MOSFIT: MAGNETIC OBSERVATORIES AND STATIONS FILTERING TOOL

Marcos Vinicius S. da Silva

Dissertação apresentada ao Programa de Pós-graduação em Geofísica do Observatório Nacional, como parte dos requisitos necessários à obtenção do título de Mestre em Geofísica.

Orientador(a): Dra. Katia Jasbinschek dos Reis Pinheiro

Co-orientador(a): Dr. Achim Morschhauser

Rio de Janeiro

Abril de 2023

"MOSFIT: MAGNETIC OBSERVATORIES AND STATIONS FILTERING TOOL"

Marcos Vinicius Siqueira da Silva

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM GEOFÍSICA DO OBSERVATÓRIO NACIONAL COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM GEOFÍSICA.

Aprovada por:

Dra. Katia Jasbinschek dos Reis Pinheiro - (Orientadora) - (ON)

Dr. André Wiermann – (ON)

Dr. Fernando Jorge Gutiérrez Pinheiro - (Universidade de Coimbra, Portugal)

RIO DE JANEIRO – BRASIL

25 DE ABRIL DE 2023

Clicksign

Marcos Vinicius Siqueira da Silva - Ata.pdf

Documento número #b620c84e-f583-4926-9f74-a07f8f1b985a

Hash do documento original (SHA256): d27921a5efd737cd72832953eea9498cf1084ddd8404ab7a1595b3bd40b180dd

Assinaturas

KATIA JASBINSCHEK DOS REIS PINHEIRO
CPF: 086.710.057-55
Assinou em 28 abr 2023 às 14:23:39

ANDRE WIERMANN
CPF: 669.385.037-53
Assinou em 26 abr 2023 às 16:35:40

GIANE DO CARMO BOLDRIM
CPF: 004.446.047-31
Assinou em 03 mai 2023 às 13:15:28

Fernando Jorge Gutiérrez Pinheiro
Assinou em 28 abr 2023 às 07:24:46

MARCOS VINICIUS SIQUEIRA DA SILVA
CPF: 168.736.797-39
Assinou em 27 abr 2023 às 09:11:37

Log

26 abr 2023, 16:23:07
Operador com email gianeboldrim@on.br na Conta d4a6a640-21d1-4ea0-ac83-474533e86e3c criou este documento número b620c84e-f583-4926-9f74-a07f8f1b985a. Data limite para assinatura do documento: 26 de maio de 2023 (16:13). Finalização automática após a última assinatura: habilitada. Idioma: Português brasileiro.

26 abr 2023, 16:23:08
Operador com email gianeboldrim@on.br na Conta d4a6a640-21d1-4ea0-ac83-474533e86e3c adicionou à Lista de Assinatura:
kpinheiro@on.br para assinar, via E-mail, com os pontos de autenticação: Token via E-mail; Nome Completo; CPF; endereço de IP. Dados informados pelo Operador para validação do signatário: nome completo KATIA JASBINSCHEK DOS REIS PINHEIRO e CPF 086.710.057-55.

...

Resumo da Dissertação apresentada ao Programa de Pós-Graduação em Geofísica do Observatório Nacional como parte dos requisitos necessários para a obtenção do título de Mestre em Geofísica.

MOSFIT: MAGNETIC OBSERVATORIES AND STATIONS FILTERING TOOL

Marcos Vinicius S. da Silva

Abril/2023

MOSFiT (Magnetic Observatories and Stations Filtering Tool) é um pacote desenvolvido em Python para visualizar e filtrar dados de observatórios e estações geomagnéticas. O objetivo do MOSFiT é isolar e analisar as informações de variação secular (VS) registradas pelos dados de observatórios geomagnéticos de forma automática. As contribuições de campo externo podem ser filtradas selecionando dados de acordo com o horário local, índices geomagnéticos e subtraindo as fontes do campo magnetosférico previstas pelo modelo CHAOS-7. MOSFiT calcula a VS através de diferenças anuais de médias mensais e detecta jerks geomagnéticos ajustando dois segmentos de reta em um intervalo de tempo definido pelo usuário analisando o comportamento da VS, os parâmetros de detecção dos jerks geomagnéticos, tempo de ocorrêcia e amplitude, são calculados automaticamente. Foi apresentado um novo pacote em Python, validado contra resultados independentes de publicações anteriores e uma aplicação do pacote na detecção de jerks geomagnético foi desenvolvida. Em particular, foi quantificado a discrepância RMS entre a VS filtrada e a VS prevista pelo modelo geomagnético CHAOS-7. A análise dos dados quasi-definitivos do INTERMAGNET com o MOSFIT permite uma investigação eficaz da VS em tempo quase real, assim como a detecção de jerks geomagnéticos recentes. O pacote também pode ser usado para seleção de dados, por exemplo, para estudos de campo externo, bem como para controle de qualidade de dados de observatórios geomagnéticos.

Abstract of the Dissertation presented to the National Observatory's Graduate Program in Geophysics as a partial fulfillment of the requirements for the degree of Master in Geophysics.

MOSFIT: MAGNETIC OBSERVATORIES AND STATIONS FILTERING TOOL

Marcos Vinicius S. da Silva

April/2023

MOSFiT (Magnetic Observatories and Stations Filtering Tool) is a python package to visualize and filter data from magnetic observatories and magnetometer stations. The purpose of MOSFiT is to automatically isolate and analyze the secular variation (SV) information measured by geomagnetic observatories data. External field contributions may be reduced by selecting data according to local time and geomagnetic indices and by subtracting the magnetospheric field predictions of the CHAOS-7 model. MOSFiT calculates the SV by annual differences of monthly means and geomagnetic jerk occurrence time and amplitude are automatically calculated by fitting two straight-line segments in a user-defined time interval of the SV time series. Here, we present the new python package, validate it against independent results from previous publications, and show its application. In particular, we quantify the RMS misfit between SV derived from processing schemes and the SV predicted by CHAOS-7. Analysing the INTERMAGNET quasi-definitive data with MOSFIT allows for a timely investigation of SV such as the detection of recent geomagnetic jerks. It can also be used for data selection for e.g., external field studies as well as for quality control of geomagnetic observatory data.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Earth's magnetic field is composed by the superposition of internal and external sources, varying in a wide range of timescales: from milliseconds to millions of years (MATZKA *et al.*, 2010). The geomagnetic field is produced by convection currents in the molten outer core, known as geodynamo. The core field represents more than 95% of the total field measured in the Earth's surface. This main field is on the order of 60.000 nT at the poles and 26.000 nT in the equatorial region. The time variation of the core field is called secular variation (SV) and it ranges from few years to millions of years. The other part that composes the internal sources are the crustal field, characterized by rock magnetization in the lithosphere originated during cooling below the Curie temperature and the induced field in the crust and mantle (CHULLIAT *et al.*, 2017).

The external field is produced by the movement of charged particles (electric currents) in the ionosphere and magnetosphere. Generally the external part of the magnetic field represents only a few tens of nanoteslas on the measured magnetic field during the quiet days, and may reach a thousand of nanoteslas during disturbed days. These variations range from seconds to tens of years (KONO, 2010). Figure 1.1 illustrates the different sources of the geomagnetic field.
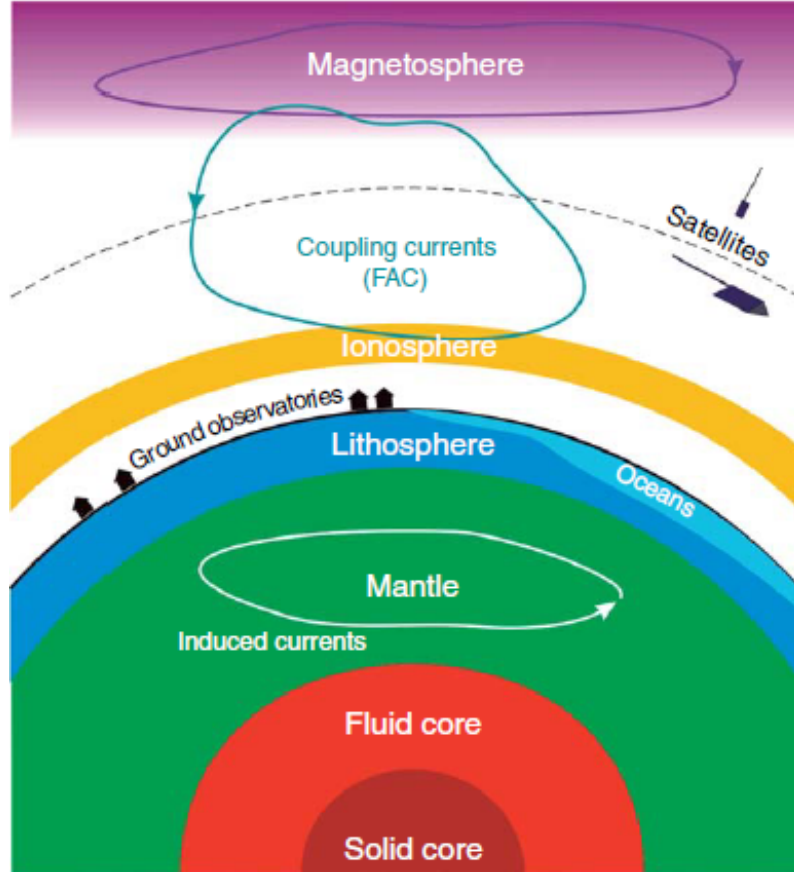
Figure 1.1: Illustration of the geomagnetic field sources. From KONO (2010).

The geodetic coordinate system, which considers the Earth's shape as a geoid, is the most used for the geomagnetic field measurements. The geomagnetic field vector is commonly expressed in Cartesian coordinates, with the components being geographic North (X), geographic East (Y), and vertical downwards (Z). Another approach, known as cylindrical coordinates, uses the components H (horizontal magnetic component), D (magnetic declination), and Z (vertical downwards). Declination is the angle between geographic north and the horizontal component (CAMPBELL, 2003). The total field intensity, represented as F, is calculated as the square root of the sum of the squares of the X, Y, and Z components (F = $\sqrt{X^2 + Y^2 + Z^2}$). The representation of these magnetic elements is shown in Figure 1.2. Magnetic measurements taken by satellites use a different coordinate system known as the geocentric coordinate system. In this system, the spherical components are B_r, B_θ, and B_ϕ. The conversion from spherical to Cartesian coordinates is given by:

$$B_\theta = -X, \; B_\phi = Y \, and \, Br = -Z. \tag{1.1}$$

2
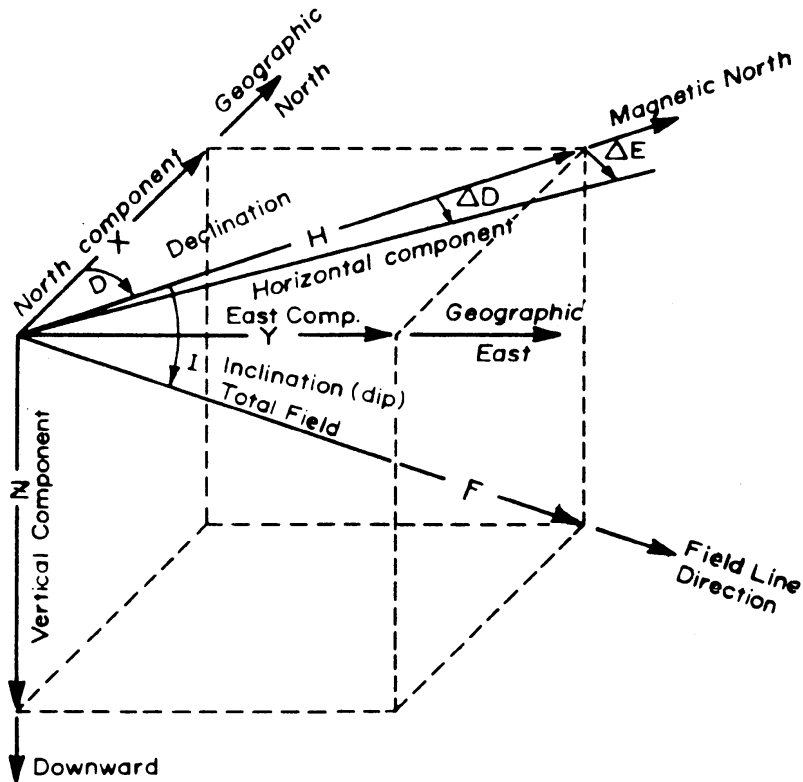
Figure 1.2: Components of the geomagnetic field. From CAMPBELL (2003).

1.1 Geomagnetic secular variation (SV)

The geomagnetic SV refers to the long-term changes in the Earth's magnetic field that occur over a period of decades or centuries. These changes can be observed in time variations in the strength and direction of the geomagnetic field around the globe. The geomagnetic SV is calculated by the first time derivative of the main field, which changes on secular timescales, from years to centuries (BLOXHAM *et al.*, 1989; KONO, 2010). Eventually shorter timescales changes can also occur, being defined as geomagnetic jerks. The SV has some specific characteristics: the westward drift, the slow decay of the geomagnetic dipole and the drift of the North magnetic pole (CHULLIAT e MAUS, 2014). The second time derivative of the geomagnetic field is known as geomagnetic secular acceleration (SA) and can also indicates this shorter timescale changes (see geomagnetic jerk section)

Figure 1.3 shows the SV of the East geomagnetic component (dY/dt) for Chambon la Forêt observatory (CLF) and Niemegk observatory (NGK). It shows the abrupt change in the SV around 1900, 1970 and 2000.

1.1.1 Geomagnetic Jerks

The study of geomagnetic jerks has become an important area of research as they can provide valuable insights into the dynamics of the Earth's core. Geomagnetic

3
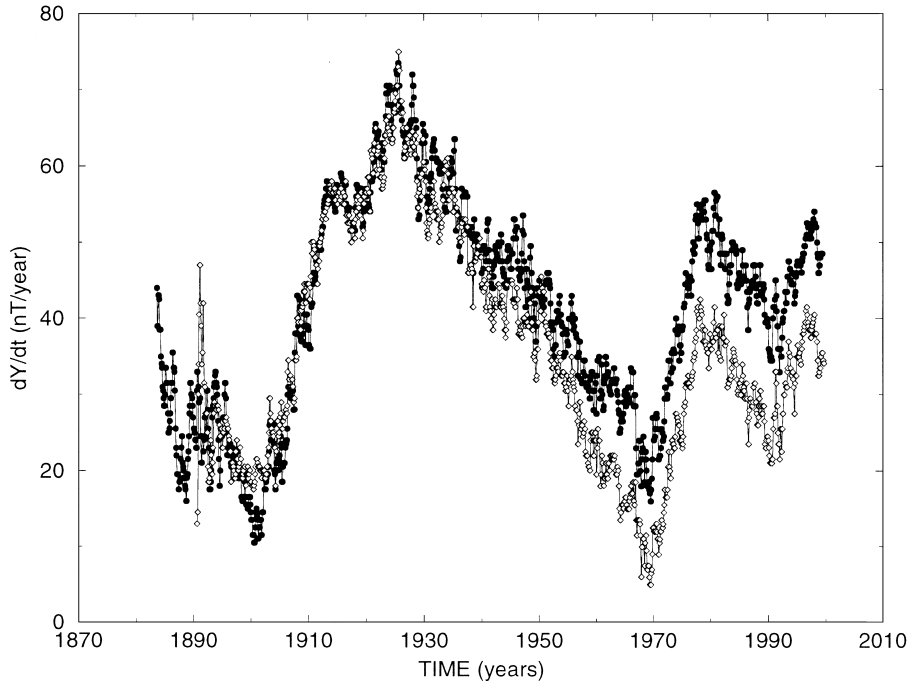
Figure 1.3: Chambon la Forêt (CLF, solid circles) and Niemegk (opaque circles) observatories SV (dY/dT). From MANDEA *et al.* (2000).

jerks can be viewed as a V-shape abrupt change in the first time derivative (SV) and as a step-like variation in the second time derivative (secular acceleration SA). Jerks are not typically simultaneously observed around the globe, meaning that the same event may be observed at different times and locations (BLOXHAM *et al.*, 2002; COURTILLOT *et al.*, 1978; PINHEIRO *et al.*, 2019). Figure 1.4 illustrates how a geomagnetic jerk is ideally represented in the SV (upper panel) and SA (lower panel). A sudden change in the SV trend, characterized by the V-shape or a change in the SA sign (step-like). The cause of geomagnetic jerks is still not known, but they are related to short timescales of core dynamics. Recent research has suggested possible explanations, as for example AUBERT e FINLAY (2019), that reported possible options: rapid local changes (accelerations) in core flow driven by intense convective fluctuations, or by rapid fluctuations in the Lorentz force due to reorganization events in the core field or to diffusional processes such as a change in the rate of flux expulsion, that may occur on short timescales if very large radial field gradients are present.

4
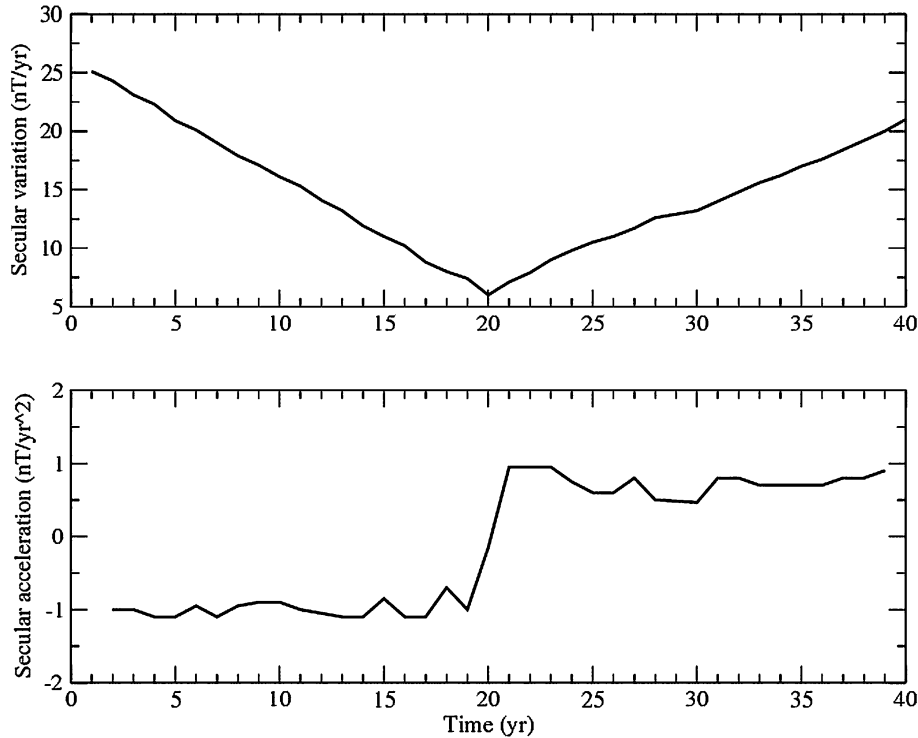
Figure 1.4: Synthetic representation of a geomagnetic jerk in the SV (upper panel), defined as the V-shape, and in the SA (lower panel), characterized by the sign change (step-like). From MANDEA e KORTE (2010).

1.2 Measuring the geomagnetic field

Direct measurements of the geomagnetic field dates back to 400 years ago, when the declination was first determined. Over time, this developed into full vector measurements taken by geomagnetic observatories and repeat stations surveys in the 19th century. In the second half of the 20th century, low-Earth-orbit satellites provides a global coverage of the geomagnetic field (MATZKA *et al.*, 2010).

1.2.1 Geomagnetic observatories

Geomagnetic observatories are fixed locations at the Earth's surface that measure the geomagnetic field for long periods, with some having data that goes back more than 150 years. Their data is of high quality and is useful for studying both SV and space weather. At a magnetic observatory, the three orthogonal components of the geomagnetic field, and its total intensity, are measured continuously using a triaxial fluxgate magnetometer and a proton or Overhauser magnetometer, respectively. Declination-Inclination absolute measurements, which are taken twice a week, are made using a DI-flux theodolite and a magnetometer (CHULLIAT *et al.*, 2017).

The general layout of the geomagnetic observatories consists of two main buildings: a variometer hut with a fluxgate and Overhauser, and the absolute hut, where

absolute measurements of D and I are performed leastways twice a week, using instruments the DI-flux and Overhauser. The sensor and the DI-flux are placed in stable pillars. Baselines are calculated by combining absolute with relative (or variometer) measurements. The relative data is a combination of X, Y and Z components, or H and Z components and D, accompanied by total field (F) recording.

