hs and P_n. Both cases result in a larger MCC with comparable or better class probabilities $\tilde{P}_0$, $\tilde{P}_1$. However, this is based on significantly reduced target sets, which include well below 50 % of the reference events. Additionally, the training sets consist of notably larger ICME driver fractions (smaller γ_{tr}), accompanied by a shifted HMC distribution towards larger absolute values. Consequently, the inferred driver class statistics are compromised: At solar minimum years, the fraction of ICME driven weak storms is comparable or even larger than that of C/SIR driven storms. This is not in compliance with what has been found in previous studies (e.g., Richardson and Cane, 2012) and we therefore argue in favor of event selections that are.

2. Since variation of Δt does not change γ_{tr}, its minor effect on the classification result is explained solely by a change of N_{tr} (runs 5, 6). Considering average recovery times (see Figure 5.1b of the main text), Δt=27 hours is reasonable.

3. Without the solar cycle scaling of $\boldsymbol{Hl}$, much less events would be detected during solar minimum years, e.g., only two instead of 36 events in 2009. The scaling allows the detection of $\approx$100 more reference events in the target set, resulting in a larger training set, which is beneficial to the classification.

4. Regardless of the event selection, the first three features (k=1,2,3 in Table 5.1 of the main text) are robustly identified as the most important ones in that order. The only exception occurs in run 8, for which all reference events are used to train the classifier, even those not in the target set (d_{ta}). In that case, the classifier performs clearly worse, which means that the chosen geomagnetic data is less sensitive to the drivers of the events in $d_{re} \setminus d_{tr}$. These geomagnetic disturbances could be more pronounced at high latitudes, which are not covered by our input data set. From that perspective, it makes sense to exclude such events.

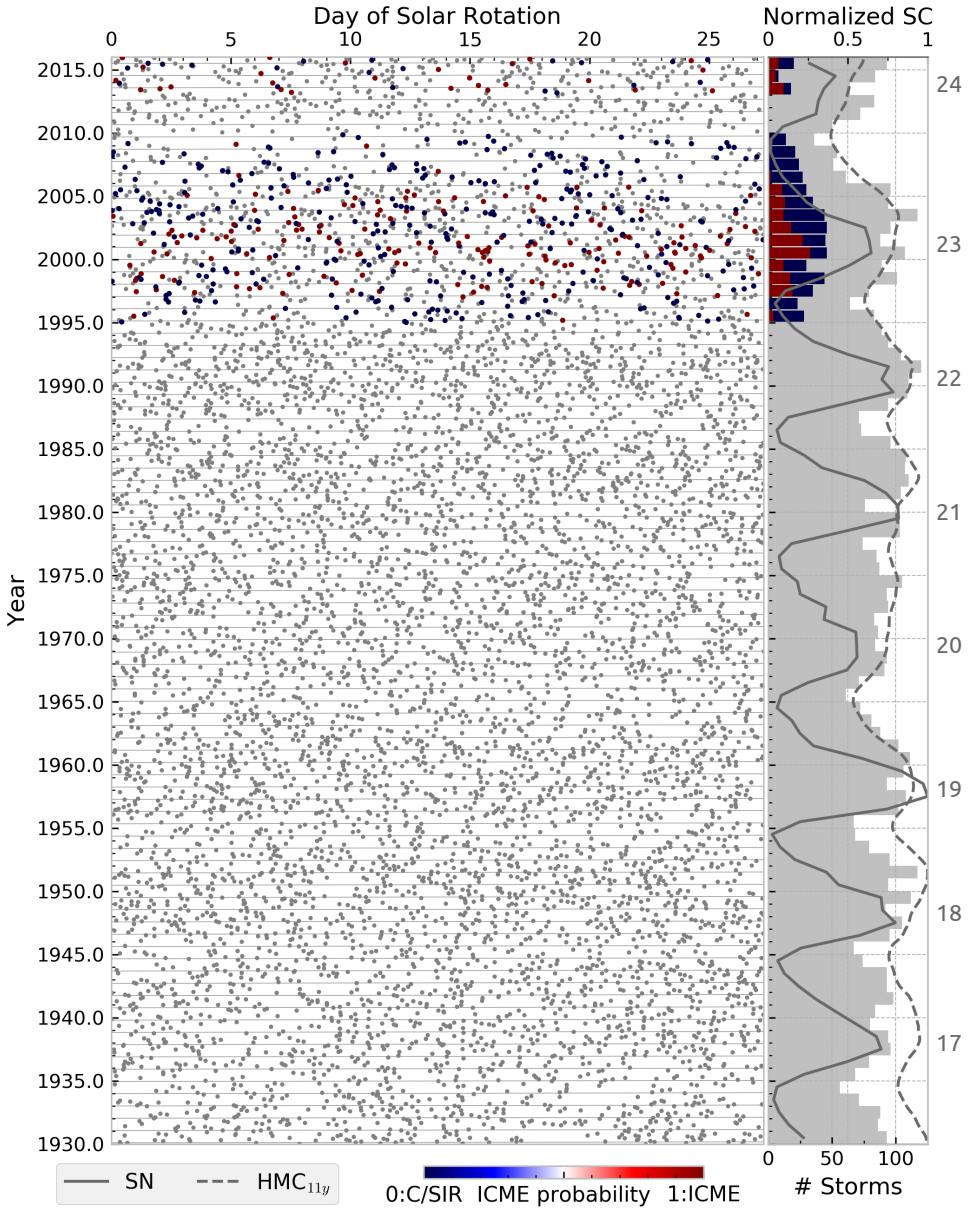

Figure 5.S1: *Classification setup.* Left: Overview of target (t_{ta}, gray) and training events (t_{tr}, colored) in a year vs. day of solar rotation plot. The color bar indicates the ICME driver probability $P(Y_i = 1)$ between zero (100 % C/SIR driver probability) and one. Right: Number of events per year (bars) with normalized yearly sunspot number (SN, solid line) and low-pass filtered HMC (dashed line). Solar cycles (SC) are numbered along the y-axis.

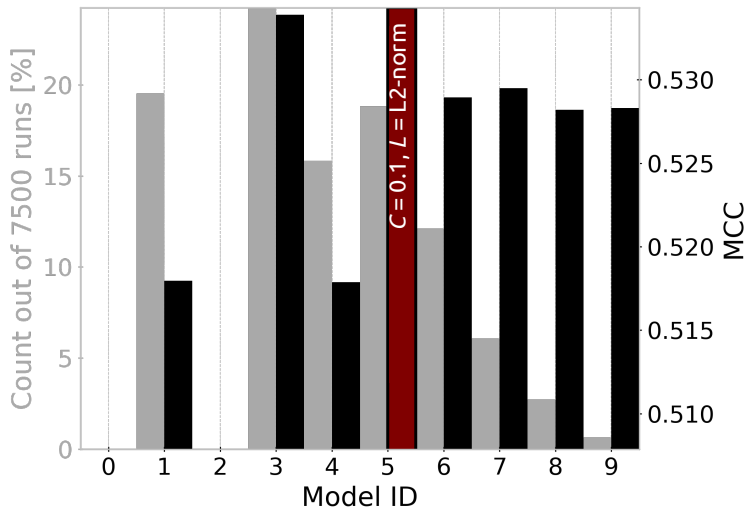

Figure 5.S2: *Model optimization.* Ten combinations x=0,...,9 of parameters C (inverse regularization strength) and L (norm), tested in a nested K-fold cross-validation with 7500 runs per combination. Gray bars show the fraction of total runs for which combination x has the highest Matthew's correlation coefficient (MCC), black bars show MCC of combination x averaged over all runs. With regard to a consistently good performance on all folds, we chose combination 5, i.e., $C = 0.1$ and L=L2-norm (red). For combinations 0 & 2 the averaged MCC falls below the displayed y-range.

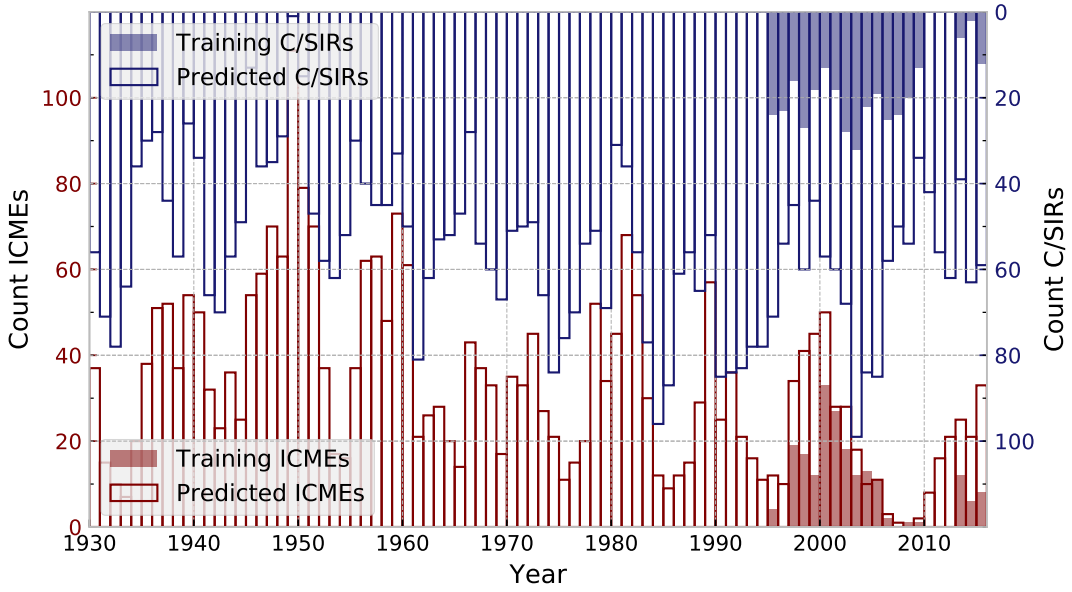

Figure 5.S4: *Yearly histogram of geomagnetic storm drivers.* Yearly numbers of identified ICME (red, left y-axis) and C/SIR (blue, right y-axis) drivers for the target (unfilled) and training events (filled; from Jian et al., (2011, 2006a,b), Turner et al., (2009), Shen et al., (2017)). Since the training events are only a subset of the target events, the total number of events indicated by the two sets differ (see also Figure 5.S1).