Magnetic observatories are not evenly distributed around the globe, and Figure 1.5 demonstrates the distribution of INTERMAGNET magnetic observatories in operation. Gaps in coverage are supplemented by satellite measurements, such as those taken by Swarm, which fly at altitudes between 350km and 850km. The combination of data from magnetic observatories and satellites allows the calculation of geomagnetic models and helps separating internal and external contributions. These global models are useful for expanding knowledge in various areas, such as geomagnetic jerks, South Atlantic Anomaly (SAA) and navigation.
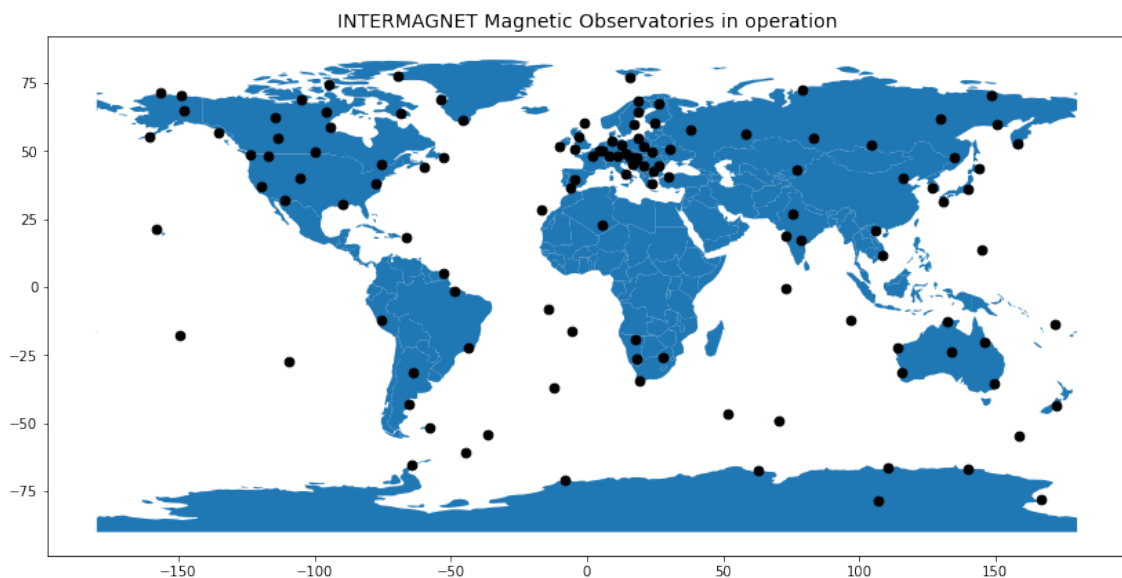
Figure 1.5: Global distribution of INTERMAGNET observatories in operation in 2021.

1.2.2 Repeat stations

Geomagnetic repeat stations are fixed locations at remote sites where detailed measurements and observations of the geomagnetic field are taken every few years. These sites must be precisely located to ensure that measurements are always taken in the same location and should be free from human-made magnetic interference. The purpose of repeat station measurements is to monitor changes in the SV of the geomagnetic field and to use them in core field modeling (KONO, 2010; MANDEA e KORTE, 2010; NEWITT *et al.*, 1996).

1.2.3 Satellites

Prior to the first space-based measurements made by Sputnik 3 satellite in 1958, ground-based data was the only available magnetic measurements. Due to the uneven distribution of geomagnetic observatories, satellites play a crucial role in understanding SV and modelling the geomagnetic field, in addition to allow a better understanding of the magnetosphere. Some advantages of satellites in relation to ground observatories are: global coverage, same magnetometer measuring the magnetic field all around the world, less influence of local anomalies in the field measured. However, since the satellite is in movement it is not possible to measure time variation or spatial changes in the same point with same accuracy as ground observatories (KONO, 2010).

The Figure 1.6 demonstrate the ϕrsted satellite mission ground track coverage of 24 hours on January 2, 2021. The red dots are the in operation geomagnetic observatories from INTERMAGNET/WDC between 1997 and 2012, highlighting the important gap in global coverage provided by satellite missions.
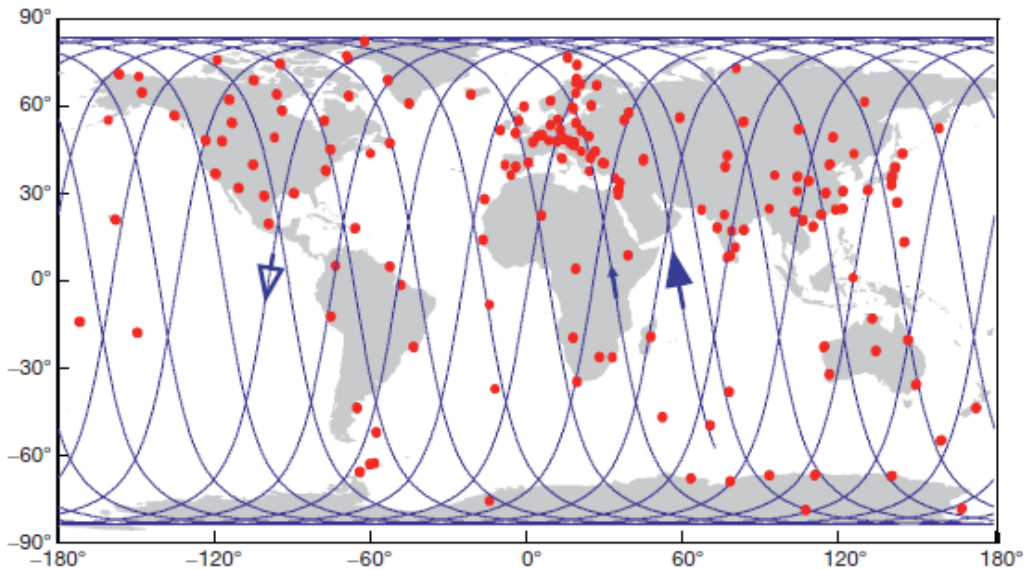
Figure 1.6: Observatories that provided data to the WDC/INTERMAGNET system between 1997 and 2012 (red dots) and ground track of 24 h of the ϕrsted satellite on 2 January 2001 (blue curve). From KONO (2010).

1.3 International Real-time Magnetic Observatory Network (INTERMAGNET)

The International Real-time Magnetic Observatory Network (INTERMAGNET) is a global network of digital geomagnetic observatories that provide high-quality geo-

magnetic data for the scientific community. The network provides standard specifications for equipment and measurements, and the data provided by INTERMAGNET observatories are filtered to avoid aliasing effects. The network offers three types of data: preliminary, quasi-definitive (QD) and definitive.

Preliminary data are sent in real-time with a preliminary baseline and are often used for forecasting. Quasi-definitive data, which were first available in 2008 (PELTIER e CHULLIAT, 2010), are intended to accelerate the distribution of high-quality data for the scientific community. Prior to the availability of QD data, researchers had to wait approximately one year for the publication of definitive data, which is calculated using final baselines that take into account all months of the year. QD data must be corrected using temporary baselines, and must be delivered at least three months after their acquisition, with a difference compared to the definitive data of less than 5 nT for every monthly mean of the year (for geomagnetic components X, Y, and Z).

Finally, definitive data are produced on an annual basis, and are corrected with definitive baselines, with all spikes in the data removed and gaps filled, when possible. The definitive data provides the most accurate representation of the geomagnetic field, and is used for long-term studies. The INTERMAGNET network is important for the study of the geomagnetic field, as the global distribution of geomagnetic observatories is uneven, thus the INTERMAGNET data provides a comprehensive coverage of the geomagnetic field.

1.4 Geomagnetic indices

The goal of an index is to describe or give an information about a phenomena which varies with time at certain time interval. In geomagnetism, indices measures the geomagnetic activity, i.e. the magnetic contributions from the ionospheric and magnetospheric sources, these indices are generally recorded at geomagnetic observatories and are typically updated at regular, constant intervals (MAYAUD, 1980).

There are several different geomagnetic indices that have been developed to describe different phenomena. For example, the K and Kp indices are used to measure planetary-scale magnetic activity, while the Dst, RC, and A_{sym} indices are used to describe equatorial storms and magnetospheric ring currents. Additionally, the AE-indices are used to describe auroral electrojets.

In this study, I will focus on the Kp and Dst (as well as the refined version of the Dst, the RC index) as they are the most relevant to this work.

1.4.1 The Kp index

The Kp index concept was introduced by BARTELS (1938), as a means of quantifying the level of irregular geomagnetic disturbances caused by solar particle radiation on a global scale. The index is a 3-hour interval derived from measurements taken at 13 sub-auroral geomagnetic observatories worldwide (MATZKA *et al.*, 2021). The process of deriving the K value involves determining the disturbance level range (see a in Fig 1.7). by measuring the difference between the highest and lowest values (during the 3-hour intervals) of the horizontal component at each geomagnetic observatory, after subtracting the solar quiet variation (S_q), which is the mean daily variation of the 5 quietest days of each month (Figure 1.7). The resulting a range is obtained is converted to a K value that varies from 0 to 9 in a quasi-logarithmic scale, with 0 indicating the lowest geomagnetic activity and 9 the highest. The K values are converted to Ks values (conversion tables are used to perform such conversion, find details on `https://www.gfz-potsdam.de/en/section/geomagnetism/data-products-services/geomagnetic-kp-index`) in a scale of thirds (28 values varying from 0 to 9). The Kp index, that represents the geomagnetic disturbance in a global scale, is the mean of the Ks values from the 13 sub-auroral stations (KAURISTIE *et al.*, 2017; MATZKA *et al.*, 2021). The Kp index is made available in real-time, being provided by GFZ-Potsdam on `https://kp.gfz-potsdam.de/en/`.
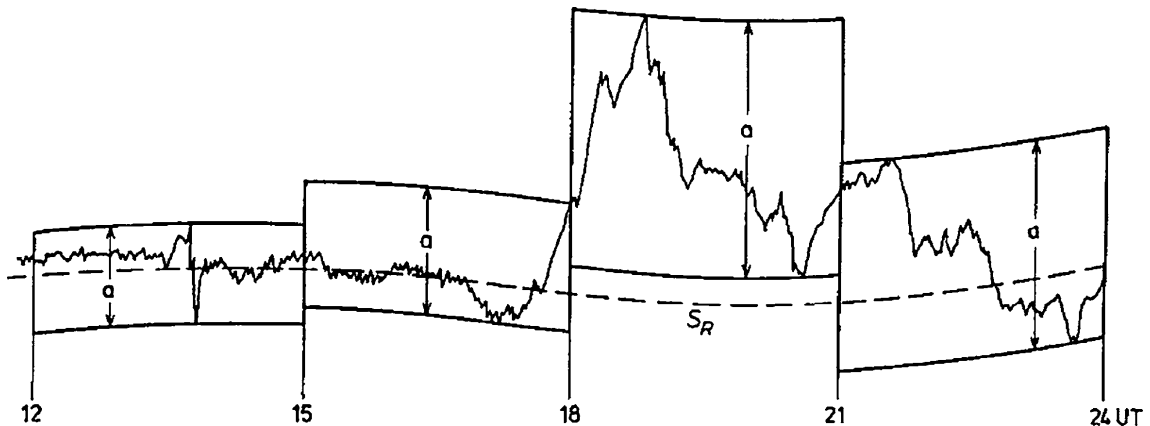
Figure 1.7:   Illustration of the 3-hour intervals used on Kp index calculation. The dashed line is the S_q elimination (S_r in the figure). The differences between the highest and lowest values represents the a range. From KAURISTIE *et al.* (2017)

1.4.2 Magnetospheric ring-current indices

The Disturbance storm time (Dst) index is a measure of the strength decrease in the main geomagnetic field caused by the magnetospheric ring current and its induced counterpart. This phenomenon is characterized by the north component of

the axially symmetrical part of the equatorial disturbance field at the Earth's surface, as described by KONO (2010). To determine the Dst index, data from four low-latitude observatories (Honolulu, Hermanos, San Juan and Kakioka) are used. These observatories are strategically located in such a way as to be well distributed longitudinally and far enough from the influence of auroral and equatorial electrojets. The data is collected on an hourly basis and undergoes a process of SV and S_q removal, followed by a geometric adjustment to normalize it to the magnetic equator. The Dst index is then calculated as the average of the normalized data. It is represented mathematically by (1.2):

$$Dst(t) = \frac{\sum_{i=1}^{4} D_i(t)}{\sum_{i=1}^{4} cos(\lambda_i)}, \tag{1.2}$$

Where Dst(t) is the Dst index at a time "t", D_i(t) is the disturbance variation of the H component of the four (i = 1 - 4) mentioned observatories and λ is the dipole latitudes of the observatories.

The Figure 1.8 shows an example of the Dst index when a strong magnetic storm occurred in June 2015. The y-axis indicates the index intensity, LOEWE e PRÖLSS (1997) classifies geomagnetic storm intensities as: for a small storm -30 nT > Dst > -50 nT, for a moderate storm -50 nT > Dst > - 100 nT, for a strong storm -100 nT > Dst > -200 nT, and a great storm Dst < -350 nT.

Figure 1.8: Dst index final hourly timeseries for June, 2015. Available on `https://wdc.kugi.kyoto-u.ac.jp/dst_final/index.html`.

The RC index, introduced by OLSEN *et al.* (2014), aims to overcome limitations of the Dst index in accurately describing the strength of the magnetospheric ring-current during geomagnetic quiet conditions. The Dst index has been known to under-represent the intensity of the ring-current during such conditions, with estimates suggesting it can be as much as 20 nT, as suggested by LANGEL e ESTES (1985). In contrast to the Dst index, which is based on measurements from four low-latitude observatories, the RC index utilizes data from 21 observatories located

at both low and mid latitudes, and is derived through Spherical Harmonic Analysis (SHA) of their hourly mean values.

1.5   Global geomagnetic models

The ability to represent the time-dependent geomagnetic field and its various sources (such as the core, crust, ionosphere, and magnetosphere) at any location on and near Earth's surface is a fundamental aspect of global geomagnetic modeling. Various mathematical methods have been developed for this purpose, but the most commonly used technique is spherical harmonic analysis (SHA), which is a mathematical method used to represent and analyze geomagnetic fields. This method is based on the expansion of the geomagnetic field into a series of spherical harmonics (i.e., functions defined on the surface of a sphere). The SHA is particularly useful for modeling geomagnetic fields because it is able to accurately represent the field when the modeled data has a good global distribution, which enables the separation and analysis of different sources of the geomagnetic field. Additionally, the SHA is widely used for interpreting geomagnetic observations and studying geomagnetic SV. The use of satellite data, which provides global coverage, is crucial in this context as the distribution of geomagnetic observatories on Earth is often uneven (KONO, 2010).

The magnetic scalar potential is written in terms of internal and external contributions, $V = V_{int} + V_{ext}$. The internal and external scalar potentials (V_{int} and V_{ext}) are represented by spherical harmonic expansions (SABAKA *et al.*, 2010).

For internal contributions

$$V_{int} = a \sum_{n=1}^{N_{int}} \sum_{m=0}^{n} (g_n^m \cos m\phi + h_n^m \sin m\phi) \left(\frac{a}{r}\right)^{n+1} P_n^m(\cos\theta) \qquad (1.3)$$

For external contributions

$$V_{ext} = a \sum_{n=1}^{N_{ext}} \sum_{m=0}^{n} (q_n^m \cos m\phi + s_n^m \sin m\phi) \left(\frac{r}{a}\right)^{n+1} P_n^m(\cos\theta) \qquad (1.4)$$

Where a is Earth's reference radius in kilometers. The (g_n^m, h_n^m) and (q_n^m, s_n^m) are the Gauss spherical harmonic coefficients for internal and external sources, respectively, (r, θ, ϕ) are coordinates in a geocentric spherical coordinate systems, P_n^m are semi-normalized associated Legendre functions of degree n and order m, the N indicates the maximum spherical harmonic expansion degree.

Chapter 2

Objectives

The objective of this project is to design and implement a new Python automated tool for quality control and detection of geomagnetic jerks in geomagnetic data, validate the package over two case studies, and demonstrate the use of the package presenting a new geophysical application detecting jerks. The tool uses recent quasi-definitive minute mean data from INTERMAGNET observatories, and allows the user to select various processing options to minimize the external field contributions in the data. The user can automatically detect geomagnetic jerks by fitting two straight-line segments or through visual inspection, thus facilitating the investigation of core field phenomena, particularly geomagnetic jerks.

To validate the effectiveness of the Python tool, we reproduce the investigations of FINLAY *et al.* (2020) and TORTA *et al.* (2015) by comparing the results obtained from my automated methods with those from the original studies. The validation process is presented in the form of two case studies.

In Case Study 1, I reproduced the investigation of FINLAY *et al.* (2020) by validating the automatic CHAOS-7 model geomagnetic field prediction for INTER-MAGNET observatories. This is achieved by comparing the Root Mean Squared error (RMSe) between the predicted SV and observatory data SV for over 100 observatories. Additionally, I present a global misfit of the Python package's automatic magnetospheric correction method and quantify how the correction improves the SV.

In Case Study 2, I detect the geomagnetic jerks that occurred in 2007, 2011, and 2014 in the observatory data from NGK, EBR, TAM, and ASC. I compare the occurrence time and amplitude statistics of the jerks obtained from the package's automatic method with those from the published investigation of TORTA *et al.* (2015), validating the data resampling, SV calculation, and automatic jerk detection.

Furthermore, I conduct a new investigation to evaluate the impact of various external field filtering methods on geomagnetic jerk detection. This study aims to understand how different methods can affect the detection of amplitude and

occurrence time of geomagnetic jerks.

Chapter 3

Magnetic Observatories and Stations Filtering Tool (MOSFiT) methods

In this chapter we describe the most important methods present in MOSFiT package. Our goal is to demonstrate and explain the features that were developed and applied during this study.

3.1 Outlier detection - Hampel filter

The Hampel filter is a robust outlier detector for time series. It is based on the median absolute deviation (MAD), since it is less affected by outliers than the mean. See HAMPEL (1974) and PEARSON *et al.* (2016) for a deep explanation about outlier detection. Equation 3.1 is the MAD, where x_i is the ith observed data point and m the median for each window.

$$MAD = median(|x_i - m|) \tag{3.1}$$

Basically, it works like a sliding window (analogous to the moving average method) over the time series. The user specifies the sliding window size and the threshold to tune the filter sensibility. Any data point exceeding the threshold is flagged as an outlier and replaced by the window median value. This approach follows COX *et al.* (2018) method of denoising the WDC data, in order to preliminary filter the SV.

Hourly mean values are used in order to decrease computational cost, while applying the Hampel filter. For example, it would take around 1.5 minutes to denoise 10 years of hourly mean values, while for minute means it would be 50 times slower.

Figure 3.1 is an application of the MOSFiT Hampel filter method in the Neumayer Station III (VNA) magnetic observatory, from 2014 to 2022. The hampel filter window size was 24 hours and the a threshold of 2.
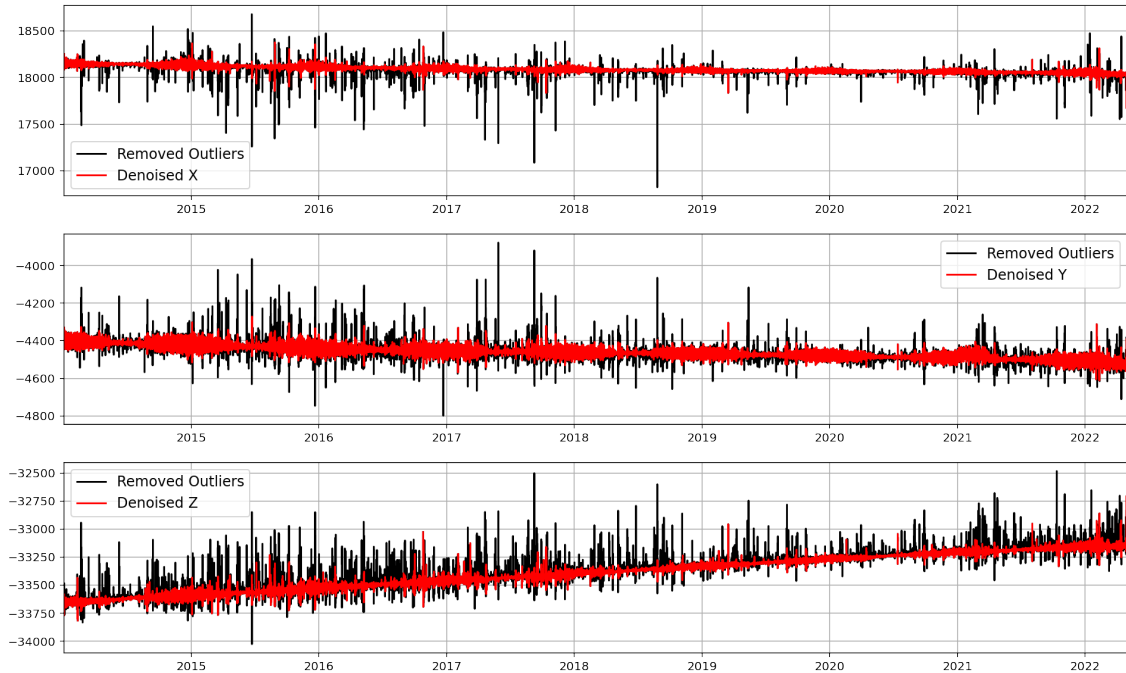
Figure 3.1: MOSFiT hampel filter method application to VNA magnetic observatory hourly mean data. Black lines are the X, Y and Z geomagnetic components before the Hampel filter application, from the top to the bottom, respectively. Red curves are the filtered components. Figure produced automatically by MOSFiT.

3.2 External field filtering - Data selection

MOSFiT provides different options of geomagnetic data selection. One method to mitigate external field contributions, is by selecting days or periods when the geomagnetic activity is weak (KAURISTIE *et al.*, 2017). The MOSFiT options are: (i) Kp index ≤ 2, (ii) selection of only nighttime period, (iii) keep only top 10 quiet days of each month or by (iv) removing top 5 disturbed days of each month. Although the main focus of the package is the SV, data selection is widely applied for many different purposes, e.g., Solar quiet daily variation (Sq) (YAMAZAKI e MAUTE, 2017) and internal field modelling (KAURISTIE *et al.*, 2017).

3.2.1 Kp index

The Kp is a three-hour interval index representing all irregular geomagnetic disturbances caused by solar particle radiation, distributed by GFZ-Potsdam (MATZKA

et al., 2021). It is derived from 13 sub-auroral stations [1].

In combination with others indices, Kp is widely applied to data selection in geomagnetic field modelling and for studies of the ionosphere, thermosphere and magnetosphere (MATZKA *et al.*, 2021). The Kp data selection criteria for days with weak geomagnetic activity is usually set to Kp ≤ 2 (defaut of MOSFiT) that constitutes in average 30% of the data availability (YAMAZAKI e MAUTE, 2017). But in MOSFiT, the data selection can be used from $Kp=0$ to $Kp=9$. The days with Kp value above the defined limit are removed from the dataset. The Kp index is automatically updated when the function is used, which provides a near real-time application of the method. Figure 3.2 illustrates the Kp index selection ($Kp \leq 2$) using the MOSFiT function, applied to Vassouras magnetic observatory (VSS) from January 2010 to May 2022.
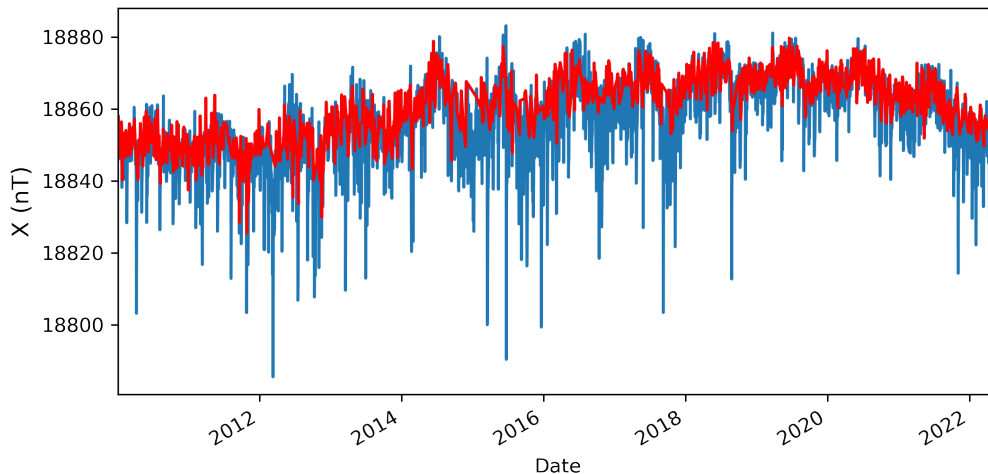
Figure 3.2: Vassouras magnetic observatory (VSS) X component from January 2010 to May 2022. Blue curves are the observed daily mean data and the orange curves represents the observed data after applying data selection for Kp-index $\leq 2^{\text{o}}$. These results were produced using MOSFiT.

3.2.2 International quiet and disturbed days - IQD's and IDD's

The IQD's and IDD's are derived from the Kp index and delivered by the GFZ-Potsdam, on a monthly basis (find the list of IQD's and IDD's since 1932, on `https://www.gfz-potsdam.de/en/kp-index`). BARTELS (1957) determined three crite-

---

[1]The list of those stations and their locations can be found on `https://www.gfz-potsdam.de/en/kp-index`.

ria to establish the most quiet and disturbed days for each month: the sum of the Kp values, the sum of squares of Kp values and the maximum Kp value, all calculated for each day.

MOSFiT provides options to select data based on IDD's or IQD's, which completely remove the top 10 IDD's and keep the top 5 IQD's for each month of the dataset. Similar to the Kp index, the IDD's and IQD's lists are updated automatically when accessed by the package.

3.2.3   Nighttime selection

The nighttime selection is usually selected as from 23:00 to 05:00 LT (default of MOSFiT) since in low and mid latitudes the external fields are minimum at that time, therefore the external field influence is decreased. MOSFiT automatically select the nighttime for all analysed observatories. The user can modify the default MOSFiT definition of nighttime, according to the objectives. INTERMAGNET data is distributed in universal time (UT), before select the nighttime period, it is necessary to shift the data to LT. The function finds the correct difference of time between UT and LT, shift the data to LT, select the nighttime interval, and shift it back to UT, in order to keep all the data in UT. Equation 3.2 determine the correct time to be shifted, where ϕ is the observatory longitude in degrees.

$$T_s = \frac{\phi}{15} \qquad (3.2)$$

Figure 3.3 represent the nighttime selection method applied to Tatuoca magnetic observatory (TTB) data from July, 2022. The orange dots (nighttime period) fits well to the quiet period of each day, as expected.
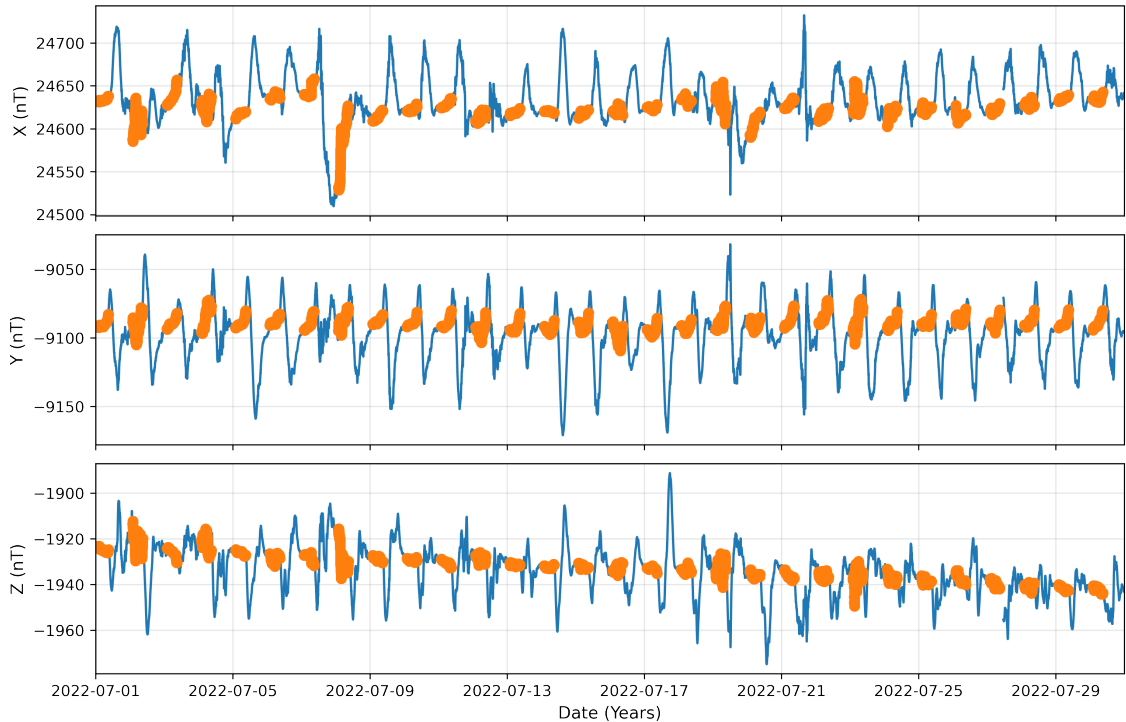
Figure 3.3: MOSFiT nighttime selection on TTB data from July, 2022. Blue lines are the geomagnetic components X, Y and Z, top, middle and bottom, respectively. The orange dots are the nighttime selection period.

3.3 CHAOS-7 model correction

CHAOS-7 geomagnetic model (FINLAY *et al.*, 2020) is derived from Swarm, CHAMP, Orsted and SAC-C satellites data and on ground observatory data, aiming to describe the different magnetic field sources. MOSFiT package includes automatically the most recent version, which is CHAOS-7.13, spanning from 1999 until November, 2022 (see `http://www.spacecenter.dk/files/magnetic-models/CHAOS-7` for release information).

In order to predict different sources, CHAOS-7 model needs as input: longitude, colatitude, radius from the Earth's center including the local altitude and temporal resolution of the field variation that will be predicted. The outputs of the model are: radial, colatitude and azimuthal components (B_r, B_θ, B_ϕ, respectively). However, in MOSFiT the user does not need to provide these inputs since the package has automatically all the information about each INTERMAGNET observatory. The B_r, B_θ, B_ϕ are converted to XYZ-components.

MOSFiT package uses CHAOS-7 model to predict different geomagnetic sources varying with time. The total geomagnetic field prediction is computed by: core field

up to spherical harmonic degree 20, internal static field above degree 25, geocentric solar magnetospheric (GSM) and solar magnetic (SM) coordinate systems, both up to degree 2. The time-dependent core field is used to calculate the secular variation (SV) and compare to observed data. The external part (GSM+SM) of the geomagnetic field is used to filter the external field contribution from the observed data.

Following FINLAY *et al.* (2020), the internal part of the geomagnetic field (V^{int}) is calculated in a spherical harmonic expansion as

$$V_{int} = a \sum_{n=1}^{N_{int}} \sum_{m=0}^{n} (g_n^m \cos m\phi + h_n^m \sin m\phi) \left(\frac{a}{r}\right)^{n+1} P_n^m(\cos\theta) \qquad (3.3)$$

Where a = 6371.2 km is the chosen Earth's spherical radius, P_n^m are the Schimdt semi-normalized Legendre functions, (r, θ, ϕ) are the spherical coordinates of the chosen location, N_{int} is the maximum degree and order of the internal expansion, g_n^m, h_n^m are the Gauss coefficients that describes the internal sources.

The external part (V^{ext}) is an adopted expansion in the SM (up to n = 2) and GSM (up to n = 2 and m = 0) coordinate systems, described as

$$
\begin{aligned}
V_{ext} = a \sum_{m=0}^{1} [q_1^{m,SM} \cos mT_d + s_1^{m,SM} \sin mT_d] \left(\frac{r}{a}\right) P_1^m(\cos\theta_d), \\
+a \sum_{m=1}^{1} [\Delta q_1^{q,SM}(t) R_{1,c}^{m,SM}(t,r,\theta,\phi)] + [\Delta s_1^{m,SM}(t) R_{1,s}^{m,SM}(t,r,\theta,\phi)], \\
+a \sum_{m=0}^{2} [q_2^{m,SM} R_{2,c}^{m,SM}(t,r,\theta,\phi) + s_2^{m,SM} R_{2,s}^{m,SM}(t,r,\theta,\phi)], \\
+a \sum_{n=1}^{2} q_n^{0,GSM} R_n^{0,GSM}(t,r,\theta,\phi),
\end{aligned}
\qquad (3.4)
$$

where θ_d is the dipole colatitude, T_d is the dipole local time represented in radians units, the $R_{n,c/s}^{m,GSM}$ and $R_{n,c/s}^{m,SM}$ are modification of the solid harmonics in SM and GSM systems.

Figure 3.4 is a sketch of the total field composition while predicting the sources using MOSFiT, which provides a dataset with the total field result, but also the different sources. The default spherical harmonic (SH) degree used to predict each component is also illustrated in this figure.
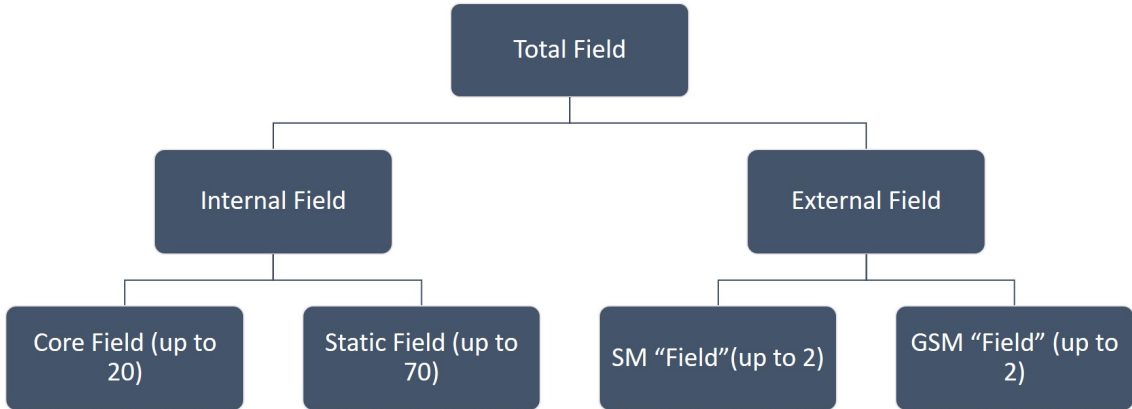
Figure 3.4: Sketch of the total field decomposition predicted using CHAOS-7 model by MOSFiT function. The SH degree for each component is shown inside parentheses.

To better understand the dynamics of the near Earth plasma and the eletromagnetic field disturbances, it is important to take into account the geometry of the Earth's geomagnetic field. The external field calculated by the CHAOS-7 model use two coordinate systems SM and GSM (FINLAY *et al.*, 2020). The GSM represents the interaction between the solar wind (dominates the interaction at large distances, high altitudes) with the magnetosphere (magnetopause region and magnetotail, distances greater than 8 Earth radii). This coordinate system has the Earth-Sun line along one of the axes and the Earth's magnetic dipole in one of the coordinate planes, being suitable to work with the mentioned interaction (LAUNDAL e RICHMOND, 2017). On the other hand, the SM coordinates describes the electrodynamics in the near-Earth space, where the Earth's magnetic field is stronger and dominates, such as in the inner magnetosphere, it has one axis along the geomagnetic dipole axis, and the Earth–Sun line in a coordinate plane (KONO, 2010; LAUNDAL e RICHMOND, 2017). An important point using the SM CHAOS-7 model function, is the possibility to use the RC-index to describe the magnetic field contribution due to the magnetospheric ring-current and its Earth-induced counterpart. The RC-index is linear interpolated from the hourly values (FINLAY *et al.*, 2020).

The next step is to subtract the predicted external field by CHAOS model from the observed dataset. Once the prediction is calculated in an hourly sample rate, the observed dataset is resampled to hourly means for comparison.

Figures 3.5 is a comparison of the CHAOS-7 model filtering, applied to the SV of Kakioka magnetic observatory (KAK). The left panel is the KAK SV (black line) produced with the observed data without any filter. In the right panel, are KAK SV after filter the external field contribution using MOSFiT CHAOS-7 method. The

red lines in the figure are the time-dependent CHAOS-7 model prediction. Such comparison demonstrates the better characterization of the SV after the filtering, explained by the good agreement with the predicted SV.
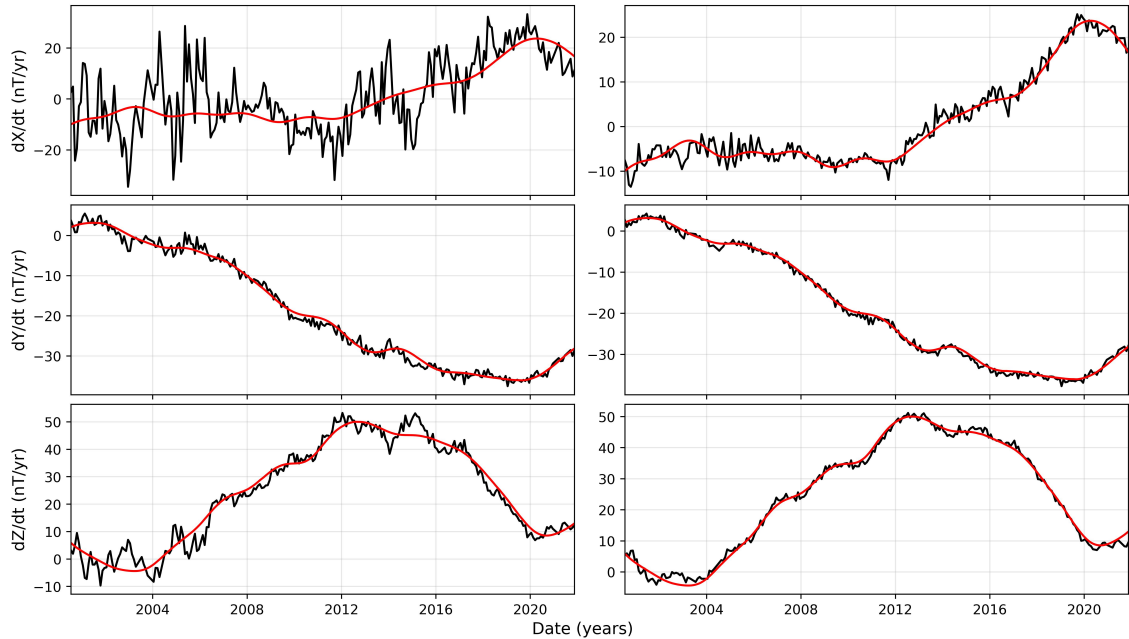
Figure 3.5: Kakioka magnetic observatory (KAK) secular variation (SV, black curves) calculated by MOSFiT before (left panels) and after (right panels) subtraction of magnetospheric field as predicted by CHAOS. SV calculated from CHAOS core field predictions are shown by red curves.

3.4 Data resampling

One of the package's feature is the possibility to resample the geomagnetic data in different samples, which are: hourly, daily, monthly and annual means. SV investigation is generally performed using monthly or annual means that are centered in the middle of the interval. For example, the monthly means are calculated for the middle of the month and annual means on July 2nd for normal years and July 1st for leap years (REDA *et al.*, 2011; ST-LOUIS *et al.*, 2020). MOSFiT only calculates the mean values if at least 90% of data is available. For example, to resample from minute means to hourly means, it is necessary to have a minimum of 54 minutes of data, following the general INTERMAGNET standard (ST-LOUIS *et al.*, 2020). I recommend the use of the default resampling condition, but there is the option to disable by the user, that might be the case when the observatory has low data availability. It is important to highlight that when there is with low data availability the means can be more affected by disturbed days.