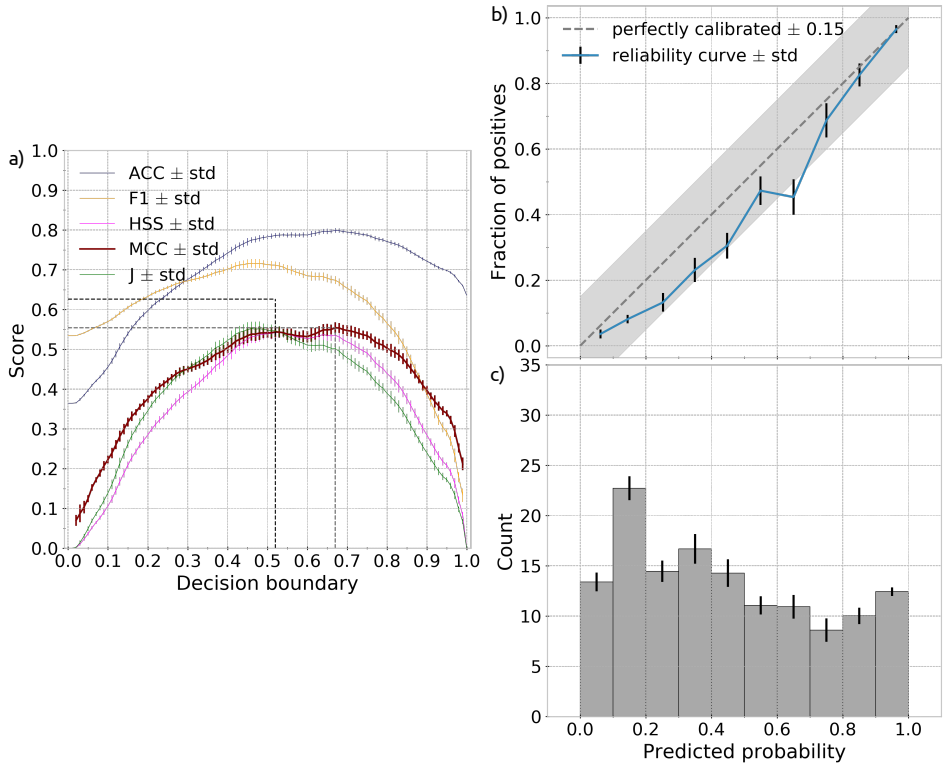

Figure 5.S3: *Decision boundary and reliability curve.* **a)** Five cross-calibrated performance scores (cf. Table 5.B1 in the main text) in dependence of the decision boundary. Dashed lines mark the maxima of MCC (gray) and of the averaged scores (black). Since there is no significant difference between MCC at ≈0.7 and ≈0.5, we set the final decision boundary to the latter. **b)** The reliability curve shows that fraction of events with a certain predicted probability $P(Y_i = 1) = [0, 1]$, that are truly ICME-driven (blue line). A perfectly calibrated, i.e. reliable, classifier would produce the dashed line, shown here with a tolerance of ±15%. **c)** Histogram of predicted probabilities for the positive class (1: ICME-driven events).

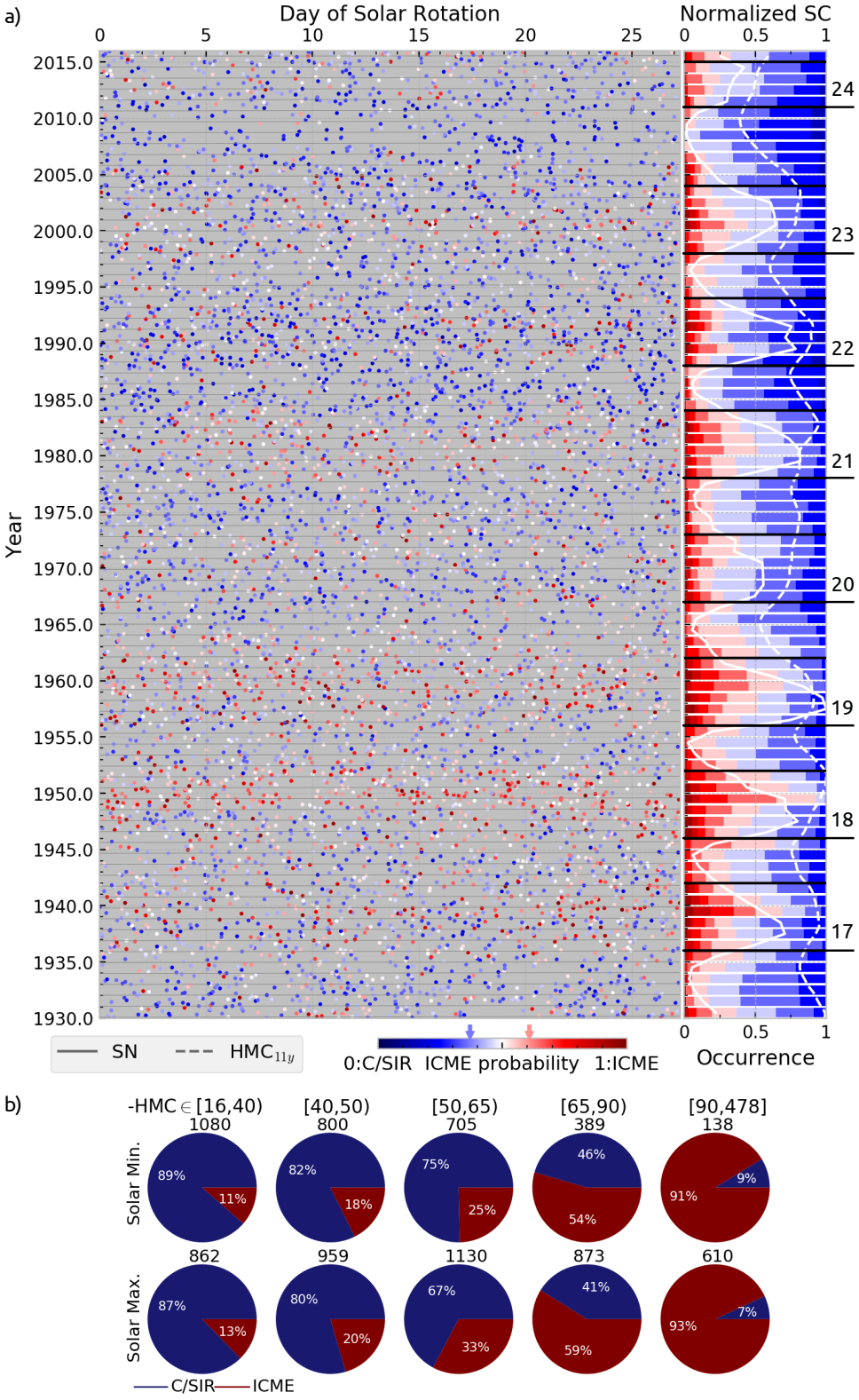

Figure 5.S5: *Classification result without features no. 4 & 5.* Same as Figure 5.5 of the main text, but here the classification algorithm was run without features no. 4 (Solar cycle phase) and no. 5 (Recurrence).

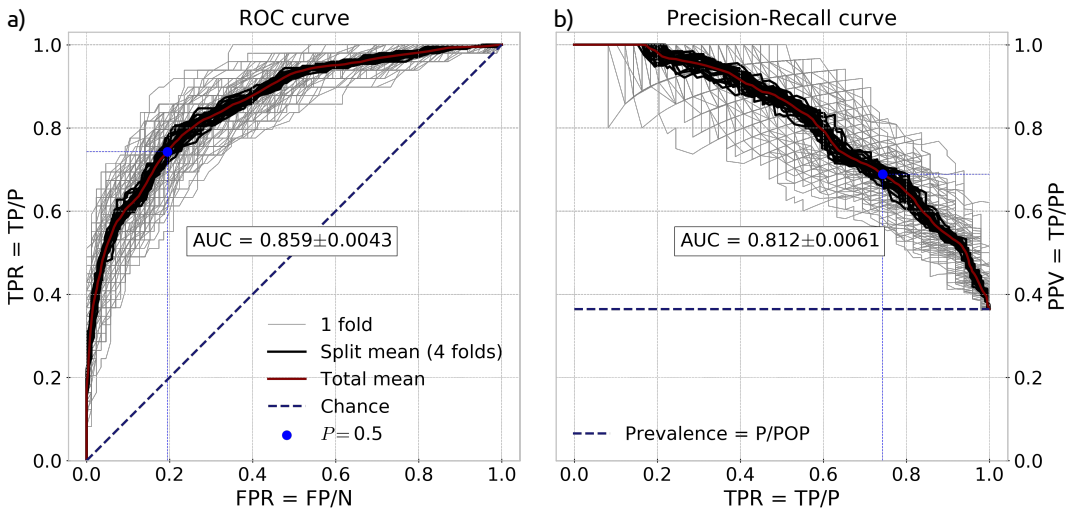

Figure 5.S6: *Diagnostic curves.* **a)** Receiver Operating Characteristic and **b)** Precision-Recall curves from 4-fold cross-validation, repeated 25 times (100 curves each). The curves are constructed by calculating the tuples (TPR-Recall, FPR-Fallout) and (PPV-Precision, TPR-Recall) as a function of the decision boundary, varying between probabilities $P(Y_i = 1) = 0$ and $P(Y_i = 1) = 1$ in steps of 0.01. The boundary $P(Y_i = 1) = 0.5$ is used for the final classifier (blue dots). A perfect classifier gives points in the upper left (a) and right (b) corners, while random guesses are located along the dashed lines. **Abbr.** P: Positives; N: Negatives; POP=N+P: Population; T: True; F: False; R: Rate; $[P/N]PV$: $[P/N]$ Predictive Value; AUC: Area Under Curve estimate; see Figures 5.4c & 5.B1 of the main text.