Figure 3.6 shows the data resampling function application. The X component of VSS was resampled from 2010 to 2022 by the different options (Hourly, Daily, Monthly and Annual). Resampling the data, the high frequencies from external field are filtered, as shown in the monthly and annual means subplots.
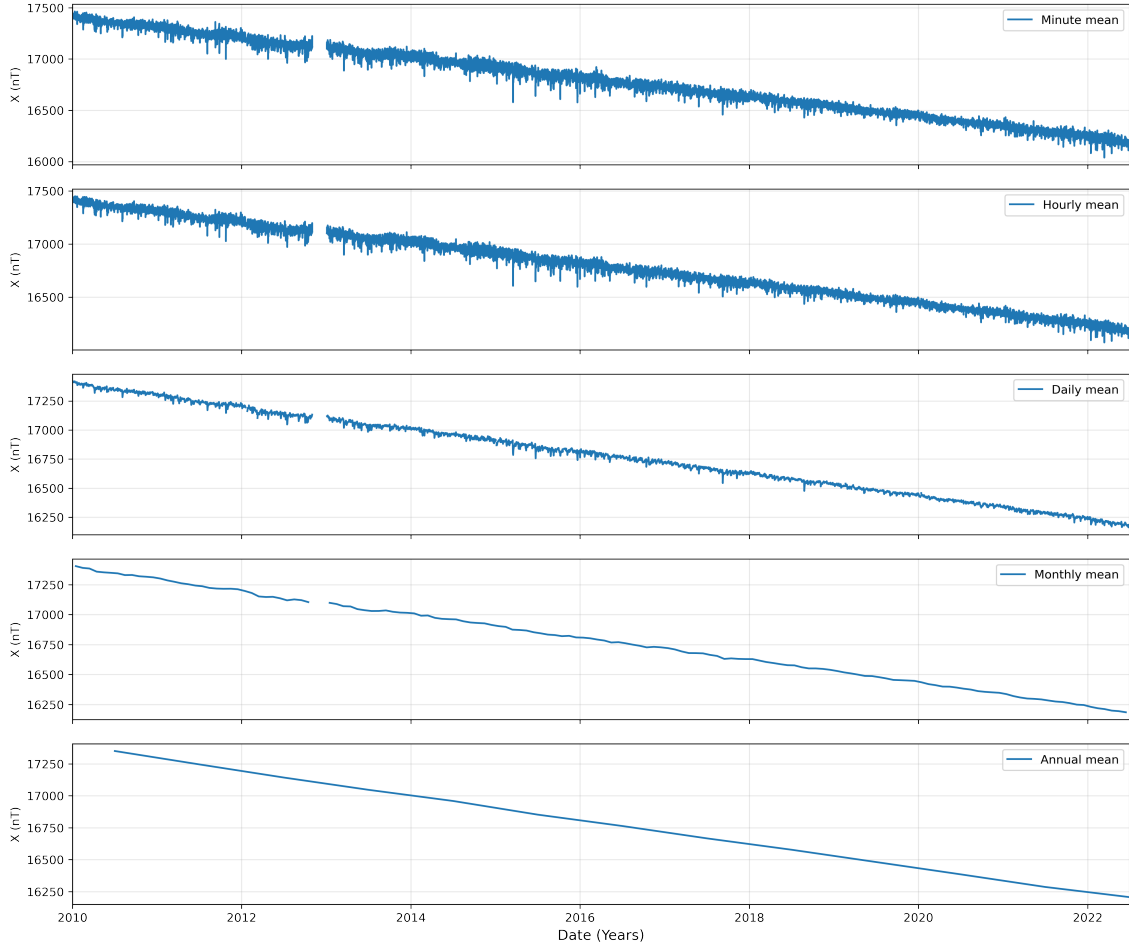
Figure 3.6: MOSFiT resample geomagnetic data function application at Vassouras magnetic observatory (VSS) data. From the top to the bottom are the minute, hourly, daily, monthly and annual means.

3.5 Secular variation calculation

The MOSFiT default method to calculate the SV is the monthly mean differences between t+6 months and t-6 months (CHULLIAT *et al.*, 2010; FENG *et al.*, 2018), well known as Annual Differences of Monthly Means (ADMM). This implies that the SV is limited to 6 months prior to the last available data. Equation 3.5 represents the SV calculated from monthly means, where t is time in months and B is the monthly mean magnetic field of any geomagnetic component. In most of the cases,

monthly means are calculated from minute means (IAGA-2002, definitive and quasi-definitive data). Monthly means are a better option for SV investigation because of the higher time resolution, therefore it is possible to identify abrupt changes into more detail. This approach mitigates the seasonal variation from magnetospheric and ionospheric currents (MANDEA *et al.*, 2000; MATZKA *et al.*, 2010).

$$\frac{dB_{(t)}}{dt} = B_{(t+6)} - B_{(t-6)} \tag{3.5}$$

A different approach while calculating SV is to use annual means, it was widely used in past. MOSFiT also provides an option to calculate the SV from annual means. But the default method is the ADMM.

Equation 3.6 shows the SV calculated from annual means, where B is the annual mean magnetic field of any geomagnetic component and t is time in years.

$$\frac{dB_{(0.5)}}{dt} = B_{(t)} - B_{(t-1)} \tag{3.6}$$

Figure 3.7 demonstrate the a SV calculation using MOSFiT. VSS one minute data from January, 2000 to July, 2022 was used. First the data is automatically resampled to monthly means and them calculated the SV, following eq. 3.5.
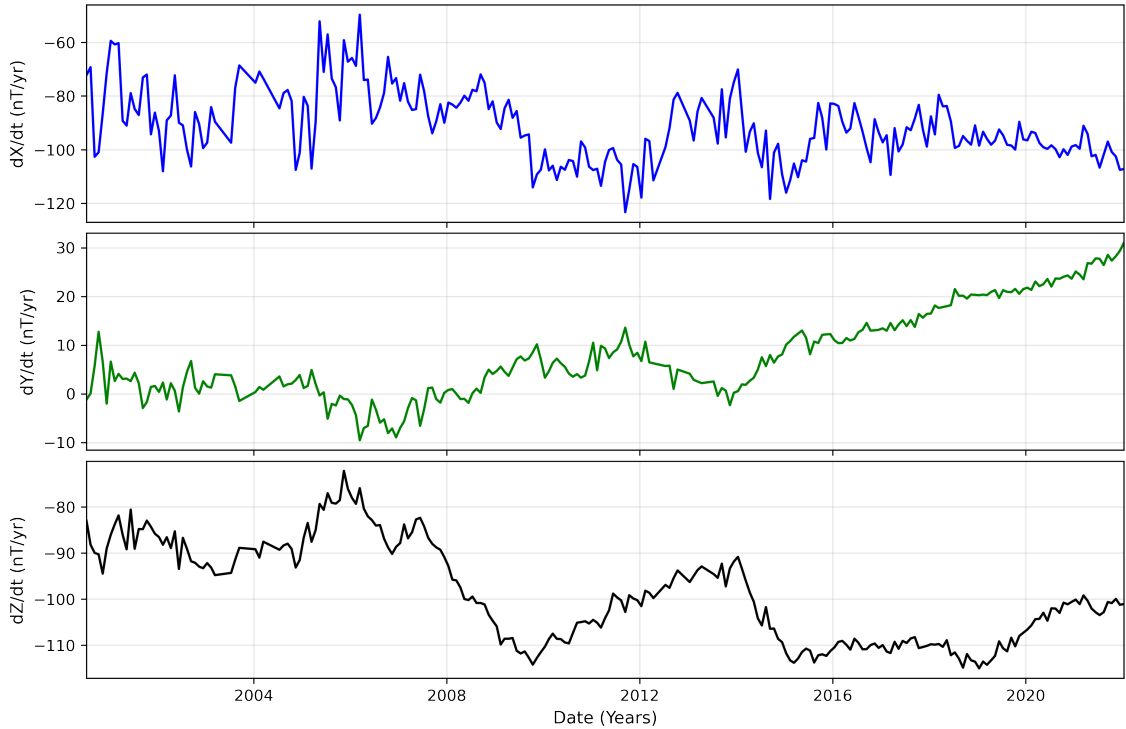
Figure 3.7: Vassouras magnetic observatory (VSS) SV calculated using MOSFiT package. Blue, green and black lines are the X, Y and Z SV, respectively.

3.6 Geomagnetic jerk detection

Geomagnetic jerks can be viewed as a sudden change in the first time derivative (SV) trend, characterized by a V-shape, or an abrupt change in the second time derivative (secular acceleration, SA) of the magnetic field. Usually, they are not observed simultaneously around the globe, i.e. the same event is observed in different times at different observatories (BLOXHAM *et al.*, 2002; COURTILLOT *et al.*, 1978; PINHEIRO *et al.*, 2019).

Many methods to detect geomagnetic jerks have been applied in the last decades. BROWN *et al.* (2013) presented an overview of geomagnetic jerk detection techniques, events and used data. Concerning the techniques, in the last decades the most applied were: fitting two straight lines (least-squares) (LE HUY *et al.*, 1998; LE MOUËL *et al.*, 1982; PINHEIRO *et al.*, 2011), wavelet analysis (ALEXANDRESCU *et al.*, 1996) and visual detection (CHULLIAT *et al.*, 2010; MANDEA *et al.*, 2000; OLSEN e MANDEA, 2007).

In addition to the visual detection, in MOSFiT I provide an automatic fitting of two straight lines segments by least squares, considering the user chosen time window. The package makes use of PWLF (JEKEL e VENTER, 2019), a python library that fits continuous piecewise linear functions to one dimensional (1D) dependent variables. In this case, I want to fit two piecewise linear segments for an unknown breakpoint (the straight lines intersection), which is the geomagnetic jerk occurrence time (t_0). The jerk amplitude (A) is the difference between the slopes of the two linear trends (LE MOUËL *et al.*, 1982; PINHEIRO *et al.*, 2011). The geomagnetic jerk detection is shown in Figure 3.8, the A is the result of A_2 - A_1, indicated in the figure, the t_0 is also demonstrated.
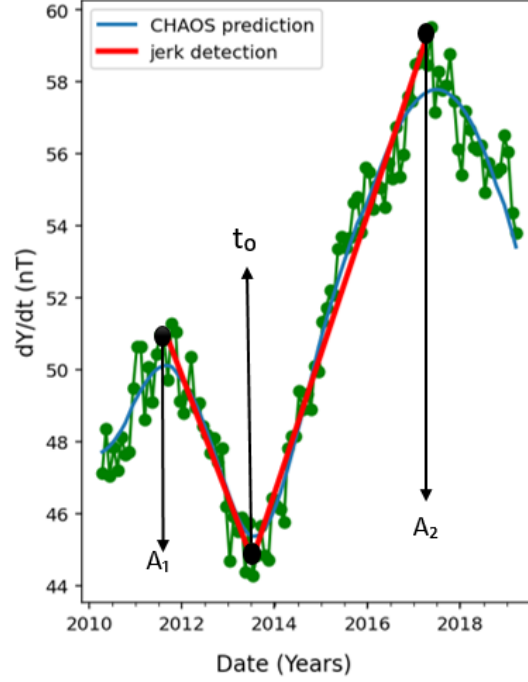
Figure 3.8: MOSFiT geomagnetic jerk detection parameters schema. The red lines are the straight line segments fitted to the Y SV (green lines) automatically by MOSFiT. The black dots and arrows are indicating the detection parameters A_1, A_2 and t_0. The A is the difference between the slopes A_2 and A_1.

Following JEKEL e VENTER (2019), lets assume a dataset with n data points (x_n, y_n). The set of piecewise linear functions are described as

$$
y(x) = \begin{cases}
n_1 + m_1(x - b_1) & b_1 < x \le b_2 \\
n_2 + m_2(x - b_2) & b_1 < x \le b_3 \\
\quad\vdots \\
n_{nb-1} + m_{mb-1}(x - b_{nb-1}) & b_{nb-1} < x \le b_{nb}
\end{cases}
\tag{3.7}
$$

Where b indicates the breakpoints, being b_1 the first and b_{nb} the last breakpoint. The number of linear segments is nb - 1. Enforcing that the functions are continuous over the domain, the slopes and intercepts of each piecewise is dependent of the previous values, reducing the eq.3.7 to

$$
y(x) = \begin{cases}
\beta_1 + \beta_2(x - b_1) & b_1 \le x \le b_2 \\
\beta_1 + \beta_2(x - b_1) + \beta_3(x - b_2) & b_2 < x \le b_3 \\
\quad\vdots \\
\beta_1 + \beta_2(x - b_1) + \beta_3(x - b_2) + ... + \beta_{nb}(x - b_{nb-1}) & b_{n-1} < x \le b_{nb}
\end{cases}
\tag{3.8}
$$

25

The number of unknown model parameters β is the same as the number of break-points, and the piecewise functions in the eq.3.8 can be expressed as the following matrix

$$
\begin{bmatrix}
1 & x_1 + b_1 & (x_1 + b_2)i_{x_1 > b_2} & \cdots & (x_1 + b_{nb-1})i_{x_1 > b_{nb-1}} \\
1 & x_2 + b_2 & (x_2 + b_2)i_{x_2 > b_2} & \cdots & (x_2 + b_{nb-1})i_{x_2 > b_{nb-1}} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & x_n + b_1 & (x_n + b_2)i_{x_n > b_2} & \cdots & (x_n + b_{nb-1})i_{x_n > b_{nb-1}}
\end{bmatrix}
\times
\begin{bmatrix}
\beta_1 \\
\beta_2 \\
\vdots \\
\beta_{nb}
\end{bmatrix}
=
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{bmatrix}
\quad (3.9)
$$

Where $i_{x_n > b_2}$ is the indicator function, which is as piecewise functions that are 0 or 1, as exemplified by eq.3.10

$$
i_{x_n > b_1} = \begin{cases} 0 & x_n \leq b_2 \\ 1 & x_n > b_2 \end{cases} \quad (3.10)
$$

The eq.3.9 is solved as a linear system, expressed by

$$
A\beta = y \quad (3.11)
$$

Where A is the n x n_b regression matrix, β is the vector of unknown parameters, and y is the vector of data points. The least squares problem solves for the unknown that reduces the sum-of-square of the residuals, being solves as

$$
\beta = (A^T A)^{-1} A^T y \quad (3.12)
$$

Given the β, the residual vector is obtained by

$$
e = A\beta - y \quad (3.13)
$$

Global optimization is used to find the best breakpoint (t_0), by minimizing the sum-of-square error of the residuals, it solves the least squares fit several times until find the best breakpoint location. The detection is applied to XYZ components, giving as output statistics: coefficient of determination (R^2), jerk amplitude (A) and occurrence time (t_0). R^2 compares the correlation between the fitted segments and the observed SV, following JEKEL e VENTER (2019) approach. First, the total sum-of-square (T) and the sum-of-square of the residuals (R) are calculated.

$$
T = \sum_{i}^{n} (y_i - \bar{y}) \quad (3.14)
$$

Where the $\bar{y}$ is the mean of y, depending only on the observed data.

$$R = e^T e \tag{3.15}$$

Where e denotes the residual vector (eq. 3.13), R^2 is the:

$$R^2 = 1 - \frac{R}{T} \tag{3.16}$$

Figure 3.9 demonstrates the geomagnetic jerk detection application for an event around 2014. The detection was performed in the SV of Niemegk Magnetic Observatory (NGK), computed using MOSFiT and corrected by CHAOS-7 model external field. The detection was applied to a time window from June 2012 to January 2018. The statistics about the detection are listed in the Table 3.1. As expected, the X SV presented the largest misfit since it is the most affected by the external field contribution. Otherwise, Y and Z SV present a very good fit. The occurrence time (t_0) was the same for X and Z components (2014.62), and a bit earlier in the Y component (2014.27). The user should take into account that the chosen window may strongly affect the jerk characterization and the interpretation of the results should be done by the user. The straight-line segments are adjusted in the time window independent of the existence of a jerk sign. The existence of a V-shape is fundamental as well as check the jerk amplitude and the R^2.
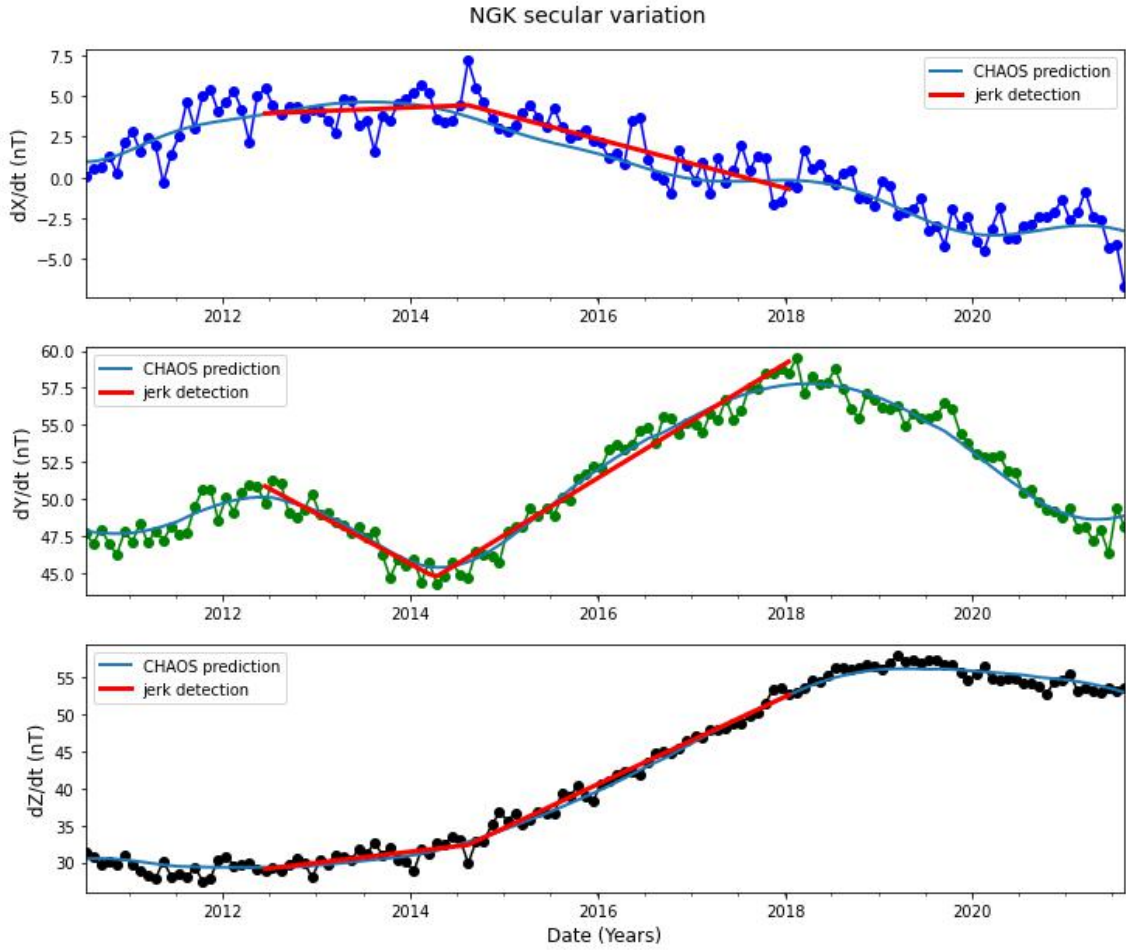
Figure 3.9: Niemegk magnetic observatory (NGK) SV from July 2010 to July 2021. Dark blue, green and black are the SV for X, Y and Z geomagnetic components, respectively. Light blue lines give the SV from the CHAOS-7 predicted core field. The red straight lines are the geomagnetic jerk automatic detection using MOSFiT.

Table 3.1: MOSFiT geomagnetic jerk detection statistics for X, Y and Z secular variation. The parameters R^2, A and t_0 are the coefficient of determination, amplitude and occurrence time, respectively.

SV	R^2 %	A (nT/yr)	t_0 (time)
X	0,73	-1,73	2014,62
Y	0,96	7,26	2014,27
Z	0,98	4,43	2014,62

Chapter 4

Results

In this chapter I present the results of the MOSFiT package developments, as well as give information about the package use and set up. The results are presented the same structure of the chapter 4, first I present case studies 1 and 2, followed by MOSFiT applications.

4.1 Magnetic Observatories and Stations Filtering Tool (MOSFiT) package

MOSFiT python package was designed to work with INTERMAGNET minute mean data in the IAGA-2002 format, in order to analyse the SV and check the INTERMAGNET observatories data quality. The definitive and quasi-definitive data are mainly used because of higher quality and reliability, especially for SV studies. However, the package can also be used with others types of IAGA-2002 data (i.e. provisional and variation). In this chapter, I give instructions about the package use and details about the data processing methods. In the Appendix A I present the package documentation with more details about the functionalities.

4.1.1 Requirements and package installation

MOSFiT is developed in Python 3 language. The package can be downloaded in: `https://github.com/marcosv9/MOSFiT-package`. In the same link and in the Appendix A, there is a documentation of how to use the package functions and usage examples. MOSFiT has different modules that must be imported before the use [1].

---

[1] Here are importing suggestion to use the package: "import data_processing_tools as dpt","import utilities_tools as utt", "import main_functions as mvs" and "import support_functions as spf".

4.1.2 Downloading IAGA-2002 data

In order to use MOSFiT, the data that will be investigated must be stored in the local computer. This data can be downloaded on the INTERMAGNET website (`https://www.intermagnet.org/`), directly in the INTERMAGNET ftp server (ftp://ftp.seismo.nrcan.gc.ca/intermagnet/) or by using the MOSFiT function called "download_data_intermagnet". MOSFiT will only read filenames in the same format of INTERMAGNET IAGA-2002 [2]. After the data is downloaded, the user may organize all files from different observatories in a single or different folders.

4.1.3 MOSFiT modules and functionalities

MOSFiT package is organized in four different modules: main_functions, data_processing_tools, support_functions and utilities_tools. Their functions are designed to automatically work with all INTERMAGNET observatories. The CHAOS-7 model (FINLAY *et al.*, 2020) is integrated into the package, since MOSFiT provides options to predict different sources of the geomagnetic field and correct the magnestospheric field using this model. Here I will describe the most important developed functions and their functionalities.

Data management and visualization:

• load_intermagnet_files - Read e transform into a single dataset all the data for any IAGA-2002 data format.

• download_data_INTERMAGNET - Download automatically, QD or Definitive data for all the INTERMAGNET observatories or for a passed list, based on the selected period, organizing in the correct structure.

• plot_samples - Plot minute, hourly, daily, monthly and annual samples for X, Y and Z components.

• plot_sv - Plot the secular variation of the geomagnetic data.

• plot_tdep_map - Plot a global map of the secular variation or secular acceleration predicted using CHAOS-7 model for a specific data.

• IMO - Class representing IMO. Can be used to check IMO information on MOSFiT database (IMO existence, latitude, longitude, altitude) as well as add a new IMO or delete

---

[2]The structure is: IAGA code, year, month, day, type of data and time sample. For example: vss20090809dmin.min (IAGA code = vss, year=2009, month=08, day=09, "d" of definite data and "min" indicating minute means

Data processing options:

- resample_obs_data - Allows the user to resample the geomagnetic data to hourly, daily, monthly or annual means.

- calculate_sv - Compute the secular variation from minute means.

- hdz_to_xyz_conversion - Automatically identify the existence of HDZ in the data series, and convert to XYZ.

- remove_disturbed_days - Drop the top 5 disturbed days from the dataset.

- keep_quiet_days - Remove from the dataset all the days outside the top 10 quiet days.

- night_time_selection - Select the nighttime period based on the local time.

- kp_index_correction - Filter the observed data based on the Kp-index. The user must specify the Kp value limit.

- chaos_model_prediction - Predict the different geomagnetic sources using the CHAOS-7 model.

- chaos_model_external_field_correction - Uses the predicted external part of the geomagnetic field, and correct the observed data.

- jerk_detection - Fit two straight line segments in each geomagnetic components, for a specified time window (adapted from (JEKEL e VENTER, 2019), updated).

- MOSFIT_SV - The main function, presented in the sequence, includes many functions and automate the process.

4.1.4   Running MOSFiT_SV function

MOSFiT_SV is an interactive function of the package that merges different functions into a single workflow. Figure 4.1 shows a flowchart illustrating a possible sequence of the most important data processing options. However, the user can combine any of the processing steps in any possible order or combination. MOSFiT_SV starts with user inputs and finish with the geomagnetic jerk detection. During the process the user can choose about how the data will be processed. The diamond-shapes represent the interactions in the code, in which the user needs to choose the options to process the data, the red rectangles represent the automatic calculations and the green rectangles are the output (figures, files and statistics).

31

All daily files will be read and concatenated into a single dataset, organized by columns (that are the geomagnetic components X, Y and Z) indexed by date and time. The missing values, filled as 99999.9 in INTERMAGNET standard format, are automatically replaced by "Not a Number (NaN)", since they would not be appropriated for the SV analysis and statistics. The Hampel filter is the first interaction of the code, used for outlier detection. Following the flowchart, the second interaction is the external field filtering. There are the following options: (i) Kp index ≤ 2 (KP), (ii) selection of only nighttime period (NT), (iii) keep only quiet days (QD) or by (iv) removing disturbed days (DD). The third step (Figure 4.1) is the CHAOS-7 model external field filtering. After all chosen options in the workflow corrections, the SV is calculated. The data can be visualized by using different means (minute, hourly, daily, monthly, annual), while the SV is only visualized by monthly means. The user can choose if these plots and files should be saved or not. A specific directory is automatically created to store the saved files/plots. The outputs are separated by observatories, and all the data files contain a header with the observatory information and the chosen processing options. Even if the user decides not to save the plots, they will be displayed in the screen. In this way, the user can visually inspect the SV before deciding the suitable geomagnetic jerk window detection. The last interaction is the geomagnetic jerk detection, based on the fitting of two straight-linear segments. The user just needs to inform the start and end time window for the jerk detection. The outputs are plots with the information of geomagnetic jerk occurrence time (t_0) and geomagnetic jerk amplitude (A) and coefficient of determination (R^2).
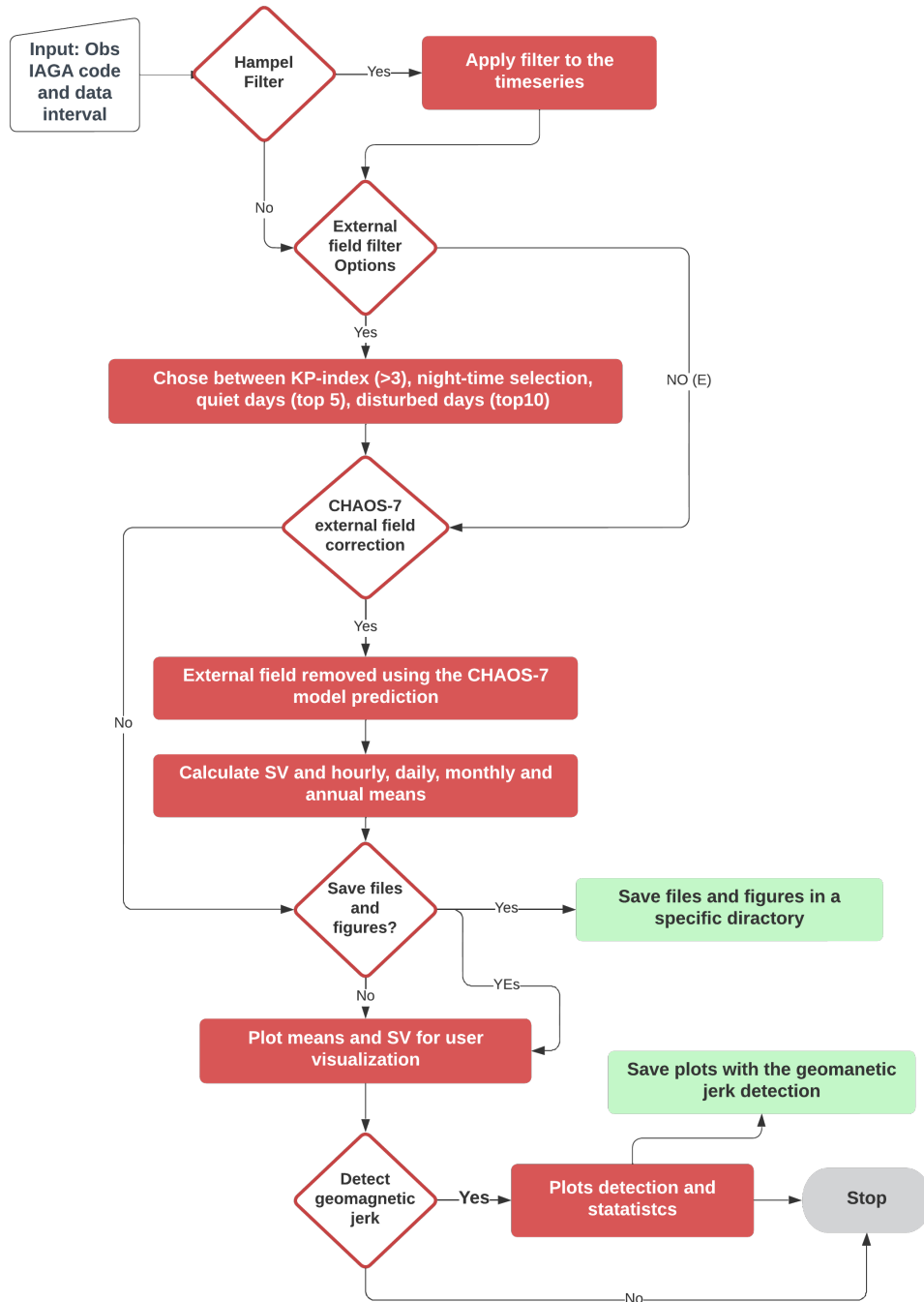
Figure 4.1: Flowchart of MOSFiT, as implemented in the interactive function sv_-obs. The diamond shapes are the data processing sequence interactions, where the user chooses the options of data processing. Diamonds, red boxes, and green box represents user selections, processing steps, and outputs, respectively.

4.2 Case Study 1: Global misfit with and without magnetospheric correction

In case study 1, I present a "Global misfit with and without magnetospheric correction" of the SV for 115 INTERMAGNET observatories. This case study is divided in two parts. In the first part, I focus on validating the automatic prediction of the geomagnetic field sources and the magnetospheric correction using the CHAOS-7 model. The results obtained using MOSFiT were compared with the one's from FINLAY *et al.* (2020). In the second part, I present a study of the magnetospheric correction effectiveness, by producing global misfit of the observed versus predicted SV, before and after the magnetospheric correction.

4.2.1 Case study 1: part 1

To validate the latest version of the CHAOS model, FINLAY *et al.* (2020) used ground observatory SV data from 181 observatories spanning the time period from 1997.5 to 2019.5. The monthly means were derived from hourly means that were corrected for magnetospheric sources. The misfit between the model's internal field predictions and the observatory SV was calculated by taking the root mean square error (RMSE) between them.

In the part 1, I reproduced FINLAY *et al.* (2020) validation, with the objective of verifing MOSFiT automatic methods. Comparing my results with FINLAY *et al.* (2020), I can check the resample observatory data , calculate SV, automatic prediction of CHAOS-7 model geomagnetic sources and magnetospheric field correction methods.

Data and observatories selection

I select 115 INTERMAGNET observatories and data from January, 2000 to June, 2022 (referred to as the observed data). Quasi-definitive data was used when definitive data was unavailable, it means that for 2021 and 2022 most of the data is quasi-definitive. Observatories with less than 3 years of data availability throughout the entire time interval were excluded.

Using MOSFiT, I predicted the internal field (core + crustal) of the CHAOS-7 model for the same observatories and time interval (referred to as the predicted data).

The 115 selected INTERMAGNET observatories are depicted on a world map in Figure 4.2.
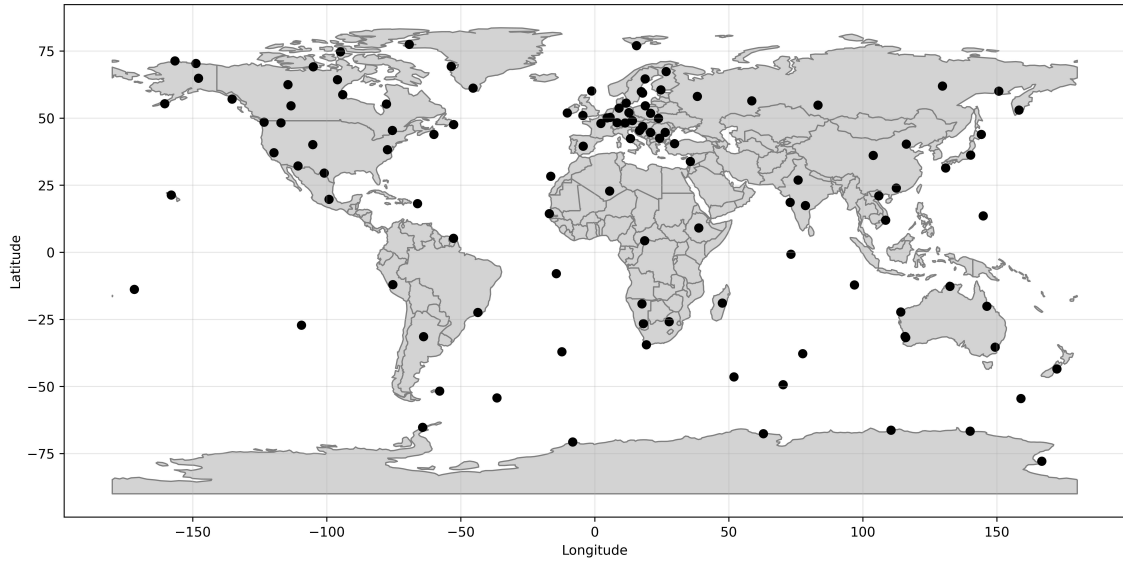
Figure 4.2: Worldmap displaying the 115 INTERMAGNET observatories selected to compare the MOSFiT methods with FINLAY *et al.* (2020).

CHAOS-7 model prediction

Using MOSFiT, I predicted the CHAOS-7 model geomagnetic field for the 115 IN-TERMAGNET observatories. The core field source was considered up to the 20th harmonic spherical degree, the crustal field up to the 110th degree, and the magnetospheric sources (GSM + SM) up to the 2nd degree. The internal sources were used to calculate the SV, while the magnetospheric field was used to correct the observed data.

Secular Variation

To calculate the observed SV, I first correct the observed data from the selected observatories using the MOSFiT magnetospheric field correction method (described in section 3.3). This yields the observed internal field. I then use the MOSFiT function to calculate the SV from the selected observatories, following eq.3.5, which results in a spanning time interval from July, 2000 to October 2021. This involves resampling the corrected observed data to monthly means. The predicted SV is also calculated using MOSFiT.

Misfit computation

To obtain the misfit, I calculate the RMSe between the observed corrected SV and the predicted SV, for X, Y and Z geomagnetic components, following eq. 4.1. The RMSe is defined as the square root of the average of the squares of the differences

35

between the predicted SV ($\hat{y}$) and the observed corrected SV (y) over the total number of observations (n). The misfit here described (observed corrected vs predicted) is referred to as $RMSe_1$. After calculate the $RMSe_1$ for all observatories, I calculated their mean, which will be compared with results of FINLAY $et\ al.$ (2020).

$$RMSe = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}} \tag{4.1}$$

Figure 4.3 shows an example of the calculation of the $RMSe_1$ for the Honolulu observatory (HON). The black dots correspond to the observed corrected SV, and the red lines correspond to the predicted SV. The $RMSe_1$ is labelled in the figure for X, Y and Z geomagnetic components.
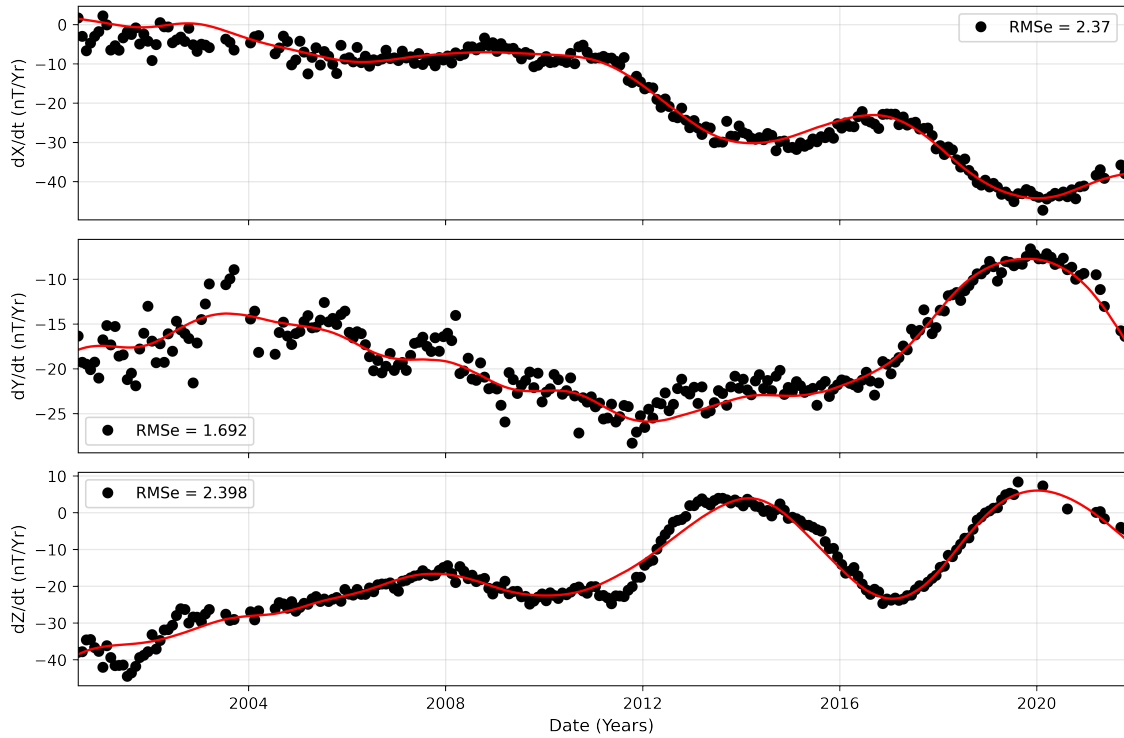
Figure 4.3: Computation of $RMSe_1$ for Honolulu magnetic observatory (HON). Black dots are the observed corrected SV, while red lines are the predicted SV (core field).

Case study 1 (part 1): discussion

The main goal of case study 1 is to validate part of MOSFiT package implementation, i.e., geomagnetic observatory monthly means and SV calculation, CHAOS-7 model geomagnetic field prediction and magnetospheric field correction. In search of that, was produced the misfit (later in the text, this misfit is reffered as $RMSe_1$)

between 115 INTERMAGNET geomagnetic observatories SV and the CHAOS-7 model predicted internal field SV. The investigation is based on FINLAY *et al.* (2020) study, where was introduced the first version of the model that is accessed by MOSFiT.

The MOSFiT automatic magnetospheric field correction validation against the results from FINLAY *et al.* (2020) are presented in Table 4.1. The results produced using MOSFit (also in Table 4.1) are very similar though not identical as there are some differences in the observatories, time interval and CHAOS model version. These differences exist because FINLAY *et al.* (2020) selected not only IN-TERMAGNET observatories, while MOSFiT was designed to automatically work only with INTERMAGNET observatories, although it is possible to insert different observatories or stations. Also the data availability in the epoch is not the same as the one in the present. Even taking into account the differences, the results indicates that the CHAOS implementation in MOSFiT works properly. In Figure 4.4 is showed the global distribution of this RMS misfit for the individual observatories. As expected, largest misfits are observed at high latitudes, where there are the strongest unmodelled external field variations. The smallest misfits are observed in Europe, likely because here the density of geomagnetic observatories is the highest.