Table 5.S1: *Model test runs.* Results of model runs with varying setups. Run 0 is the chosen model described in the main text. In each of the runs 1–7, one of the event selection parameters is altered. Run 8 is executed with default settings, but uses all reference events for training. **Abbr.** *Hs*: HMC drop; P_n: n-th HMC percentile; Δt: minimum inter-event time; p_{sc}: solar cycle scaling power; $N_{tr \vee ta}$: number of training or target events; $\gamma_{tr \vee ta}$: class ratios for training or target events; N_f: number of selected features; *MCC*: Matthew's correlation coefficient; $\tilde{P}_0$, $\tilde{P}_1$: median class probabilities (see section 5.4 of the main text).

#	Test	N_{tr}	N_{ta}	γ_{tr}	γ_{ta}	N_f	Feature ranking, k=	*MCC*	$\tilde{P}_0$	$\tilde{P}_1$
0	Default	538	7547	1.74	1.70	8	1,2,3,4,5,6,7,8	0.54	0.74	0.73
1	Hs = 0 nT	561	11863	1.76	2.57	11	1,2,3,4,6,9,7,8, 5,10,11	0.52	0.75	0.68
2	Hs = 18 nT	300	2481	0.99	0.92	11	1,2,3,5,4,7,6,8, 10,11,9	0.55	0.69	0.73
3	P_{50} = -28.4 nT	547	8011	1.79	1.94	7	1,2,3,4,6,5,7	0.53	0.65	0.62
4	P_5 = -64.7 nT	389	3665	1.29	1.04	6	1,2,3,5,4,6	0.59	0.78	0.81
5	Δt = 2 h	602	15217	1.74	1.63	8	1,2,3,9,4,7,6,8	0.52	0.63	0.62
6	Δt = 56 h	467	5277	1.73	1.94	10	1,2,3,4,6,7,5,8, 10,11	0.50	0.64	0.62
7	$p_{sc} = 0$	435	6041	1.30	1.70	10	1,2,3,4,5,6,7,8, 11,9	0.53	0.74	0.71
8	$d_{tr} = d_{re}$	868	7547	1.92	2.03	10	1,9,2,5,3,4,8,7, 11,6	0.46	0.55	0.54

Part III

DISCUSSION

6 CONTEXTUALIZATION OF THE RESULTS

In this chapter, the newly derived indices are placed in the context of the ring current index family. The inferred answers to the research question combine the results of the individual publications under a common scope. Their discussion in light of the current state of knowledge is limited to the two most interesting findings.

6.1 NEW GEOMAGNETIC INDICES AS ANALYSIS BASIS

AMC is derived from annual means of five mid-latitude, northern-hemispheric observatories (including their predecessors) in publication 1 and, as an advancement, HMC is calculated from hourly means of 21 bi-hemispheric observatories in publication 2. Accessing the long-term evolution of inner magnetospheric currents through a Dst-like geomagnetic index is the most basic approach and physically rooted in the DPS relation. However, a global measure is insufficient to capture the currents' dynamics, particularly during geomagnetically disturbed times. Therefore, the MLT dependence of the observatory residuals, which form the basis of HMC (here called $\Delta \boldsymbol{B}_{\mathrm{MLT}}$), is studied in publication 3 (cf. matrix $\boldsymbol{Z}$ in section 5.2.1, Figure 5.1c).

Figure 6.1 places AMC, HMC, and $\Delta \boldsymbol{B}_{\mathrm{MLT}}$ in the context of the ring current index family whose members are designed to quantify the level of mid-latitude geomagnetic activity either in the discipline of geomagnetic field modeling or space weather research (cf. Table 1.1). These fields pose different requirements for the indices. On the one hand, SH-based geomagnetic field modeling requires long-lasting (preferably a full solar cycle), continuous indices that parametrize the temporal variability of the large-scale external geomagnetic field on the time scale of hours. The availability of satellite measurements is the limiting factor for this index class, of which RC can be understood as an end-member. On the other hand, space weather phenomena are location-dependent, and occur on time scales ranging from minutes (e.g., substorms) to days (e.g., geomagnetic storms) so that their event-oriented investigation requires spatially resolved, high cadence indices. The availability of observatory one-minute means is the limiting factor for this second index class, of which the SMR indices (Newell and Gjerloev, 2012) represent end-members.

In this context, AMC and HMC represent results in their own right as they populate the so-far untapped domain of long-term (>100 years), ring current targeting indices. Although not designated a geomagnetic activity index per se, the same is true for $\Delta \boldsymbol{B}_{\mathrm{MLT}}$ (85 years), being the only spatially resolved measure of ring current activity that lasts longer than SMR. Apart from the time series' length, A/HMC combines traits of both ring current index classes to address the research question in light of the known Dst-related problems. These traits arise from the application of the following principles during the indices' derivations:

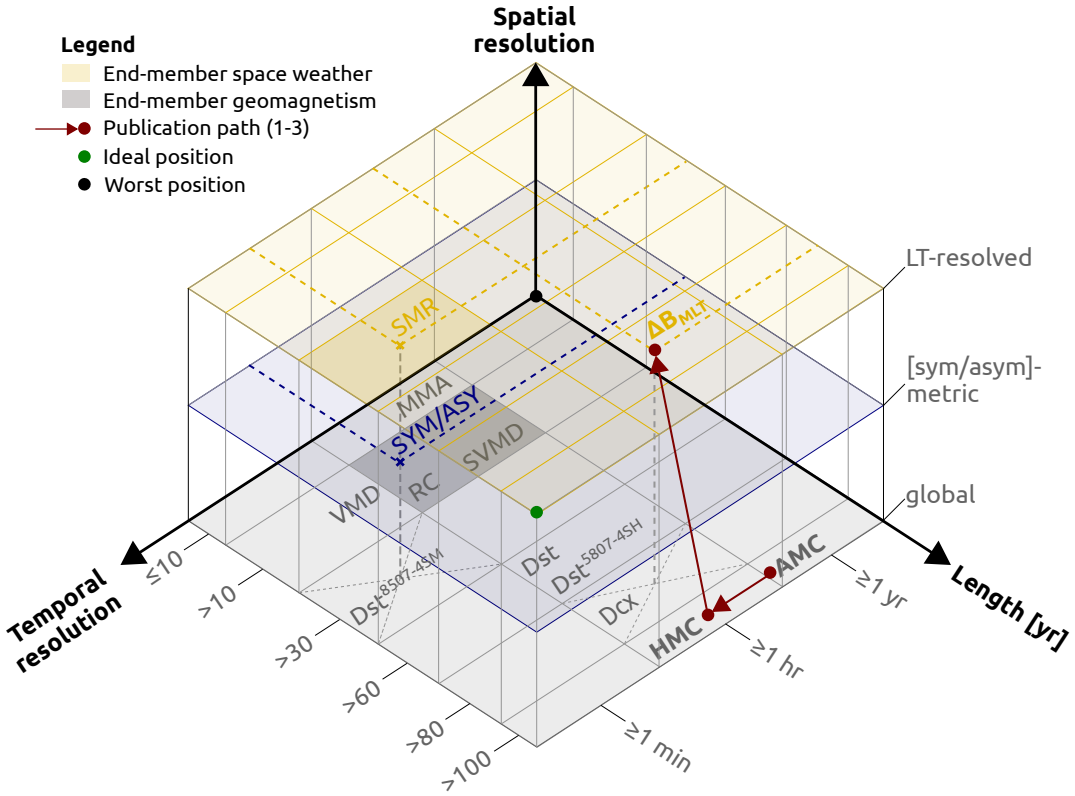

Figure 6.1: The ring current index family (cf. Table 1.1), including the results of publications 1-3 (red dots & arrows), sorted in respect of temporal resolution, index length (horizontal plane), and spatial resolution (vertical axis with discrete levels in gray, blue, and yellow). Each axis points from the worst to the best available state so that the origin represents the worst (black), and the diagonally opposite vertex the ideal (green) overall position. Highlighted squares represent end-member positions in space weather research and geomagnetic field modeling.

1. *Integration of third-party geomagnetic field models*
   The geomagnetic field residuals are formed by subtracting the predictions of third-party geomagnetic field models for the core (i.e., the baseline), the large-scale external (for the crustal biases), and the ionospheric field contributions from the observatory measurements.

2. *Generality*
   No restrictions to geomagnetically quiet and/or dark conditions are imposed on the final index.

3. *Spatial homogeneity*
   The spatial distribution of the contributing observatories is kept relatively stable over time.

4. *Compactness of the representation*
   A transformation into geomagnetic coordinates focuses the relevant signal along just one coordinate axis, the dipole axis (MAG Z_M-component).

While principles 1 and 2 arrange for the quiet-time level and solar-cycle variability to be represented realistically by A/HMC (and $\Delta \boldsymbol{B}_{\mathrm{MLT}}$, see section 1.2.3), a consequent

implementation of principle 3 serves the interpretability of the indices in terms of solar wind forcing (see section 6.3.1) as demanded by the second part of the research question. Principle 4 justifies the designation of A/HMC as conventional indices, i.e., *scalar* time series. The appropriateness and consequences of applying principle 4 are assessed, i.a., in chapter 7.

6.2 ANSWERS TO THE RESEARCH QUESTION

The level of detail at which the centennial evolution of the external magnetic field, ascribed to inner magnetospheric currents, can be characterized naturally depends on the temporal and spatial resolutions of the underlying geomagnetic observatory measurements and the applied analysis methods.

At the lowest level of complexity, AMC describes the annually-averaged magnitude and temporal evolution of the spatially uniform external magnetic field along the dipole axis between 1900 and 2010. This field is shown to be directed southward (negative sign) with a magnitude variation composed of an upward linear trend of 0.13 nT/year (35.44 % absolute decline since 1900) on top of a solar cycle related variability as indicated by comparison to the total yearly sunspot number (Figure 3.6). The latter ranges between amplitudes of 18.14 nT (SC 16) and 44.58 nT (SC 15) while the average absolute magnitude of the field over the whole period amounts to 26.1 nT after subtraction of the drift. The credibility of these findings is rated by AMC's standard deviation, which decreases from 34.01 % of the field's magnitude in 1900 to 3.63 % in 2010.