Table 4.1: Validation of the MOSFiT implementation for external field correction by CHAOS-7, against results from FINLAY *et al.* (2020)

	MOSFit ($RMSe_1$)	FINLAY *et al.* (2020) ($RMSe_1$)
dB_r/dt	3,73	3,73
dB_θ/dt	2,90	3,59
dB_ϕ/dt	3,50	3,31

4.2.2 Case study 1: part 2

In addition to the validation, I analyzed the effectiveness of the magnetospheric field correction method by producing the misfit between observed SV (no correction) and predicted SV (referred to as $RMSe_2$) for each geomagnetic component (X, Y and Z). I divided the observatories into three latitude zones: low (between 0 and 30 degrees to North and South), mid (between 31 and 59 degrees, North and South), and high (from 60 to 90 degrees, North and South). The external field has varying impacts on the Earth's magnetic field at different latitudes, and I aim to quantify the extent to which the magnetospheric field can affect ground geomagnetic measurements based on their location. The magnetospheric correction effectiveness was evaluated by the percentage of change between the mean $RMSe_2$ and $RMSe_1$, for each latitude zone.
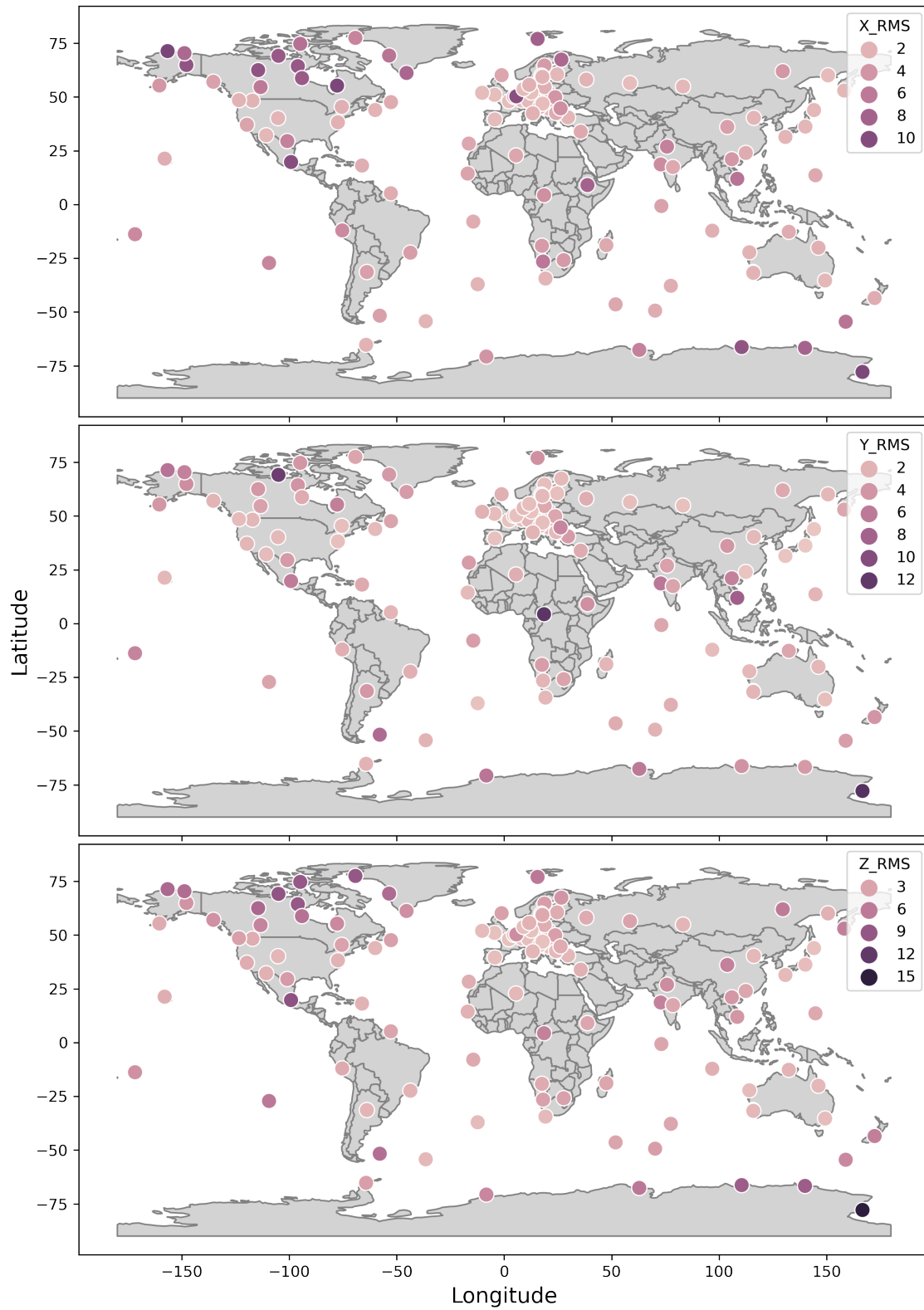
Figure 4.4: INTERMAGNET observatories RMSe between CHAOS-7 model internal field prediction SV and observed (filtered) SV. Maps from X, Y and Z components SV are showed from the top to the bottom, respectively. These results were produced by using MOSFiT package.

Case study 1 (part 2): discussion

In the second part of Case Study 1, I aim to explore the effectiveness of magneto-spheric field filtering in improving the secular variation (SV) at each geomagnetic component across 115 INTERMAGNET observatories. Since external field sources have uneven global impacts, I categorized the observatories according to their latitude zones. This investigation seeks to highlight the significance of data filtering in studying core field phenomena. The analysis involves computing the difference between $RMSe_1$ (as previously introduced) and $RMSe_2$ (SV without filtering vs predicted core field).

Figure 4.5 presents an example of the observed improvement in secular variation (SV) resulting from filtering the magnetospheric field. This is demonstrated through the display of the RMSe in two cases at the Honolulu observatory (HON): $RMSe_1$ and $RMSe_2$. The left panel shows the results of $RMSe_2$, which compares the observed SV to the predicted SV for the geomagnetic components dX, dY, and dZ. The right panel displays $RMSe_1$, which is the same data and plots as depicted in Figure 4.3. I note that the improvement in SV is evident across all geomagnetic components. Notably, the X component exhibits the largest improvement at approximately 352%, as expected due to its sensitivity to the external field. Even the Y component, which is less impacted by the external field, showed a significant improvement of 55%.
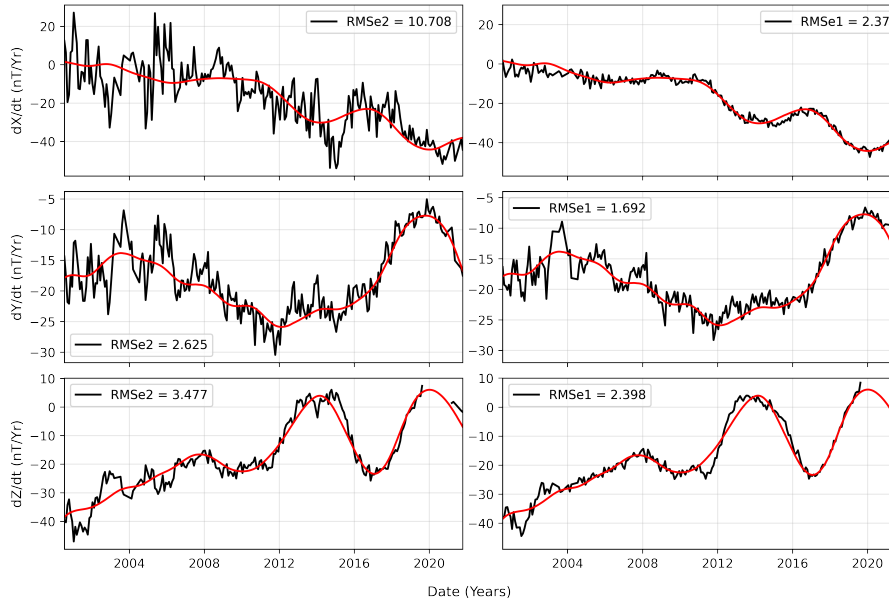
Figure 4.5:   Comparison between $RMSe_2$ and $RMSe_1$ for HON. In the left panel are the $RMSe_2$ computed between observed and predicted (core field) SV. In the right panel are the $RMSe_1$ computed from observed corrected and predicted SV. The black lines are always representing the SV from observed data (filtered in the right panel) and the red lines are the predicted SV (both panels).

Table 4.2 shows the percentage of improvement caused by the correction method. Most improvement is seen for dX/dt at mid and low latitudes as the magnetospheric signal, which is modelled by CHAOS-7 and then subtracted, is strongest in the X component. At high latitudes the improvement is only around 30% since signals from the auroral electrojet and field aligned currents are not modelled by CHAOS-7. The Y component is the least affected by external fields and shows the lowest RMS values as well as the smallest improvements. Despite the smallest improvements for dY/dt, it is important to check each observatory individually before decide or not to correct the data. Individual observatories may present a higher improvement, as for example the mid latitude observatory HON, that present more than 50% of improvement for dY/dt (see Figure 4.3).

Table 4.2: CHAOS-7 model external field filtering effectiveness. Percentage of change between $RMSe_2$ and $RMSe^1$ (see text) for low, mid and high latitudes, for the SV of the X, Y and Z component.

	Mid Latitude			Low Latitude			High Latitude		
	$RMSe_2$	$RMSe_1$	%	$RMSe_2$	$RMSe_1$	%	$RMSe_2$	$RMSe_2$	%
dX/dt	7,99	2,62	204,96%	11,46	3,90	193,85%	8,44	6,43	31,26%
dY/dt	2,68	2,10	27,62%	3,73	3,42	9,06%	4,84	4,36	11,01%
dZ/dt	5,36	2,5	114,40%	4,04	3,27	23,55%	9,65	6,45	49,61%

4.3 Case Study 2: Detection of the geomagnetic jerks in 2007, 2011 and 2014

TORTA *et al.* (2015) detected geomagnetic jerk events in 2007, 2011, and 2014 using observatory data from NGK, EBR, TAM, and ASC, by fitting two straight line segments to a selected time window. In their publication, the detection was carried out only for the Y component, as it is less affected by external field contributions and no external field correction was applied. The geomagnetic jerk detection parameters occurrence time (t_0) and amplitude (A) will be take into account in the comparison.

To achieve the purpose of validate MOSFiT automatic jerk detection method, I detected the 2007, 2011 and 2014 geomagnetic jerk events for the same geomagnetic observatories.

4.3.1 Secular Variation calculation

Following TORTA *et al.* (2015), the SV was calculated from monthly means, only for the Y geomagnetic component. No external field filtering was performed in the data.

4.3.2 Geomagnetic jerk detection and validation

Using the MOSFiT geomagnetic jerk detection function, I detected geomagnetic jerk events and automatically obtained the values of t_0, A, and R^2. In order to reproduce the detection carried out by TORTA *et al.* (2015), I adopted the following criteria for selecting the time window:

- Search for the characteristic V-shape.

- Repeat the detection slightly shifting the window until the best misfit (R^2) is found.

- For the 2014 event, the detection time window was limited to 2015 to simulate the data availability at that time.

I conducted the validation by comparing TORTA *et al.* (2015) t_0 and A with MOSFiT detections. Additionally, the R^2 and the detection window were provided.

4.3.3 Case Study 2: discussion

The case study 2 presents an analysis of the automatic detections of geomagnetic jerks using the MOSFiT method. The aim of this analysis was to validate the automatic geomagnetic jerk detection method by comparing it to the jerk detection parameters (t_0 and A) results of a previous investigation conducted by TORTA *et al.* (2015) of the 2007, 2011, and 2014 geomagnetic jerk events.

Although TORTA *et al.* (2015) did not include the detections figure in their investigation, I chose to demonstrate them in Figure 4.6 in order to facilitate the visual characterization of the event and to validate the results of the MOSFiT package. Figure 4.6 presents the Y SV (blue dots), CHAOS-7 model core field SV prediction (green curves), and jerk detection (orange lines), all of which were automatically calculated using MOSFiT. The figure displays the geomagnetic jerk detections from top to bottom for the years 2007, 2011, and 2014, and from left to right for the observatories NGK, EBR, TAM, and ASC. The characteristic V-shape is present in all detections, and there is a good agreement between the observed SV and the CHAOS-7 model predicted SV, as well as with the straight line segments.

The geomagnetic jerk detection parameters (t_0, A, R^2), and the chosen time window are presented in Table 4.3, for each geomagnetic jerk event and observatory. I compared my results of t_0 and A against results from TORTA *et al.* (2015) (inside parenthesis in Table 4.3).

The largest occurrence time difference between MOSFiT and TORTA *et al.* (2015) for the jerk in 2007 is 0.15 years or 55 days for NGK, while the amplitudes are very similar. For the jerk in 2011 occurrence time reported by TORTA

et al. (2015) for ASC is 2012.0, against 2010.02 detected by MOSFiT. However, Figure 4.6 shows that both the observatory data and the CHAOS model prediction identify a jerk around 2010, which is the same date as detected by MOSFiT. Jerk amplitudes detected by MOSFiT are very similar to those published by TORTA *et al.* (2015). Finally, for the jerk 2014 the results shows a very good agreement for the occurrence time of the jerk as well as a good agreement for the amplitude at all four observatories.
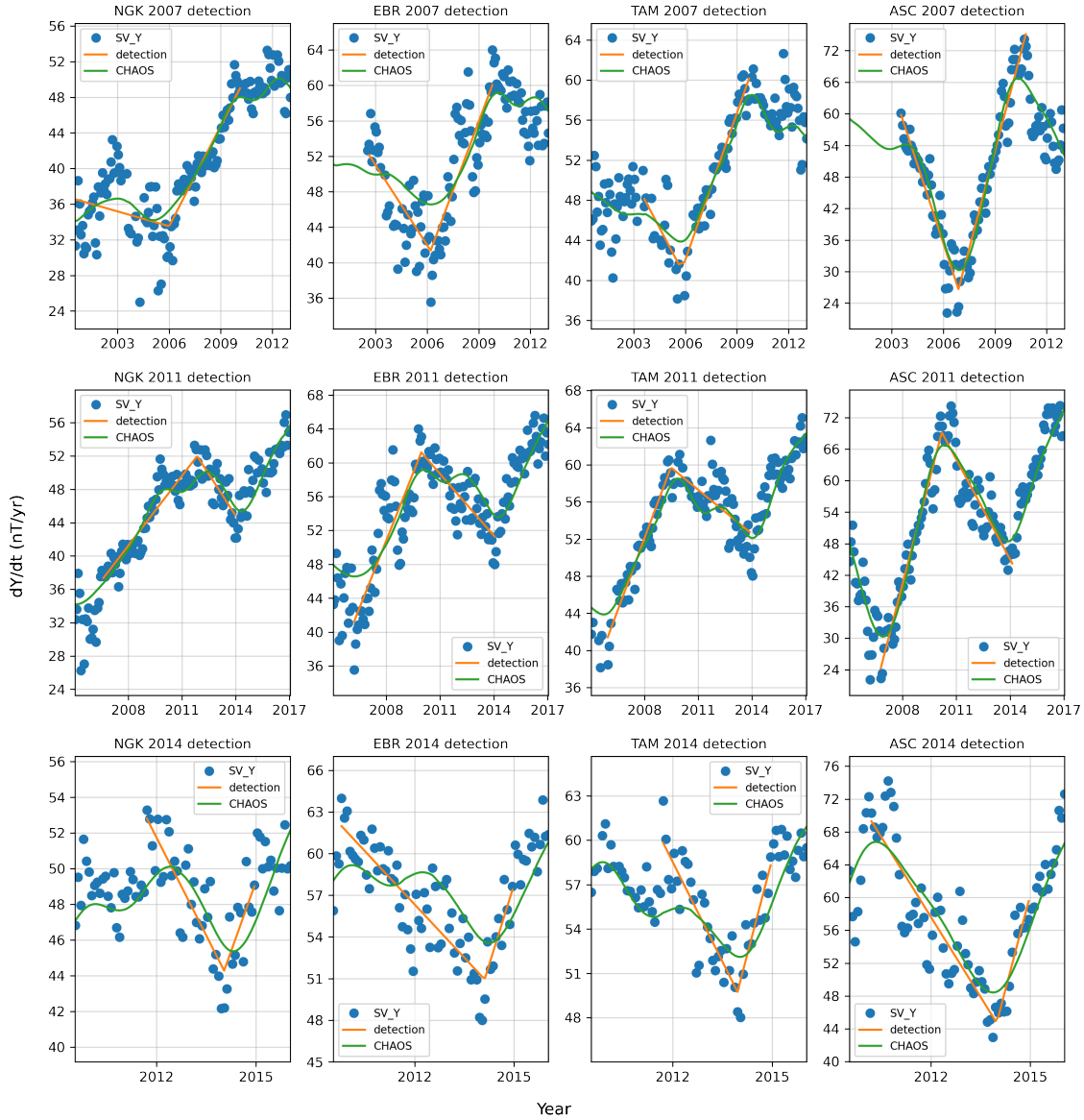
Figure 4.6: Geomagnetic jerk detection of the 2007, 2011 and 2014 events, using the MOSFiT automatic method, for NGK, EBR, TAM and ASC. Blue dots are the non-filtered Y component SV, green lines are the CHAOS-7 model internal field SV prediction and the orange straight lines are the MOSFiT automatic jerk detection.

Table 4.3: MOSFiT detection of geomagnetic jerks in 2007, 2011 and 2014 at NGK, EBR, TAM and ASC observatories. Occurrence time (t_0) and amplitude (A) are shown for MOSFiT and for TORTA *et al.* (2015) in parentheses. The coefficient of determination R^2 and the start and end year/month of the detection window are also shown, for this study.

IMO	t_0 (yr)	A (nT/yr)	R^2	initial window	final window
Jerk 2007					
NGK	2005.85 (2006.0)	4,75 (4,6)	0,83	2003-05	2010-02
EBR	2006.28 (2006.4)	6,16 (5,6)	0,67	2003-05	2009-10
TAM	2005.79 (2005.7)	8,15 (8,2)	0,90	2003-08	2009-10
ASC	2006.9 (2006.9)	22,72 (23,4)	0,91	2003-07	2010-01
Jerk 2011					
NGK	2011.84 (2011.8)	-6,08 (-6,2)	0,84	2006-08	2014-04
EBR	2009.95 (2011.0)	-7,76 (-6,7)	0,74	2006-02	2014-01
TAM	2009.48 (2009.5)	-6,78 (-6,8)	0,86	2005-10	2013-11
ASC	2010.02 (2012.0)	-19,35 (-19,5)	0,90	2006-12	2014-02
Jerk 2014					
NGK	2014.04 (2014.0)	8.82 (7.2)	0.68	2011-09	2014-12
EBR	2014.02 (2014.0)	15.28 (12.7)	0.80	2011-06	2014-12
TAM	2013.96 (2014.0)	13.03 (15.2)	0.74	2011-09	2014-12
ASC	2013.99 (2014.01)	21.87 (24.9)	0.79	2010-02	2014-12

4.4 MOSFiT application: determining the influence of the external field filtering methods on geomagnetic jerk detection

In this study, I evaluate the influence of different processing methods, i.e. no data selection and no CHAOS correction (original data), Kp data selection (later labelled KP), IQDs data selection (labelled QD) and nighttime data selection (labelled NT) and the CHAOS-7 magnetospheric field reduction (labelled CHAOS) in the detection of 2007, 2011 and 2014 geomagnetic jerk events in the X, Y and Z components. Analysing how different processing methods can change t_0 and A parameters statistics, or making it possible the detection of jerks previously not detectable.

4.4.1 INTERMAGNET observatories selection

As geomagnetic jerk can be a global or local event, we select 10 INTERMAGNET observatories longitudinal and latitudinal distributed, which allows a better comprehension. In Figure 4.7 are the 10 selected observatories location and represented by their IAGA code.
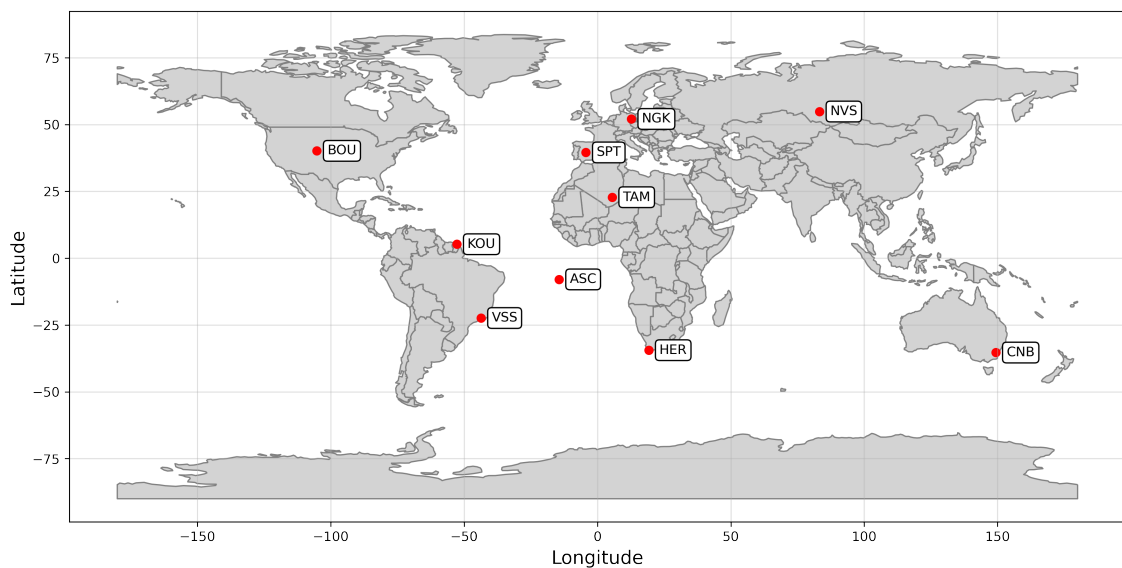

Figure 4.7: INTERMAGNET observatories used in the application for determining the influence of the external field filtering methods on geomagnetic jerk detection. The three letter IAGA code indicates the observatory.