At the next higher level of complexity, the temporal resolution of the analysis is lowered down to hourly averages while the simplifying assumption of a spatially uniform external field is kept for the calculation of HMC between 1900 and 2015. A careful revision of the underlying hourly mean observatory data is speculated to be the prime cause for the marked reduction of the linear trend to 0.01 nT/year (4.76 % absolute decline between 1900 and 2010) in annual averages of HMC, which is shown to be strongly interval-dependent. The absolute, trend-free average magnitude in the case of HMC amounts to 30.59 nT. Although the OSF is argued to explain $\sim$61 % of HMC's annual variability, the long-term drifts disagree concerning the sign and magnitude (Figure 4.7), particularly within the period 1935–1965. Hence, the long-term change seen in AMC/HMC can not straightforwardly be traced back to that in the OSF.

An analysis of HMC's power spectral density quantitatively confirms an SC related variability that is more pronounced than for rivaling members of the ring current index family (especially Dcx; Figure 4.5). The trend-free amplitudes of the annual SC variation range from 14.9 nT (SC 24) to 31.81 nT (SC 22) with an average value of 26.42 nT (SC 14–23), which agrees well with what is given by AMC (26.38 nT). Based on annual means calculated exclusively from the five quietest days per month (or all hours for which $Kp < 2$) since 1932, a realistic average quiet-time activity level of 17.45 nT (18.66 nT) becomes accessible via HMC.

At the highest level of complexity explored within this thesis, the longitudinal variability of the external disturbance field is taken into account by binning the signals obtained at individual observatories with respect to MLT. The storm-time

disturbance pattern is shown to be statistically sensitive to the kind of solar wind driver (ICMEs vs. CIRs/SIRs). Apart from the size of the HMC-drop at the storm peak, the most discriminative features are derived from the difference of the disturbances in the dawn and dusk sectors within 24 hours after the storm peak (recovery phase), which is speculated to reflect the MLT-dependent effect of the PRC. The predicted annual numbers of ICME- and C/SIR-driven storms show SC-dependent fluctuations while, intriguingly, the number of ICMEs seems to have decreased during 1930–2015.
The entity of the results implies that the mid-latitude geomagnetic activity, ascribed to the ring current system, has slightly decreased from 1900 until 2015 and that the responses of the ring current system to different solar wind drivers are statistically distinguishable at ground level. These two findings are further discussed below.

6.3 THE CURRENT STATE OF KNOWLEDGE

6.3.1 *Long-term trend of geomagnetic activity*

A long-standing, fundamental result in space climate science is the discovery of a notable increase of geomagnetic activity over the last century. The trend was first registered in the aa index (Mayaud, 1972), which is derived homogeneously from just two almost antipodal mid-latitude stations in England (Hartland, HAD) and Australia (Canberra, CNB) and their predecessors since 1868. Being a "range" index (Bartels et al., 1939), aa is defined as the arithmetic mean of the ranges of magnetic field variation detected at the two sites within three-hourly intervals, both of which are manually normalized to that at Niemegk observatory (NGK, Germany) using a quasi-logarithmic scale (giving analog local "*K*" indices). Resulting from a long debate on the trend's reliability, it is now generally accepted (e.g., Lockwood, 2013; Martini et al., 2012a) mainly because the long-term increase could be verified by other global activity indices based on digitized hourly values, e.g., the inter-diurnal and inter-hourly variation indices (IHV, IDV; Svalgaard and Cliver, 2005, 2007; Svalgaard et al., 2004) as well as the Ah and AhK indices (Martini et al., 2011; Mursula and Martini, 2007). This is consistent with the long-term increase registered in the OSF, which is reconstructed based on aa and/or variants of the aforementioned indices (cf. publication 2, section 4.2.2.2; Lockwood et al., 1999).
As can be expected from a complex system, the long-term "trend" in global geomagnetic activity is neither linear nor monotonic (e.g., Love, 2011; Martini et al., 2012a). The latter is demonstrated to apply also for a restriction to low- and mid-latitudes in the form of HMC (publication 2, Table 4.1). When limited to the period 1900–2000, annual averages of HMC display a 9 % absolute linear increase in compliance with the global activity increase as reported above. In fact, most of the HMC-increase occurs until 1941 (SC 17), after which the activity level recovers slowly until compensation is reached during 2008–2009 (SC 23). The net absolute decrease, as reported in the answers to the research question, results from the continuously decreasing activity ever since.
Differences between long-term trends seen in global geomagnetic activity measures and OSF on the one hand, and HMC on the other hand, have various reasons.

On the technical side, differences arise due to the nature of the indices (hourly mean vs. range indices; Lockwood, 2013, chapter 6) and the recording method of the corresponding observatory measurements (analog vs. digital; Martini et al., 2012b). On the physical side, the long-term evolution of local geomagnetic activity depends systematically on geomagnetic latitude. Based on local IHV and Ah indices derived from individual observatory measurements at different MAG latitudes, the centennial increasing trend is shown to be largest at high latitudes, smallest at mid-latitudes, and in between at low latitudes (Martini et al., 2012a; Mursula and Martini, 2006). The underlying mechanisms are addressed by Holappa et al., (2015, 2014), who show that the annually-averaged geomagnetic activity, as measured by local IHV and Ah indices at 26 stations from all latitudes, can be decomposed into a linear combination of two independent components: one that is correlated with the IMF magnitude (B_{sw}) and another correlated with the solar wind speed (v_{sw}). Given that the coupling function $B_{sw}v_{sw}^2$ (see section 1.2.2.2) is dominated by an enhanced B_{sw} for ICMEs and an enhanced v_{sw} for HSSs (or CIRs/SIRs; I. G. Richardson and Cane, 2012; I. G. Richardson et al., 2000), these components can be understood as the ICME- and HSS-related contributions to geomagnetic activity, respectively. The latitude dependence of the local geomagnetic activity is introduced by the varying sensitivities to B_{sw} and v_{sw} in the following sense (see also Finch et al., 2008): They behave oppositely, at which the sensitivity to B_{sw} (v_{sw}) is generally larger (smaller) with a local decrease (increase) at $\sim$50° latitude and a broader minimum (maximum) across the auroral region ($\sim$60–80° latitude). These results are complemented by those of Martini et al., (2015), who demonstrate that the smaller centennial increase of geomagnetic activity at mid-latitudes, as compared to higher latitudes, is caused by a relative overestimation (underestimation) of activity during the first (second) half of the century. Furthermore, the authors show that the differences in geomagnetic activities at mid- and low latitudes are also MLT-dependent (largest at dawn) and maximize in the declining phases of the solar cycles, which is explained by a declining sensitivity to HSSs at mid-latitudes as compared to high latitudes.

The assembled information implies that the geomagnetic activity measured by HMC is more sensitive to B_{sw} (ICMEs) than to v_{sw} (HSSs, C/SIRs). The decrease of mid-latitude geomagnetic activity during 1930–2015 as measured by HMC (publication 2) is therefore consistent with the identified decline of the annual fraction of ICME-driven storms (publication 3).

6.3.2 *Ring current dynamics in response to solar wind drivers*

The conclusion drawn regarding the different responses to ICMEs and C/SIRs measured at the HMC observatory sites should be evaluated in the face of studies that assess the ring current response either via in-situ measurements or global drift physics-based numerical simulations.

In the case of the drift physics models, the issue is commonly tackled by comparing the so-called pressure-corrected Dst (Dst*, see section 7) to simulations of Dst* based on the DPS relation, using the modeled ion fluxes (kinetic energy $\lesssim$500 keV) for several storm events. One example of such a model is the Hot Electron

and Ion Drift Integrator (HEIDI; based on e.g., Fok et al., 1993; Jordanova et al., 1996), which simulates the plasma inflow from the magnetotail, particle acceleration and loss (mainly by charge exchange) along the drift paths, as well as the outflow at the dayside magnetopause. Another one is the Comprehensive Ring Current Model (CRCM; Fok et al., 2001) that uses a self-consistently calculated electric field. A common result of studies based on these models is that Dst^* can be simulated better for ICME-driven storms than for HSS-driven ones of comparable size, i.e., with similar peak Dst^* values (e.g., Cramer et al., 2013; Liemohn et al., 2010), particularly during the recovery phase (Jordanova et al., 2009). The aforementioned authors speculate that, for HSSs, HEIDI either misses plasma injections from substorm-induced electric fields and/or radial diffusion. In relation to the peculiarity associated with the recovery phase, Miyoshi and Kataoka, (2005) evaluate the ion flux from particle measurements of the NOAA (National Oceanic and Atmospheric Administration) satellites and conclude that, during the recovery phase, ring current flux enhancements in the outer ring largely depend on the different solar wind structures, while those in the inner ring can only be caused by strong ICMEs.

The perhaps most important study concerning this thesis (publication3) is the work by Liemohn and Katus, (2012), which builds on the previous study (Liemohn et al., 2010), but uses the pressure-corrected SYM-H index instead of Dst (see table 1.1). The authors conclude that "the inner magnetosphere and near-Earth plasma sheet are responding in fundamentally different ways to the two solar wind driver structures [ICMEs; C/SIRs and HSSs], even for the same level of driving". The reason for this is speculated to be connected to the stronger driving during ICME-storms (see e.g., Borovsky and Denton, 2006), which causes the plasma sheet hot ions to penetrate deeper into the inner magnetosphere where the feedback of the FAC closure is more pronounced.

Although none of the above-mentioned studies deliver concrete evidence in support of the findings of publication 3, two points are consistent. Firstly, the differences in the ring current response to ICMEs and C/SIRs seem to be particularly large during the storm recovery phases. Secondly, these differences are connected to the spatially resolved magnetic disturbance pattern, particularly to the effect of the Region-2 FACs as part of the PRC system (see section 6.2).

7

OUTLOOK

To meet the demands imposed by the integrative research question, two lines of development have been initiated in this study, each of which has proven to be worth pursuing. Basically, the direction of future efforts depends on whether or not the concept of a geomagnetic index (HMC) should be kept (principle 4).

Regardless of the specific road taken, there are some considerations of general applicability when focusing on the long-term magnetic field signal of the ring current. First off, more internationally coordinated efforts toward the complete digitization and publication of the hourly measurements recorded in observatory yearbooks should be initiated. In doing so, the metadata needs to inform clearly on the sampling rate, i.e., whether just spot values or actual hourly means are reported (see Mursula and Martini, 2006), a point not handled consistently in the currently available data catalogs.

A second point of concern is the so-far unaccounted field component that is induced in the electrically conducting mantle and oceans in response to the time-varying ring current (geomagnetic storms). Assuming a 1-D electrical conductivity structure, which depends solely on r, this signal can be eliminated from the observatory residuals (or HMC) with knowledge of the "Q-response" as is shown for Dst and RC (e.g., Olsen et al., 2005). The Q-response is the complex-valued ratio of induced (internal) and inducing (external) Gauss coefficients, which depends on frequency and the spherical harmonic degree (e.g., Kuvshinov, 2008, 2012). However, due to the large influence of the laterally heterogeneous ocean conductivity on the coastal observatory sites ("ocean effect", e.g., Kuvshinov et al., 2002), 3-D effects need to be considered in order to properly assess the induction effect of geomagnetic storms (Olsen and Kuvshinov, 2004; Püthe and Kuvshinov, 2013). However, as geomagnetic storms are relatively rare ($\sim$1 % of all UT hours during 1930–2015 are initially flagged as disturbed in publication 3) and the electrical conductivity structure relatively stable over time, the influence of the induced component on the derived long-term evolution of the magnetospheric magnetic field should be small.

Lastly, an elimination of the fluctuations caused by the solar wind dynamic pressure, transferred to the observatory residuals (and HMC) by the Chapman-Ferraro currents (in analogy to Dst; Siscoe et al., 2005), could be attempted. In the case of an annual time scale, a crude approach is to eliminate all positive values from the whole time series, so that the annual averages of the remaining ones are assured to be *dominated* by the ring current effect (Lockwood, 2013, chapter 2.3). A more sophisticated approach is to model the strength of the pressure contribution to the ground-based magnetic field measurements based on Burton et al., (1975) and then subtract it to get the "pressure-corrected" measurements (see also O'Brien and McPherron, 2000). Of course, the latter is only possible for times when appropriate measurements of solar wind speed and density are available (since $\sim$1960).

Moving on to the different lines of development, it has been shown that HMC holds a unique position within the ring current family (Figure 6.1). As such, it has similarly diverse areas of application as the other indices, e.g.,: selection of geomagnetically quiet times (Kauristie et al., 2017), temporal parametrization of the external magnetic field potential (Finlay et al., 2017), geomagnetic storm detection (particularly across several SCs, publication 3), initialization of drift particle models for the magnetosphere (Fok et al., 2001). Additionally, A/HMC can be used to clean observatory data directly from the magnetospheric magnetic field, which is important, e.g., for the study of geomagnetic jerks (publication 1). To enhance HMC's attractiveness in these areas, particularly for event-based studies, the following steps could be taken:

1. An up-to-date version of HMC for the current year could be delivered using (a) one-hour means of the INTERMAGNET one-minute provisional data (`https://www.intermagnet.org`; 12 HMC-stations available in 2019) (b) a baseline extrapolated from the IGRF and the corresponding, predicted SV (c) crustal biases and the Sq-field estimates as calculated in publications 1 and 2.

2. The Sq-field could be calculated from the data itself, i.e., without the need for additional input parameters (F10.7) to run a third-party model (CM4). A very interesting prospect is the method developed by Martini et al., (2011), who use a Kalman filter technique to obtain the solar daily regular geomagnetic variation from hourly geomagnetic observatory measurements. Applying this method may help eliminate the remainders of the ionospheric field, which are likely contained in the current version of HMC.

3. A homogeneous response of the HMC observatories to the solar wind driving could be insured by selecting only those observatories that show similar sensitivity to the elements of the coupling function (B_{sw}, v_{sw}) on an annual scale.

It has become apparent in publication 3 that information about the kind of the solar wind driving is not retrievable from an index alone, but requires the spatially resolved disturbance field (see also section 6.3.1). A long-term, more detailed description of the magnetospheric magnetic field on the ground would also serve electromagnetic induction studies in need of a realistic "source". In connection with this, the following steps offer themselves:

1. The results of the conducted spherical harmonic analysis (publication 2, section 4.3.2) could be fully exploited by analyzing the external Gauss coefficients perpendicular to the dipole axis (q_1^1, h_1^1). The most appropriate cut-off degree (N) of the spherical harmonic expansion can be checked using the geomagnetic power spectrum (see Maus, 2008).

2. The additional information content of $\Delta \boldsymbol{B}_{\mathrm{MLT}}$ (publication 3) calculated from the MAG X_M- and Y_M-components should be investigated. Given that the different large-scale current systems affect the magnetic field vector components to varying degrees (PRC, Region-2 FACs; see Iyemori, 1990), I expect that their comparison helps to discern the contributions from the individual current

systems (problem 4). To this end, a comparison of the results obtained after the transformation of the observatory data into different coordinate systems, particularly in the SM and GSM, is advisable.

3. The current state of knowledge (section 6.3) indicates that HMC is not as sensitive to HSSs as to ICMEs. Therefore, the results of publication 3 could be meaningfully expanded by incorporating observatory data from high geomagnetic latitudes, i.e. $\gtrsim 50°$, with systematically different sensitivities to the elements of the coupling function.

Part IV

SUMMARY

SUMMARY

This cumulative thesis is concerned with the evolution of geomagnetic activity since the beginning of the 20th century, that is, the time-dependent response of the geomagnetic field to solar forcing. The focus lies on the description of the magnetospheric response field at ground level, which is particularly sensitive to the ring current system, and an interpretation of its variability in terms of the solar wind driving. Thereby, this work contributes to a comprehensive understanding of long-term solar-terrestrial interactions.

The common basis of the presented publications is formed by a reanalysis of vector magnetic field measurements from geomagnetic observatories located at low and middle geomagnetic latitudes. In the first two studies, new ring current targeting geomagnetic activity indices are derived, the Annual and Hourly Magnetospheric Currents indices (A/HMC). Compared to existing indices (e.g., the Dst index), they do not only extend the covered period by at least three solar cycles but also constitute a qualitative improvement concerning the absolute index level and the $\sim$11-year solar cycle variability. The analysis of A/HMC shows that (a) the annual geomagnetic activity experiences an interval-dependent trend with an overall linear decline during 1900–2010 of $\sim$5 % (b) the average trend-free activity level amounts to $\sim$28 nT (c) the solar cycle related variability shows amplitudes of $\sim$15–45 nT (d) the activity level for geomagnetically quiet conditions (Kp$<$2) lies slightly below 20 nT. The plausibility of the last three points is ensured by comparison to independent estimations either based on magnetic field measurements from LEO satellite missions (since the 1990s) or the modeling of geomagnetic activity from solar wind input (since the 1960s). An independent validation of the long-term trend is problematic mainly because the sensitivity of the locally measured geomagnetic activity depends on geomagnetic latitude. Consequently, A/HMC is neither directly comparable to global geomagnetic activity indices (e.g., the aa index) nor to the partly reconstructed open solar magnetic flux, which requires a homogeneous response of the ground-based measurements to the interplanetary magnetic field and the solar wind speed.

The last study combines a consistent, HMC-based identification of geomagnetic storms from 1930–2015 with an analysis of the corresponding spatial (magnetic local time-dependent) disturbance patterns. Amongst others, the disturbances at dawn and dusk, particularly their evolution during the storm recovery phases, are shown to be indicative of the solar wind driving structure (Interplanetary Coronal Mass Ejections vs. Stream or Co-rotating Interaction Regions), which enables a backward-prediction of the storm driver classes. The results indicate that ICME-driven geomagnetic storms have decreased since 1930 which is consistent with the concurrent decrease of HMC.

Out of the collection of compiled follow-up studies the inclusion of measurements from high-latitude geomagnetic observatories into the third study's framework seems most promising at this point.

Part V

BIBLIOGRAPHY

BIBLIOGRAPHY

Alfvén, H. (1977). "Electric currents in cosmic plasmas." In: *Reviews of Geophysics* 15.3, pp. 271–284. DOI: 10.1029/RG015i003p00271 (cit. on p. 13).

Asikainen, T., V. Maliniemi, and K. Mursula (2010). "Modeling the contributions of ring, tail, and magnetopause currents to the corrected Dst index." In: *Journal of Geophysical Research: Space Physics* 115.A12. DOI: 10.1029/2010JA015774 (cit. on p. 18).

Bartels, J., N. H. Heck, and H. F. Johnston (1939). "The three-hour-range index measuring geomagnetic activity." In: *Terrestrial Magnetism and Atmospheric Electricity* 44.4, pp. 411–454. DOI: 10.1029/TE044i004p00411 (cit. on p. 114).

Baumjohann, W. and R. A. Treumann (1997). *Basic Space Plasma Physics*. Imperial College Press, London. ISBN: 1-86094-079-X (cit. on p. 14).

Birn, J., M. Hesse, and K. Schindler (1998). "Formation of thin current sheets in space plasmas." In: *Journal of Geophysical Research: Space Physics* 103.A4, pp. 6843–6852. DOI: 10.1029/97JA03602 (cit. on p. 13).

Borovsky, J. E. and Y. Y. Shprits (2017). "Is the Dst Index Sufficient to Define All Geospace Storms?" In: *Journal of Geophysical Research: Space Physics* 122.11, pp. 11, 543–11, 547. DOI: 10.1002/2017JA024679 (cit. on p. 16).

Borovsky, J. E. and M. H. Denton (2006). "Differences between CME-driven storms and CIR-driven storms." In: *Journal of Geophysical Research: Space Physics* 111.A07S08. DOI: 10.1029/2005JA011447 (cit. on p. 116).

Burton, R. K., R. L. McPherron, and C. T. Russell (1975). "An empirical relationship between interplanetary conditions and Dst." In: *Journal of Geophysical Research* 80.31, pp. 4204–4214. DOI: 10.1029/JA080i031p04204 (cit. on pp. 17, 117).