4.4.2 Data treatment and secular variation

In order to compute the influence of different external field reduction methods, the measured observatory data was processed by what I labelled as: KP, QD, CHAOS and NT. The following is a summary of how each method operates. All data processing steps were performed using the MOSFiT package, from the data reading to geomagnetic jerk detection.

- The KP method removes all days with Kp index values greater than 2 from the dataset.

- The QD method only retains the top 10 quietest days of each month in the dataset.

- The CHAOS method predicts the external part of the geomagnetic field (GSM + SM up to degree 2) for X, Y, and Z, and subtracts it from the measured data.

- The NT method only selects data from 10pm to 5am.

As a result, five datasets were produced for each observatory (including the original dataset). I subsequently calculated the SV using Annual Differences of Monthly Means (ADMM) for each dataset.

4.4.3 Geomagnetic jerk detection

The geomagnetic jerk detection of 2007, 2010 and 2014 was performed following some criteria, first of all, only jerks with high values of R^2 are considered and observatories that have the typical "V" shape in their SV. The detection window was individually chosen according to the range of the SV trend for each of the observatories, but always keeping the same for the different processing methods. For each processing method the occurrence time (t_0) and the amplitude (A) was determined.

4.4.4 Influence analyses

After detecting the geomagnetic jerks for each observatory, I obtained the parameters t_0 and A. For t_0, that is represented in fraction of year, was first converted to days and the absolute different of each processing method and original data was calculated. The parameter A, represented in nT, was evaluated by computing the absolute percentage of change. To provide a comprehensive analysis of the effect on the detection parameters, I calculated the average of the results of the three jerk events for each geomagnetic component, which minimize contamination from outliers. In order to ensure the accuracy of the global average, only geomagnetic jerks that could be detected by all processing methods were considered.

4.4.5   MOSFiT application: discussion

Since the 2000s, several jerks have been detected by various authors using ground-based and satellite data and by geomagnetic global models. Reducing the influence of external field contributions on geomagnetic data is a challenge when studying the internal field. One common approach to mitigate the effects of magnetospheric currents, such as ring and auroral currents, which primarily affect the X component, is to compute yearly differences of observatory monthly means. On the other hand, the Z component is more affected by induction effects (TORTA *et al.*, 2015), and the Y component is the least affected, making it the preferred choice for most jerk detection studies. Another option is to select data with low geomagnetic activity before calculating the monthly means.

The 2003 jerk was first reported by OLSEN e MANDEA (2007), who used CHAMP satellite data from all geomagnetic activity conditions. The 2007 jerk was detected by CHULLIAT *et al.* (2010) using observatory data from all geomagnetic activity conditions and by KOTZÉ (2011) using quiet-time monthly means, with a Kp index of less than 4. CHULLIAT e MAUS (2014) identified a jerk around 2011 using observatory monthly means from nightside and quiet-time (Kp < 2 and Dst < 20), while TORTA *et al.* (2015) detected the 2014 jerk using data from all geomagnetic activity and the CHAOS-5 model. KOTZÉ (2017) also detected the 2014 jerk event in southern African magnetic observatories using quiet-time (Kp ≤ 2 and Dst variation < 3nT/h) monthly means and the CHAOS-6 model.

Therefore, the external field contribution present in geomagnetic observatory data have been filtered by many different methods when searching for a jerk detection. In addition, the analysis take into account the different geomagnetic components, which are also affected by external contributions in different ways. Here I quantify how each method affects jerk detection parameters, which means how their amplitude and occurrence time can change..

The results of my analysis are showed in Figure 4.8. The upper panels show the mean of the absolute difference in jerk occurrence time (in days) of the processing methods KP, QD, NT and CHAOS compared to using the original data. The mean occurrence time (Δt_0) difference between CHAOS correction and original data is largest for the X component (349 days) where it is about 7 times larger than for the processing methods KP, QD and NT (44, 53, 33 days, respectively). A similar pattern is seen for Z, where CHAOS gives 105 days, which is about 5 times larger than KP, QD and NT (23, 19, 22 days, respectively). From this comparison alone one can not define if CHAOS magnetospheric correction is either a lot better or a lot worse in these components than the tested data selection methods. However, given that the ring current has a strong amplitude combined with fast temporal

changes compared to a jerk and that its signal is not removed by night time or Kp-dependent data selection, one can expect that it is the magnetospheric correction that performs best. This can also be seen in Figure 4.9 that compares SV for X at Chambon-la-Foret magnetic observatory (CLF) for each processing scheme (blue) with the SV prediction by CHAOS. It clearly shows that the magnetospheric correction by CHAOS gives much better similarity to the CHAOS SV prediction than the three data selection processing methods KP, QD and NT, which all give SV time series that are quite similar to that of the original data with the full content of external fields.
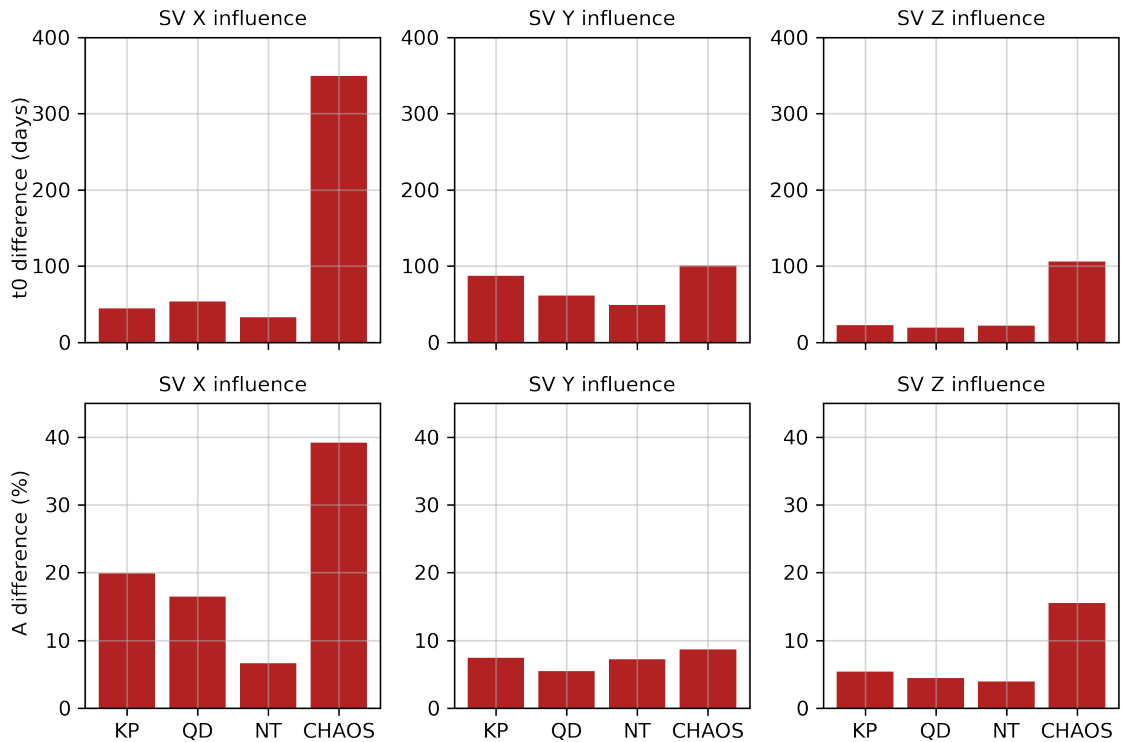
Figure 4.8: Upper panels: mean absolute difference in jerk occurrence time Δt_0 in days between processed data and original observatory data for different processing methods KP, QD, NT and CHAOS (for explanation see text) for X, Y and Z . Bottom panels: the same as upper panels but for jerk amplitude difference (ΔA) in %

Figure 4.9 shows the application of the different processing methods (KP, QD, NT and CHAOS), including original data, into the Chambon-la-Foret geomagnetic observatory (CLF) X SV calculation, compared to CHAOS-7 model core field. The effectiveness of the processing methods clearly illustrates the results obtained by the investigation.

Figure 4.9: Chambon-la-Foret geomagnetic observatory (CLF) X SV from 2000 to 2022, using the processing methods KP, QD, NT and CHAOS, and original data, compared to CHAOS-7 core field.

Chapter 5

Conclusion

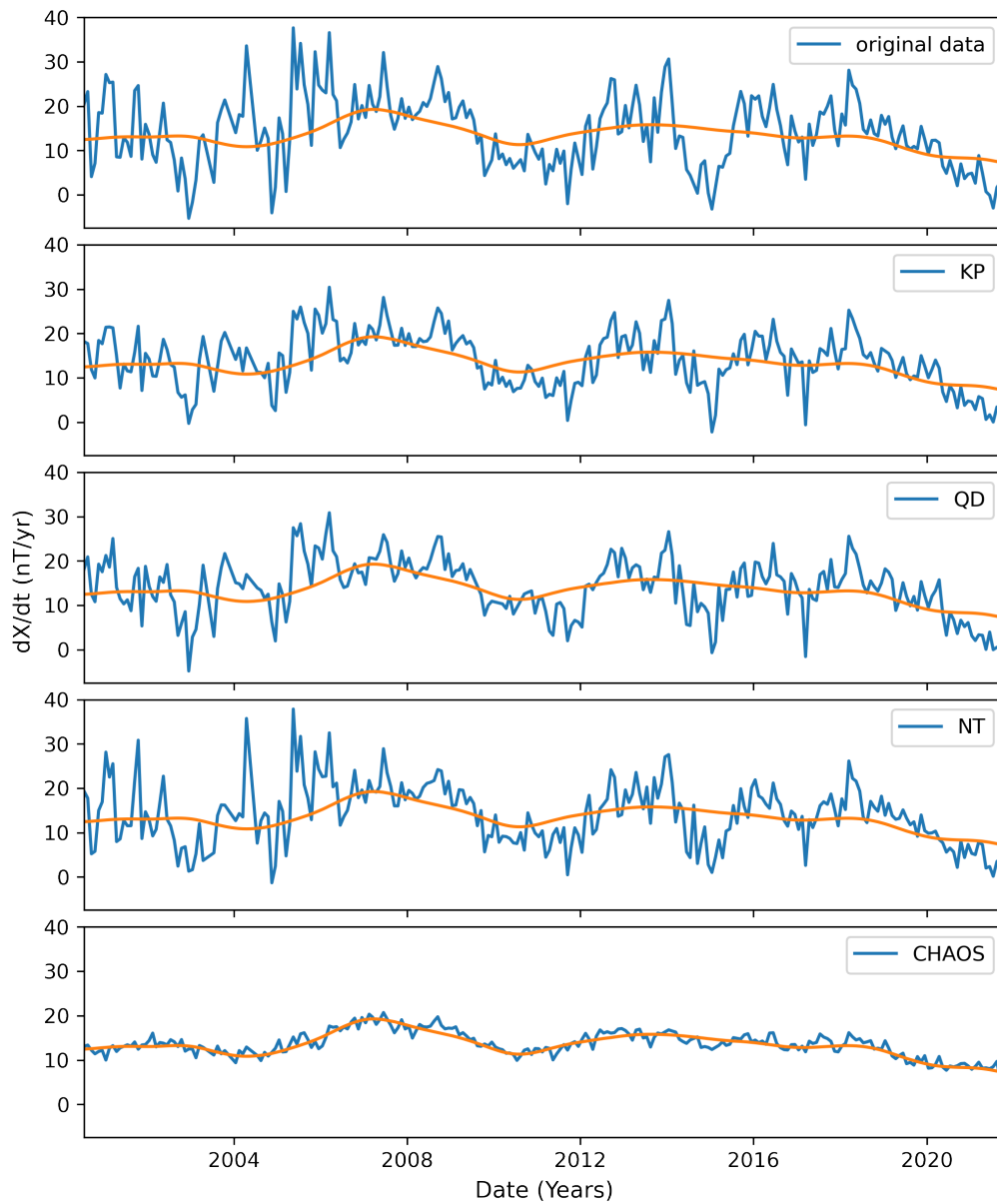I present the python package MOSFiT ("Magnetic Observatory and Stations Filtering Tool") to investigate the geomagnetic Secular Variation (SV) in observatory data. MOSFiT is designed to work with one minute INTERMAGNET definitive and quasi-definitive data. However, it can also be applied to any geomagnetic observatory and magnetometer stations data, as long as they are in the IAGA-2002 format. The package offers outlier rejection (Hampel filter), four data selection options (quiet days, disturbed days, Kp-index and nighttime period), an magnetospheric field reduction using CHAOS-7. It can resample the data to hourly, daily, monthly and annual means and provides a method to determine geomagnetic jerk occurrence time (t_0) and amplitude (A), and all steps can be visualized.

I successfully validate the implementation of CHAOS in MOSFiT against results for more than 150 observatories presented by FINLAY *et al.* (2020). This validation was also done by the MOSFiT geomagnetic jerk detection method against results for three jerks (2007, 2011, 2014) in four observatories presented by TORTA *et al.* (2015). I investigate the difference between the observatory data SV and the CHAOS core field SV (by calculating its RMS) for uncorrected and observatory data corrected by the CHAOS magnetospheric field prediction for 115 geomagnetic observatories in low, mid and high latitudes. For the uncorrected observatory data, this difference is smallest in the Y component, as expected. The magnetospheric field correction reduces this difference most strongly (by about two thirds) in the X component in low and mid latitudes. In general, the differences after the magnetospheric correction are similar for the X, Y and Z components and similar to the difference for the uncorrected Y component. I further quantified the effect of different observatory data processing methods on the determination of the jerk occurrence time (t_0) and the jerk amplitude (A) for the three jerks for a subset of 10 geomagnetic observatories and found that the CHAOS magnetospheric correction performs best, especially for the X and Z component.

MOSFiT is available at `https://github.com/marcosv9/MOSFiT-package` and

I expect it can be of a great help for geomagnetic observatory data analysis, allowing convenient and automatic access to the most recent geomagnetic observatory data and indices. In particular, it will allow a timely identification of future geomagnetic jerks and it can be used as a tool for geomagnetic observatory data quality control.

Bibliography

ALEXANDRESCU, M., GIBERT, D., HULOT, G., et al., 1996, "Worldwide wavelet analysis of geomagnetic jerks", *Journal of Geophysical Research: Solid Earth*, v. 101, n. B10, pp. 21975–21994.

AUBERT, J., FINLAY, C. C., 2019, "Geomagnetic jerks and rapid hydromagnetic waves focusing at Earth's core surface", *Nature Geoscience*, v. 12, n. 5, pp. 393–398.

BARTELS, J., 1938, "1938. Potsdamer erdmagnetische Kennziffern, 1 Mitteilung", *Zeitschrift fur Geophysik*, p. 68.

BARTELS, J., 1957, "The geomagnetic measures for the time-variations of solar corpuscular radiation, described for use in correlation studies in other geophysical fields", *Ann. Intern. Geophys.*, v. 4, pp. 227–236.

BLOXHAM, J., GUBBINS, D., JACKSON, A., 1989, "Geomagnetic secular variation", *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, v. 329, n. 1606, pp. 415–502.

BLOXHAM, J., ZATMAN, S., DUMBERRY, M., 2002, "The origin of geomagnetic jerks", *Nature*, v. 420, n. 6911, pp. 65–68.

BROWN, W., MOUND, J., LIVERMORE, P., 2013, "Jerks abound: An analysis of geomagnetic observatory data from 1957 to 2008", *Physics of the Earth and Planetary Interiors*, v. 223, pp. 62–76.

CAMPBELL, W. H., 2003, *Introduction to geomagnetic fields*. Cambridge University Press.

CHULLIAT, A., MAUS, S., 2014, "Geomagnetic secular acceleration, jerks, and a localized standing wave at the core surface from 2000 to 2010", *Journal of Geophysical Research: Solid Earth*, v. 119, n. 3, pp. 1531–1543.

CHULLIAT, A., MATZKA, J., MASSON, A., et al., 2017, "Key ground-based and space-based assets to disentangle magnetic field sources in the Earth's environment", *Space Science Reviews*, v. 206, n. 1-4, pp. 123–156.

CHULLIAT, A., THÉBAULT, E., HULOT, G., 2010, "Core field acceleration pulse as a common cause of the 2003 and 2007 geomagnetic jerks", *Geophysical Research Letters*, v. 37, n. 7.

COURTILLOT, V., LE MOUËL, J., OTHERS, 1978, "Sur une accélération récente de la variation séculaire du champ magnétique terrestre", .

COX, G., BROWN, W., BILLINGHAM, L., et al., 2018, "MagPySV: A Python package for processing and denoising geomagnetic observatory data", *Geochemistry, Geophysics, Geosystems*, v. 19, n. 9, pp. 3347–3363.

FENG, Y., HOLME, R., COX, G. A., et al., 2018, "The geomagnetic jerk of 2003.5-characterisation with regional observatory secular variation data", *Physics of the Earth and Planetary Interiors*, v. 278, pp. 47–58.

FINLAY, C. C., KLOSS, C., OLSEN, N., et al., 2020, "The CHAOS-7 geomagnetic field model and observed changes in the South Atlantic Anomaly", *Earth, Planets and Space*, v. 72, n. 1, pp. 1–31.

HAMPEL, F. R., 1974, "The influence curve and its role in robust estimation", *Journal of the american statistical association*, v. 69, n. 346, pp. 383–393.

JEKEL, C. F., VENTER, G., 2019, *pwlf: A Python Library for Fitting 1D Continuous Piecewise Linear Functions*. Disponível em: <https://github.com/cjekel/piecewise_linear_fit_py>.

KAURISTIE, K., MORSCHHAUSER, A., OLSEN, N., et al., 2017, "On the usage of geomagnetic indices for data selection in internal field modelling", *Space Science Reviews*, v. 206, n. 1, pp. 61–90.

KONO, M., 2010, *Treatise on Geophysics, Volume 5: Geomagnetism*. Amsterdam, Elsevier.

KOTZÉ, P., 2011, "Signature of the 2007 geomagnetic jerk at the Hermanus Magnetic Observatory, South Africa", *South African Journal of Geology*, v. 114, n. 2, pp. 207–210.

KOTZÉ, P., 2017, "The 2014 geomagnetic jerk as observed by southern African magnetic observatories", *Earth, Planets and Space*, v. 69, n. 1, pp. 1–7.

LANGEL, R., ESTES, R., 1985, "Large-scale, near-field magnetic fields from external sources and the corresponding induced internal field", *Journal of Geophysical Research: Solid Earth*, v. 90, n. B3, pp. 2487–2494.

LAUNDAL, K. M., RICHMOND, A. D., 2017, "Magnetic coordinate systems", *Space Science Reviews*, v. 206, n. 1-4, pp. 27–59.

LE HUY, M., ALEXANDRESCU, M., HULOT, G., et al., 1998, "On the characteristics of successive geomagnetic jerks", *Earth, planets and space*, v. 50, n. 9, pp. 723–732.

LE MOUËL, J.-L., DUCRUIX, J., DUYEN, C. H., 1982, "The worldwide character of the 1969–1970 impulse of the secular acceleration rate", *Physics of the Earth and Planetary Interiors*, v. 28, n. 4, pp. 337–350.

LOEWE, C., PRÖLSS, G., 1997, "Classification and mean behavior of magnetic storms", *Journal of Geophysical Research: Space Physics*, v. 102, n. A7, pp. 14209–14213.

MANDEA, M., KORTE, M., 2010, *Geomagnetic observations and models*, v. 5. Springer.