Campbell, W. H. (1997). *Introduction to Geomagnetic Fields*. Cambride University Press. ISBN: 0521571936 (cit. on p. 5).

Cramer, W. D., N. E. Turner, M.-C. Fok, and N. Y. Buzulukova (2013). "Effects of different geomagnetic storm drivers on the ring current: CRCM results." In: *Journal of Geophysical Research: Space Physics* 118.3, pp. 1062–1073. DOI: 10.1002/jgra.50138 (cit. on p. 116).

Cranmer, S. R., S. E. Gibson, and P. Riley (2017). "Origins of the Ambient Solar Wind: Implications for Space Weather." In: *Space Science Reviews* 212.3, pp. 1345–1384. DOI: 10.1007/s11214-017-0416-y (cit. on p. 12).

Daglis, I. A., R. M. Thorne, W. Baumjohann, and S. Orsini (1999). "The terrestrial ring current: Origin, formation, and decay." In: *Reviews of Geophysics* 37.4, pp. 407–438. DOI: 10.1029/1999RG900009 (cit. on p. 14).

Dessler, A. J. and E. N. Parker (1959). "Hydromagnetic theory of geomagnetic storms." In: *Journal of Geophysical Research* 64.12, pp. 2239–2252. DOI: 10.1029/JZ064i012p02239 (cit. on p. 16).

Dubyagin, S., N. Ganushkina, M. Kubyshkina, and M. Liemohn (2014). "Contribution from different current systems to SYM and ASY midlatitude indices." In:

Journal of Geophysical Research: Space Physics 119.9, pp. 7243–7263. DOI: 10.1002/2014JA020122 (cit. on p. 18).

Dungey, J. W. (1961). "Interplanetary Magnetic Field and the Auroral Zones." In: *Phys. Rev. Lett.* 6 (2), pp. 47–48. DOI: 10.1103/PhysRevLett.6.47 (cit. on p. 13).

– (1965). "The length of the magnetospheric tail." In: *Journal of Geophysical Research* 70.7, pp. 1753–1753. DOI: 10.1029/JZ070i007p01753 (cit. on p. 13).

Egeland, A. and W. J. Burke (2012). "The ring current: a short biography." In: *History of Geo- and Space Sciences* 3.2, pp. 131–142. DOI: 10.5194/hgss-3-131-2012 (cit. on p. 15).

Finch, I. D. and M. Lockwood (2007). "Solar wind-magnetosphere coupling functions on timescales of 1 day to 1 year." In: *Annales Geophysicae* 25.2, pp. 495–506. DOI: 10.5194/angeo-25-495-2007 (cit. on p. 14).

Finch, I. D., M. Lockwood, and A. P. Rouillard (2008). "Effects of solar wind magnetosphere coupling recorded at different geomagnetic latitudes: Separation of directly-driven and storage/release systems." In: *Geophysical Research Letters* 35.21. DOI: 10.1029/2008GL035399 (cit. on pp. 16, 115).

Finlay, C. C. et al. (2017). "Challenges Handling Magnetospheric and Ionospheric Signals in Internal Geomagnetic Field Modelling." In: *Space Science Reviews* 206.1, pp. 157–189. DOI: 10.1007/s11214-016-0285-9 (cit. on p. 118).

Finlay, C. C., N. Olsen, S. Kotsiaros, N. Gillet, and L. Tøffner-Clausen (2016). "Recent geomagnetic secular variation from Swarm and ground observatories as estimated in the CHAOS-6 geomagnetic field model." In: *Earth, Planets and Space* 68.112. DOI: 10.1186/s40623-016-0486-1 (cit. on p. 11).

Fok, M.-C., J. U. Kozyra, A. F. Nagy, C. E. Rasmussen, and G. V. Khazanov (1993). "Decay of equatorial ring current ions and associated aeronomical consequences." In: *Journal of Geophysical Research: Space Physics* 98.A11, pp. 19381–19393. DOI: 10.1029/93JA01848 (cit. on p. 116).

Fok, M.-C., R. A. Wolf, R. W. Spiro, and T. E. Moore (2001). "Comprehensive computational model of Earth's ring current." In: *Journal of Geophysical Research: Space Physics* 106.A5, pp. 8417–8424. DOI: 10.1029/2000JA000235 (cit. on pp. 116, 118).

Gannon, J. L. and J. J. Love (2011). "USGS 1-min Dst index." In: *Journal of Atmospheric and Solar-Terrestrial Physics* 73.2, pp. 323–334. DOI: 10.1016/j.jastp.2010.02.013 (cit. on pp. 17, 19).

Ganushkina, N. Y., M. W. Liemohn, and S. Dubyagin (2018). "Current Systems in the Earth's Magnetosphere." In: *Reviews of Geophysics* 56.2, pp. 309–332. DOI: 10.1002/2017RG000590 (cit. on pp. 15, 18).

Gosling, J. T. and V. J. Pizzo (1999). "Formation and Evolution of Corotating Interaction Regions and their Three Dimensional Structure." In: *Space Science Reviews* 89.1, pp. 21–52. DOI: 10.1023/A:1005291711900 (cit. on p. 12).

Hamilton, B. (2013). "Rapid modelling of the large-scale magnetospheric field from Swarm satellite data." In: *Earth, Planets and Space* 65.10, pp. 1295–1308. DOI: 10.5047/eps.2013.09.003 (cit. on pp. 17, 19).

Holappa, L., K. Mursula, and T. Asikainen (2015). "A new method to estimate annual solar wind parameters and contributions of different solar wind structures

to geomagnetic activity." In: *Journal of Geophysical Research: Space Physics* 119.12, pp. 9407–9418. DOI: 10.1002/2014JA020599 (cit. on p. 115).

Holappa, L., K. Mursula, T. Asikainen, and I. G. Richardson (2014). "Annual fractions of high-speed streams from principal component analysis of local geomagnetic activity." In: *Journal of Geophysical Research: Space Physics* 119.6, pp. 4544–4555. DOI: 10.1002/2014JA019958 (cit. on p. 115).

Iyemori, T. (1990). "Storm-Time Magnetospheric Currents Inferred from Mid-Latitude Geomagnetic Field Variations." In: *Journal of geomagnetism and geoelectricity* 42.11, pp. 1249–1265. DOI: 10.5636/jgg.42.1249 (cit. on pp. 18, 19, 118).

Iyemori, T., M. Takeda, M. Nosé, Y. Odagi, and H. Toh (2010). "Mid-latitude Geomagnetic Indices ASY and SYM for 2009 (Provisional)." In: *Internal Report of Data Analysis Center for Geomagnetism and Space Magnetism, Kyoto University, Japan* (cit. on p. 18).

Johnston, H. F. (1943). "Mean K-indices from twenty-one magnetic observatories and five quiet and five disturbed days for 1942." In: *Terrestrial Magnetism and Atmospheric Electricity* 48.4, pp. 219–227. DOI: 10.1029/TE048i004p00219 (cit. on p. 16).

Jordanova, V. K., L. M. Kistler, J. U. Kozyra, G. V. Khazanov, and A. F. Nagy (1996). "Collisional losses of ring current ions." In: *Journal of Geophysical Research: Space Physics* 101.A1, pp. 111–126. DOI: 10.1029/95JA02000 (cit. on p. 116).

Jordanova, V. K. et al. (2009). "Ring current development during high speed streams." In: *Journal of Atmospheric and Solar-Terrestrial Physics* 71.10, pp. 1093–1102. DOI: 10.1016/j.jastp.2008.09.043 (cit. on p. 116).

Karinen, A. and K. Mursula (2005). "A new reconstruction of the Dst index for 1932-2002." In: *Annales Geophysicae* 23.2, pp. 475–485. DOI: 10.5194/angeo-23-475-2005 (cit. on pp. 16, 19).

Kauristie, K. et al. (2017). "On the Usage of Geomagnetic Indices for Data Selection in Internal Field Modelling." In: *Space Science Reviews* 206.1, pp. 61–90. DOI: 10.1007/s11214-016-0301-0 (cit. on p. 118).

Keika, K. et al. (2018). "Ion Energies Dominating Energy Density in the Inner Magnetosphere: Spatial Distributions and Composition, Observed by Arase/MEP-i." In: *Geophysical Research Letters* 45.22, pp. 12, 153–12, 162. DOI: 10.1029/2018GL080047 (cit. on p. 15).

Kilpua, E. K. J., A. Balogh, R. von Steiger, and Y. D. Liu (2017). "Geoeffective Properties of Solar Transients and Stream Interaction Regions." In: *Space Science Reviews* 212.3, pp. 1271–1314. DOI: 10.1007/s11214-017-0411-3 (cit. on p. 12).

Kivelson, M. G. (1995). "Introduction to Space Physics." In: ed. by M. G. Kivelson and C. T. Russell. Cambridge University Press. Chap. Physics of Space Plasmas, pp. 27–57. ISBN: 9780521457149 (cit. on p. 12).

Kunagu, P., G. Balasis, V. Lesur, E. Chandrasekhar, and C. Papadimitriou (2013). "Wavelet characterization of external magnetic sources as observed by CHAMP satellite: evidence for unmodelled signals in geomagnetic field models." In: *Geophysical Journal International* 192.3, pp. 946–950. DOI: 10.1093/gji/ggs093 (cit. on pp. 17, 19).

Kuvshinov, A. V. (2008). "3-D Global Induction in the Oceans and Solid Earth: Recent Progress in Modeling Magnetic and Electric Fields from Sources of Mag-

netospheric, Ionospheric and Oceanic Origin." In: *Surveys in Geophysics* 29.2, pp. 139–186. DOI: 10.1007/s10712-008-9045-z (cit. on p. 117).

Kuvshinov, A. V. (2012). "Deep Electromagnetic Studies from Land, Sea, and Space: Progress Status in the Past 10 Years." In: *Surveys in Geophysics* 33.1, pp. 169–209. DOI: 10.1007/s10712-011-9118-2 (cit. on p. 117).

Kuvshinov, A. V., N. Olsen, D. B. Avdeev, and O. V. Pankratov (2002). "Electromagnetic induction in the oceans and the anomalous behaviour of coastal C-responses for periods up to 20 days." In: *Geophysical Research Letters* 29.12. DOI: 10.1029/2001GL014409 (cit. on p. 117).

Langel, R. A., R. H. Estes, G. D. Mead, E. B. Fabiano, and E. R. Lancaster (1980). "Initial geomagnetic field model from Magsat vector data." In: *Geophysical Research Letters* 7.10, pp. 793–796. DOI: 10.1029/GL007i010p00793 (cit. on p. 17).

Laundal, K. M. and A. D. Richmond (2017). "Magnetic Coordinate Systems." In: *Space Science Reviews* 206.1, pp. 27–59. DOI: 10.1007/s11214-016-0275-y (cit. on pp. 7, 8, 10, 12, 18).

Le, G., C. T. Russell, and K. Takahashi (2004). "Morphology of the ring current derived from magnetic field observations." In: *Annales Geophysicae* 22.4, pp. 1267–1295. DOI: 10.5194/angeo-22-1267-2004 (cit. on p. 15).

Lesur, V., M. Hamoudi, Y. Choi, J. Dyment, and E. Thébault (2016). "Building the second version of the World Digital Magnetic Anomaly Map (WDMAM)." In: *Earth, Planets and Space* 68.27. DOI: 10.1186/s40623-016-0404-6 (cit. on p. 8).

Liemohn, M. W., R. M. Katus, and R. Ilie (2015). "Statistical analysis of storm-time near-Earth current systems." In: *Annales Geophysicae* 33.8, pp. 965–982. DOI: 10.5194/angeo-33-965-2015 (cit. on pp. 15, 18).

Liemohn, M. W. and R. Katus (2012). "Is the storm time response of the inner magnetospheric hot ions universally similar or driver dependent?" In: *Journal of Geophysical Research: Space Physics* 117.A00L03. DOI: 10.1029/2011JA017389 (cit. on p. 116).

Liemohn, M. W. et al. (2010). "CIR versus CME drivers of the ring current during intense magnetic storms." In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 466.2123, pp. 3305–3328. DOI: 10.1098/rspa.2010.0075 (cit. on p. 116).

Lockwood, M. (2013). "Reconstruction and Prediction of Variations in the Open Solar Magnetic Flux and Interplanetary Conditions." In: *Living Reviews in Solar Physics* 10.4. DOI: 10.12942/lrsp-2013-4 (cit. on pp. 13, 114, 115, 117).

Lockwood, M., M. Owens, and A. P. Rouillard (2009a). "Excess open solar magnetic flux from satellite data: 1. Analysis of the third perihelion Ulysses pass." In: *Journal of Geophysical Research: Space Physics* 114.A11. DOI: 10.1029/2009JA014449 (cit. on p. 13).

– (2009b). "Excess open solar magnetic flux from satellite data: 2. A survey of kinematic effects." In: *Journal of Geophysical Research: Space Physics* 114.A11. DOI: 10.1029/2009JA014450 (cit. on p. 13).

Lockwood, M., R. Stamper, and M. N. Wild (1999). "A doubling of the Sun's coronal magnetic field during the past 100 years." In: *Nature* 399, pp. 437–439 (cit. on p. 114).

Love, J. J. (2011). "Secular trends in storm-level geomagnetic activity." In: *Annales Geophysicae* 29.2, pp. 251–262. DOI: 10.5194/angeo-29-251-2011 (cit. on p. 114).

Love, J. J. and J. L. Gannon (2009). "Revised D_{st} and the epicycles of magnetic disturbance: 1958–2007." In: *Annales Geophysicae* 27.8, pp. 3101–3131. DOI: 10.5194/angeo-27-3101-2009 (cit. on pp. 17, 19).

Lühr, H. and S. Maus (2010). "Solar cycle dependence of quiet-time magnetospheric currentsand a model of their near-Earth magnetic fields." In: *Earth, Planets and Space* 62, pp. 843–848. DOI: 10.5047/eps.2010.07.012 (cit. on p. 17).

Lühr, H., C. Xiong, N. Olsen, and G. Le (2017). "Near-Earth Magnetic Field Effects of Large-Scale Magnetospheric Currents." In: *Space Science Reviews* 206.1, pp. 521–545. DOI: 10.1007/s11214-016-0267-y (cit. on pp. 9, 17).

Lui, A. T. Y. (2003). "Inner magnetospheric plasma pressure distribution and its local time asymmetry." In: *Geophysical Research Letters* 30.16. DOI: 10.1029/2003GL017596 (cit. on p. 15).

Lui, A. T. Y., R. W. McEntire, and S. M. Krimigis (1987). "Evolution of the ring current during two geomagnetic storms." In: *Journal of Geophysical Research: Space Physics* 92.A7, pp. 7459–7470. DOI: 10.1029/JA092iA07p07459 (cit. on p. 15).

Martini, D., H.-J. Linthe, V. S. Pandey, and D.-H. Lee (2012a). "On the centennial trend estimates of geomagnetic activity indices." In: *Journal of Geophysical Research: Space Physics* 117.A6. DOI: 10.1029/2012JA017564 (cit. on pp. 114, 115).

Martini, D., K. Mursula, M. Orispää, and H.-J. Linthe (2015). "Long-term decrease in the response of midlatitude stations to high-speed solar wind streams in 1914–2000." In: *Journal of Geophysical Research: Space Physics* 120.4, pp. 2662–2674. DOI: 10.1002/2014JA020813 (cit. on p. 115).

Martini, D. et al. (2011). "Kalman filter technique for defining solar regular geomagnetic variations: Comparison of analog and digital methods at Sodankylä Observatory." In: *Journal of Geophysical Research: Space Physics* 116.A6. DOI: 10.1029/2010JA016343 (cit. on pp. 114, 118).

Martini, D. et al. (2012b). "Long-term changes in indices of geomagnetic activity at the auroral station Sodankylä." In: *Advances in Space Research* 50.6, pp. 690–699. DOI: 10.1016/j.asr.2012.01.013 (cit. on p. 115).

Matzka, J., A. Chulliat, M. Mandea, C. C. Finlay, and E. Qamili (2010). "Geomagnetic Observations for Main Field Studies: From Ground to Space." In: *Space Science Reviews* 155.1, pp. 29–64. ISSN: 1572-9672. DOI: 10.1007/s11214-010-9693-4 (cit. on pp. 4, 6).

Maus, S. (2008). "The geomagnetic power spectrum." In: *Geophysical Journal International* 174.1, pp. 135–142. DOI: 10.1111/j.1365-246X.2008.03820.x (cit. on p. 118).

Mayaud, P.-N. (1972). "The aa indices: A 100-year series characterizing the magnetic activity." In: *Journal of Geophysical Research* 77.34, pp. 6870–6874. DOI: 10.1029/JA077i034p06870 (cit. on p. 114).

Merrill, R. T., M. W. McElhinny, and P. L. McFadden (1998). *The Magnetic Field of the Earth - Paleomagnetism, the Core, and the Deep Mantle*. Ed. by R. Dmowska and J. R. Holton. Vol. 63. International Geophysics Series. Academic Press. ISBN: 0-12-491246-X (cit. on p. 5).

Milan, S. E. et al. (2017). "Overview of Solar Wind-Magnetosphere-Ionosphere-Atmosphere Coupling and the Generation of Magnetospheric Currents." In: *Space Science Reviews* 206, pp. 547–573. DOI: 10.1007/s11214-017-0333-0 (cit. on pp. 14, 15).

Miyoshi, Y. and R. Kataoka (2005). "Ring current ions and radiation belt electrons during geomagnetic storms driven by coronal mass ejections and corotating interaction regions." In: *Geophysical Research Letters* 32.L21105. DOI: 10.1029/2005GL024590 (cit. on p. 116).

Mursula, K. and A. Karinen (2005). "Explaining and correcting the excessive semi-annual variation in the Dst index." In: *Geophysical Research Letters* 32.14. DOI: 10.1029/2005GL023132 (cit. on pp. 17, 19).

Mursula, K. and D. Martini (2006). "Centennial increase in geomagnetic activity: Latitudinal differences and global estimates." In: *Journal of Geophysical Research: Space Physics* 111.A8. DOI: 10.1029/2005JA011549 (cit. on pp. 115, 117).

– (2007). "A new verifiable measure of centennial geomagnetic activity: Modifying the K index method for hourly data." In: *Geophysical Research Letters* 34.22. DOI: 10.1029/2007GL031123 (cit. on p. 114).

Newell, P. T. and J. W. Gjerloev (2012). "SuperMAG-based partial ring current indices." In: *Journal of Geophysical Research: Space Physics* 117.A5. DOI: 10.1029/2012JA017586 (cit. on pp. 18, 19, 111).

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. DOI: 10.1029/2006JA012015 (cit. on pp. 14, 17).

O'Brien, T. P. and R. L. McPherron (2000). "An empirical phase space analysis of ring current dynamics: Solar wind control of injection and decay." In: *Journal of Geophysical Research: Space Physics* 105.A4, pp. 7707–7719. DOI: 10.1029/1998JA000437 (cit. on p. 117).

Olsen, N., K.-H. Glassmeier, and X. Jia (2010). "Separation of the Magnetic Field into External and Internal Parts." In: *Space Science Reviews* 152.1, pp. 135–157. DOI: 10.1007/s11214-009-9563-0 (cit. on p. 11).

Olsen, N. and A. V. Kuvshinov (2004). "Modeling the ocean effect of geomagnetic storms." In: *Earth, Planets and Space* 56, pp. 525–530. DOI: 10.1186/BF03352512 (cit. on p. 117).

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. DOI: 10.1186/BF03351897 (cit. on p. 117).

Olsen, N. and C. Stolle (2012). "Satellite geomagnetism." In: *Annual Review of Earth and Planetary Sciences* 40, pp. 441–465. DOI: 10.1146/annurev-earth-042711-105540 (cit. on p. 8).

Olsen, N. et al. (2014). "The CHAOS-4 geomagnetic field model." In: *Geophysical Journal International* 197.2, pp. 815–827. DOI: 10.1093/gji/ggu033 (cit. on pp. 17, 19).

Parker, E. N. (1957). "Newtonian Development of the Dynamical Properties of Ionized Gases of Low Density." In: *Phys. Rev.* 107 (4), pp. 924–933. DOI: 10.1103/PhysRev.107.924 (cit. on p. 15).

Perez, J. D. et al. (2015). "TWINS stereoscopic imaging of multiple peaks in the ring current." In: *Journal of Geophysical Research: Space Physics* 120.1, pp. 368–383. DOI: 10.1002/2014JA020662 (cit. on p. 15).

Pick, L. (2019). "ClassifyStorms - an automatic classifier for geomagnetic storm drivers based on machine learning techniques. V. 1.0.1." In: *GFZ Data Services*. DOI: 10.5880/GFZ.2.3.2019.003 (cit. on p. 77).

Pick, L. and M. Korte (2017). "An annual proxy for the geomagnetic signal of magnetospheric currents on Earth based on observatory data from 1900–2010." In: *Geophysical Journal International* 211.2, pp. 1223–1236. DOI: 10.1093/gji/ggx367 (cit. on p. 51).

Prölss, G. W. (2004). *Physics of the Earth's Space Environment - An Introduction*. Springer. ISBN: 3-540-21426-7 (cit. on p. 12).

Püthe, C. and A. V. Kuvshinov (2013). "Towards quantitative assessment of the hazard from space weather. Global 3-D modellings of the electric field induced by a realistic geomagnetic storm." In: *Earth, Planets and Space* 65.9, pp. 1017–1025. DOI: 10.5047/eps.2013.03.003 (cit. on p. 117).

Rasson, J. L., H. Toh, and D. Yang (2011). "Geomgnetic Observations and Models." In: ed. by M. Mandea and M. Korte. Vol. 5. IAGA Special Sopron Book Series. Springer. Chap. The Global Geomagnetic Observatory Network, pp. 1–25. DOI: 10.1007/978-90-481-9858-0 (cit. on p. 4).

Rastätter, L. et al. (2013). "Geospace environment modeling 2008–2009 challenge: Dst index." In: *Space Weather* 11.4, pp. 187–205. DOI: 10.1002/swe.20036 (cit. on p. 17).

Richardson, I. G. and H. V. Cane (2012). "Near-earth solar wind flows and related geomagnetic activity during more than four solar cycles (1963-2011)." In: *Journal of Space Weather and Space Climate* 2, A02. DOI: 10.1051/swsc/2012003 (cit. on p. 115).

Richardson, I. G., E. W. Cliver, and H. V. Cane (2000). "Sources of geomagnetic activity over the solar cycle: Relative importance of coronal mass ejections, high-speed streams, and slow solar wind." In: *Journal of Geophysical Research: Space Physics* 105.A8, pp. 18203–18213. DOI: 10.1029/1999JA000400 (cit. on p. 115).

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* 70.2, pp. 219–225. DOI: 10.1016/j.jastp.2007.08.039 (cit. on p. 13).

Roederer, J. G. and H. Zhang (2014). *Dynamics of Magnetically Trapped Particles*. Second Edition. Vol. 403. Astrophysics and Space Science Library. Springer. DOI: 10.1007/978-3-642-41530-2 (cit. on pp. 7, 15).

Sabaka, T. J., N. Olsen, R. H. Tyler, and A. V. Kuvshinov (2015). "CM5, a pre-Swarm comprehensive geomagnetic field model derived from over 12 yr of CHAMP, Ørsted, SAC-C and observatory data." In: *Geophysical Journal International* 200.3, pp. 1596–1626. DOI: 10.1093/gji/ggu493 (cit. on p. 11).

Schatten, K. H. (2013). "Models for Coronal and Interplanetary Magnetic Fields: A Critical Commentary." In: *Sun-Earth Plasma Connections*. American Geophysical Union, pp. 129–142. ISBN: 9780875900926 (cit. on p. 12).

Sckopke, N. (1966). "A general relation between the energy of trapped particles and the disturbance field near the Earth." In: *Journal of Geophysical Research* 71.13, pp. 3125–3130. DOI: 10.1029/JZ071i013p03125 (cit. on p. 16).

Siscoe, G. L., R. L. McPherron, and V. K. Jordanova (2005). "Diminished contribution of ram pressure to Dst during magnetic storms." In: *Journal of Geophysical Research: Space Physics* 110.A12. DOI: 10.1029/2005JA011120 (cit. on p. 117).

Smith, E. J. (2001). "The heliospheric current sheet." In: *Journal of Geophysical Research: Space Physics* 106.A8, pp. 15819–15831. DOI: 10.1029/2000JA000120 (cit. on p. 13).

Speiser, T. W. (1973). "Magnetospheric current sheets." In: *Radio Science* 8.11, pp. 973–977. DOI: 10.1029/RS008i011p00973 (cit. on p. 13).

Sugiura, M. (1964). "Hourly values of equatorial Dst for the IGY." In: *Annales of the International Geophysical Year* 35, pp. 4–45 (cit. on pp. 16, 19).

Sugiura, M. and T. Kamei (1991). "Equatorial Dst index 1957–1986." In: *IAGA Bulletin* 40 (cit. on pp. 16, 19).

Svalgaard, L. and E. W. Cliver (2005). "The IDV index: Its derivation and use in inferring long-term variations of the interplanetary magnetic field strength." In: *Journal of Geophysical Research: Space Physics* 110.A12. DOI: 10.1029/2005JA011203 (cit. on p. 114).

– (2007). "Interhourly variability index of geomagnetic activity and its use in deriving the long-term variation of solar wind speed." In: *Journal of Geophysical Research: Space Physics* 112.A10. DOI: 10.1029/2007JA012437 (cit. on p. 114).

Svalgaard, L., E. W. Cliver, and P. Le Sager (2004). "IHV: a new long-term geomagnetic index." In: *Advances in Space Research* 34.2, pp. 436–439. DOI: 10.1016/j.asr.2003.01.029 (cit. on p. 114).

Temerin, M. and X. Li (2006). "Dst model for 1995–2002." In: *Journal of Geophysical Research: Space Physics* 111.A4. DOI: 10.1029/2005JA011257 (cit. on p. 17).

– (2015). "The Dst index underestimates the solar cycle variation of geomagnetic activity." In: *Journal of Geophysical Research: Space Physics* 120.7, pp. 5603–5607. DOI: 10.1002/2015JA021467 (cit. on p. 17).

Thébault, E. et al. (2015). "International Geomagnetic Reference Field: the 12th generation." In: *Earth, Planets and Space* 67.79. DOI: 10.1186/s40623-015-0228-9 (cit. on p. 8).

Thomson, A. W. P. and V. Lesur (2007). "An improved geomagnetic data selection algorithm for global geomagnetic field modelling." In: *Geophysical Journal International* 169.3, pp. 951–963. DOI: 10.1111/j.1365-246X.2007.03354.x (cit. on pp. 17, 19).

Treumann, R. A. and W. Baumjohann (2013). "Collisionless magnetic reconnection in space plasmas." In: *Frontiers in Physics* 1, p. 31. DOI: 10.3389/fphy.2013.00031 (cit. on p. 13).

Vasyliunas, V. M. (1970). "Mathematical Models of Magnetospheric Convection and its Coupling to the Ionosphere." In: *Particles and Fields in the Magnetosphere*. Ed. by B. M. McCormac. Dordrecht: Springer Netherlands, pp. 60–71. ISBN: 978-94-010-3284-1 (cit. on p. 15).

Vennerstroem, S. (2000). "Long-term rise in geomagnetic activity – A close connection between quiet days and storms." In: *Geophysical Research Letters* 27.1, pp. 69–72. DOI: 10.1029/1999GL010727 (cit. on p. 17).

Welling, D. T. and A. J. Ridley (2010). "Exploring sources of magnetospheric plasma using multispecies MHD." In: *Journal of Geophysical Research: Space Physics* 115.A4. DOI: 10.1029/2009JA014596 (cit. on p. 14).

Wicht, J. and S. Sanchez (2019). "Advances in geodynamo modelling." In: *Geophysical & Astrophysical Fluid Dynamics* 113.1–2, pp. 2–50. DOI: 10.1080/03091929.2019.1597074 (cit. on p. 8).

Williams, D. J. (1981). "Ring current composition and sources: An update." In: *Planetary and Space Science* 29.11, pp. 1195–1203. DOI: 10.1016/0032-0633(81)90124-0 (cit. on p. 15).

Winch, D. E., D. J. Ivers, J. P. R. Turner, and R. J. Stening (2005). "Geomagnetism and Schmidt quasi-normalization." In: *Geophysical Journal International* 160.2, pp. 487–504. DOI: 10.1111/j.1365-246X.2004.02472.x (cit. on p. 8).

Yamazaki, Y. and A. Maute (2017). "Sq and EEJ–A Review on the Daily Variation of the Geomagnetic Field Caused by Ionospheric Dynamo Currents." In: *Space Science Reviews* 206.1, pp. 299–405. DOI: 10.1007/s11214-016-0282-z (cit. on p. 9).

Zurbuchen, T. H. H. and I. G. Richardson (2006). "In-Situ Solar Wind and Magnetic Field Signatures of Interplanetary Coronal Mass Ejections." In: *Space Science Reviews* 123.1, pp. 31–43. DOI: 10.1007/s11214-006-9010-4 (cit. on p. 13).

ERKLÄRUNG (*DECLARATION*)

Ich erkläre, dass diese Promotionsschrift ausschließlich von mir verfasst wurde und dass sie bisher weder im Ganzen noch in Teilen zur Erlangung eines akademischen Grades eingereicht wurde. Abgesehen von den mit Referenzen versehenen oder in der Erklärung zu den Publikationsanteilen erwähnten Inhalten, ist die Arbeit gänzlich meine eigene.

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where stated otherwise by reference or acknowledgment, the work presented is entirely my own.

Leonie Johanna Lisa Pick
Potsdam, Juli 2020