MANDEA, M., BELLANGER, E., LE MOUEL, J.-L., 2000, "A geomagnetic jerk for the end of the 20th century?" *Earth and Planetary Science Letters*, v. 183 (12), pp. 369–373. doi: 10.1016/S0012-821X(00)00284-3.

MATZKA, J., STOLLE, C., YAMAZAKI, Y., et al., 2021, "The geomagnetic Kp index and derived indices of geomagnetic activity", *Space Weather*, v. 19, n. 5, pp. e2020SW002641.

MATZKA, J., CHULLIAT, A., MANDEA, M., et al., 2010, "Geomagnetic observations for main field studies: from ground to space", *Space science reviews*, v. 155, n. 1, pp. 29–64.

MAYAUD, P., 1980, "Derivation, Meaning and Use of Geomagnetic Indices, Volume 22", *Washington, DC: American Geophysical Union*.

NEWITT, L., BARTON, C., BITTERLY, J., 1996, *Guide for magnetic repeat station surveys*, v. 325. International Association of Geomagnetism and Aeronomy Boulder, CO.

OLSEN, N., MANDEA, M., 2007, "Investigation of a secular variation impulse using satellite data: The 2003 geomagnetic jerk", *Earth and Planetary Science Letters*, v. 255, n. 1-2, pp. 94–105.

OLSEN, N., LÜHR, H., FINLAY, C. C., et al., 2014, "The CHAOS-4 geomagnetic field model", *Geophysical Journal International*, v. 197, n. 2, pp. 815–827.

PEARSON, R. K., NEUVO, Y., ASTOLA, J., et al., 2016, "Generalized hampel filters", *EURASIP Journal on Advances in Signal Processing*, v. 2016, n. 1, pp. 1–18.

PELTIER, A., CHULLIAT, A., 2010, "On the feasibility of promptly producing quasi-definitive magnetic observatory data", *Earth, Planets and Space*, v. 62, n. 2, pp. e5–e8.

PINHEIRO, K., JACKSON, A., FINLAY, C., 2011, "Measurements and uncertainties of the occurrence time of the 1969, 1978, 1991, and 1999 geomagnetic jerks", *Geochemistry, Geophysics, Geosystems*, v. 12, n. 10.

PINHEIRO, K., AMIT, H., TERRA-NOVA, F., 2019, "Geomagnetic jerk features produced using synthetic core flow models", *Physics of the Earth and Planetary Interiors*, v. 291, pp. 35–53.

REDA, J., FOUASSIER, D., ISAC, A., et al., 2011, "Improvements in geomagnetic observatory data quality". In: *Geomagnetic observations and models*, Springer, pp. 127–148.

SABAKA, T. J., HULOT, G., OLSEN, N., 2010, "Mathematical properties relevant to geomagnetic field modeling". In: *Handbook of geomathematics*.

ST-LOUIS, B., COUNCIL, I. E., COMMITTEE, I. O., et al., 2020, "INTERMAG-NET Technical Reference Manual, Version 5.0. 0", .

TORTA, J. M., PAVÓN-CARRASCO, F. J., MARSAL, S., et al., 2015, "Evidence for a new geomagnetic jerk in 2014", *Geophysical Research Letters*, v. 42, n. 19, pp. 7933–7940.

YAMAZAKI, Y., MAUTE, A., 2017, "Sq and EEJ—A review on the daily variation of the geomagnetic field caused by ionospheric dynamo currents", *Space Science Reviews*, v. 206, n. 1, pp. 299–405.

Appendix A
MOSFiT package documentation

Magnetic Observatories and Stations Filtering Tool (MOSFiT)

MOSFiT is part of my Master's project in geophysics. A python tool developted to investigate the secular variation (SV) of the Earth's geomagnetic field, accelerate the geomagnetic jerk detection and check geomagnetic observatory data quality.

Actually works for every INTERMAGNET geomagnetic observatory (1-min IAGA-2002 data format).

The package is separetade in modules called data_processing_tools, utility_tools, support_functions and main_functions.

There are functions used to reduce the external field contribution (5 options), calculate the SV by different methods, resample data by different means, plot the data and detect geomagnetic jerks.

All functions have a docstring with the documentation, including inputs, outputs and how to use.

Documentation summary

Setting up MOSFiT package

The MOSFiT python package was built to work with INTERMAGNET minute mean data in the IAGA-2002 format, in order to analyse the SV and check INTERMAGNET Magnetic Observatory (IMO) data quality. The definitive and quasi-definitive data are mainly used because of higher quality and reliability, especially for SV studies. However, we can also apply MOSFiT to others types of IAGA-2002 data, i.e. provisional data).

Package installation

MOSFiT is developed in the Python 3 language. The package can be compiled in the command window or in a "jupyter notebook enviroment".

You can download MOSFiT in: https://github.com/marcosv9/Thesis-Marcos. In this same link, there is a documentation of how to use the package functions, with some examples.

Downloading IAGA-2002 data

In order to use MOSFiT, it is necessary to have the data stored in a local computer.

This data can be downloaded from the INTERMAGNET website (https://www.intermagnet.org/), directly from the INTERMAGNET ftp server (ftp://ftp.seismo.nrcan.gc.ca/intermagnet/) or by using the MOSFiT function called "download INTERMAGNET file" (by choosing datatype, year, month and the list of observatories).

MOSFiT will only read filenames in the same format of INTERMAG-NET IAGA-2002 2. After the data is downloaded, the user may organize all files from different observatories in a single or multiple folders.

Most MOSFIT functions require an input called 'station'. It is the 3 letter IAGA code of the INTERMAGNET observatory. In MOSFiT there is a database with all INTERMAGNET observatories registered (IAGA code, latitude, longitude and altitude, this informations are used in many data processing functions). If you want to use MOSFiT with an observatory or location that are not registered in the database, there is a MOSFiT class called IMO that includes any location into the database. See utilities_tools section for an explanation about how to include any location.

Package modules import suggestion

To use the developed funtions, is necessary to import them. I suggest to use the following statements to import the modules.

```
import data_processing_tools as dpt
import utilities_tools as utt
import main_functions as mvs
import support_functions as spf
```

Requirements

- scipy
- chaosmagpy
- numpy
- pandas
- matplotlib
- glob2
- h5py
- pathlib2
- pwlf
- chaosmagpy
- cartopy (optional)

Modules functions

Here are the principal functions of each module and a quick description.

Functions listed here are fundamental in the geomagnetic data processing.

There are others functions into the modules, including the support_functions module that was not mentioned. Most of them are internally used by the package.

main_functions

```
import main_functions as mvs
```

Function	Description
mvs.load_intermagnet_files(station, starttime, endtime, files_path)	read and merge IAGA-2002 file format into a pandas DataFrame
mvs.sv_obs(station, starttime, endtime, plot_chaos, files_path)	interactive data processing workflow ultil geomagnetic jerk detection
mvs.plot_sample(station, dataframe, save_plots ...)	automatically plot hourly, daily, monthly and annual means
mvs.plot_sv(station, starttime, endtime, df_station, df_chaos, ...)	automatically plot SV with options to correct the data and plot CHAOS prediction
mvs.plot_tdep_map(time, deriv, plot_changes, station)	plot a global map of the SV or SA as well as their changes (from CHAOS prediction)

data_processing_tools

```
import data_processing_tools as dpt
```

Function	Description
dpt.resample_obs_data(dataframe, sample, apply_percentage)	Resample obs minute or hourly means into hourly, daily, monthly or annual means
dpt.hampel_filter_denoising(dataframe, window_size, n_sigmas, ...)	Denoising filter based on median absolute deviation
dpt.night_time_selection(station, dataframe)	Select the nighttime period from geomagnetic data. Default from 22pm to 2am LT.
dpt.keep_quiet_days(dataframe)	Select only top 10 quiet days from each month
dpt.remove_disturbed_days(dataframe)	Remove top 5 disturbed days from each month
dpt.kp_index_selection(dataframe, kp)	Select only periods with Kp values <= the defined limit. Default is 2

Function	Description
dpt.calculate_sv(dataframe, method, source, ...)	Calculate SV from geomagnetic data using monthly or annual means (input must be output from load_intermagnet_files or chaos_model_prediction). Default is monthly means
dpt.chaos_model_prediction(station, starttime, endtime, n_core, ...)	Predict diferent sources of the geomagnetic from CHAOS-7 model (core, crust, magnetospheric (GSM + SM)).
dpt.external_field_correction_chaos_model(station, starttime, endtime, ...)	Correct magnetospheric field from geomagnetic data
dpt.jerk_detection_window(station, window_start, window_end, ...)	Automatically adjust two straight line segments in the SV for an user specified time window

utilities_tools

```
import utilities_tools as utt
```

Function or Class	Description
utt.download_intermagnet_data()	Download observatory quasi-definitive or defintive data from INTERMAGNET fpt server and save in the computer
utt.hdz_to_xyz_conversion(station, dataframe, files_path)	Check the existence of reported HDZ components and convert to XYZ components
utt.IMO(self, station, latitude, longitude, ...)	Class representing IMO. Can be used to check IMO informations on MOSFiT database (IMO existence, latitude, longitude, altitude) as well as add a new IMO or delete

main_functions usage

The main_functions package module consist of functions to load (read) and visualize IAGA-2002 data.

It also has an interactive function sv_obs that includes the most important data processing options.

load_intermagnet_files

This function is the most important, since it reads any geomagnetic data following the IAGA-2002 format.

The output is a pandas DataFrame indexed by time and the columns are the X, Y and Z geomagnetic components.

Its output is used as input is most of the data processing functions.

```
load_intermagnet_files(station = 'XXX', starttime = 'yyyy-mm-dd', endtime = 'yyyy-
mm-dd', files_path = 'path//to//files')
```

The returned dataframe can be manipulated by the user or processed with the others functions of the package.

You can set a specific directory or just use the default (automatically created when the files are downloaded using download_data_INTERMAGNET function).
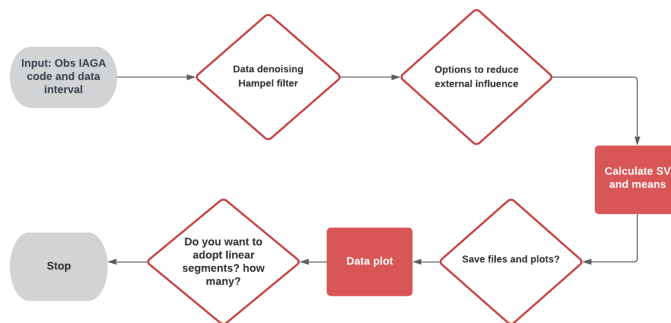
sv_obs

sv_obs is a function that includes the most important data processing options.

The processing according to the figure is already implemented in a dedicated function.

However the user can combine any of the processing steps in any possible order or combination

It allows the user to process the geomagnetic data in a interactive workflow,

using most of the available data_processing functions.


```
sv_obs(station = 'VSS', stattime = '2000-01-01', endtime = '2021-06-30',
files_path = 'path//to//files', plot_chaos = True)
```

plot_samples

Automatically plot hourly, daily, monthly and annual means for X, Y and Z

```
import main_functions as mvs
mvs.plot_samples(station = 'VSS', dataframe: df_name, save_plots = False,
plot_data_type = True, apply_percentage = False )
```

plot_sv

60

Function to automatically plot the SV for an observatory

```
import main_functions as mvs
mvs.plot_sv(station = 'NGK', starttime = None, endtime = None, files_path = None,
df_station = df_name, df_chaos = None, apply_percentage = False, plot_chaos =
True, chaos_correction = True, save_plot = False, convert_hdz_to_xyz = False)
```

Example of SV from NGK automatically created using the function. The CHAOS model internal field predictions is also an option as well as correct the magnetospheric field.
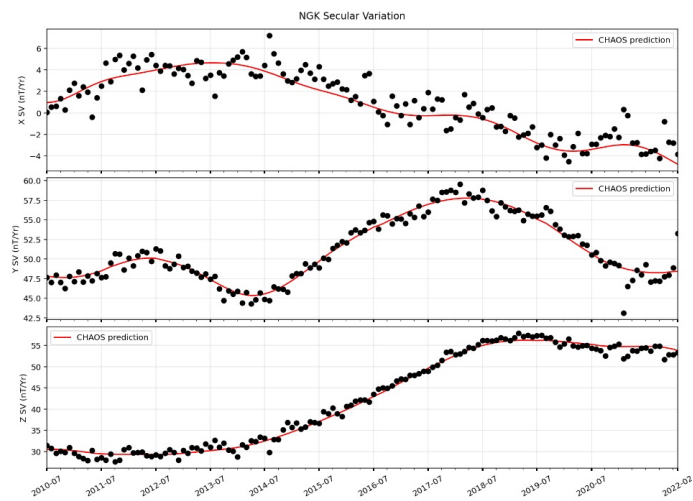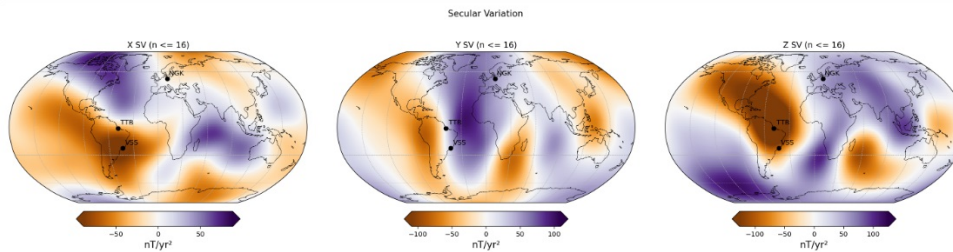

plot_tdep_map

Plot a SV or SA world map for the CHAOS model predictions. It is possible to plot SV and SA changes as well as include 'stations' in the map (see figure)

```
import main_functions as mvs
mvs.plot_tdep_map('2022-07-01', deriv = 1, plot_changes=False, station =
['VSS','NGK','TTB'])
```

Example of output from plot_tdep_map

data_processing_tools functions usage

Here I explain the principal function of the data_processing_tools module.

As the name says such functions are responsable for the data processing.

Most of them are methods to reduce the external field contribution from the observatory data. In order to investigate the SV.
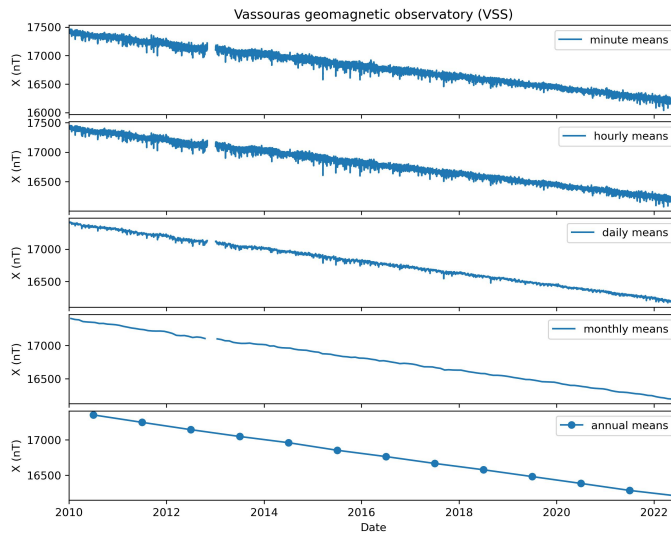
resample_obs_data

This function allows the user to resample geomagnetic observatory

data into different samples (hourly, daily, monthly and annual).

```
import data_processing_tools as dpt
dpt.resample_obs_data(dataframe = df_name, sample = 'H', apply_percentage = True)
```

Example of different data samples calculated using MOSFiT.

hampel_filter_denoising

This function to denoise geomagnetic data based on a median absolute deviation filter

In order to reduce computacional coast the function automatically resample the minute mean data (default from IAGA-2002 data and output from load_intermagnet_files) into hourly mean values

```
import data_processing_tools as dpt
dpt.hampel_filter_denoising(dataframe = df_name, window_size = 200, n_sigmas = 3,
apply_percentage = True, plot_figure = True)
```
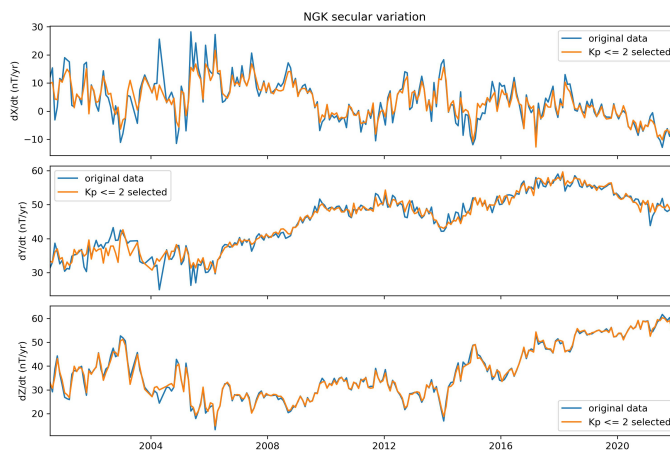
Example of denoised hourly mean data.

kp_index_correction

63
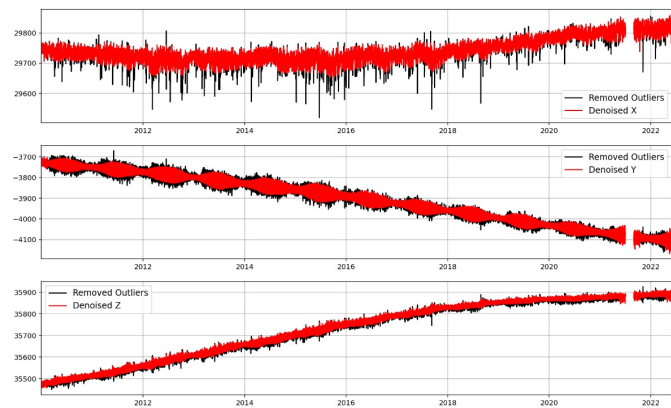
The function removes periods with Kp index values above user input limit from the geomagnetic components

Find the Kp index on https://kp.gfz-potsdam.de/en/

```
import data_processing_tools as dpt
dpt.kp_index_correction(dataframe = df_name, kp = 2)
```


NGK secular variation

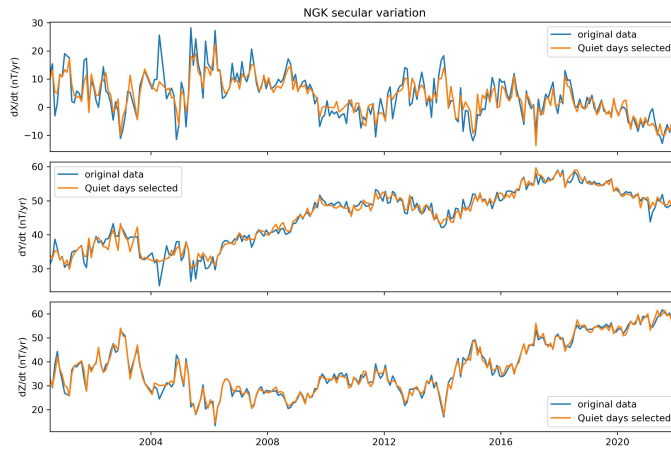## keep_quiet_days

The function select only the top 10 international quiet days from each month

Find the list of quiet days for each month on https://kp.gfz-potsdam.de/en/

```
import data_processing_tools as dpt
dpt.keep_quiet_days(dataframe = df_name)
```

Example of SV calculate using normal data and selecting quiet days for each monthly.
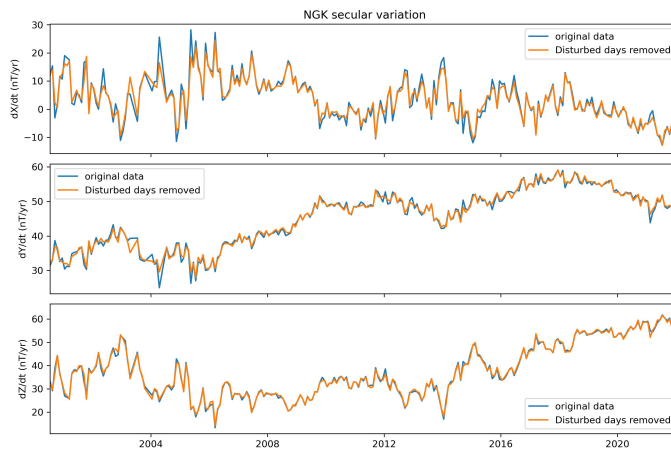
remove_disturbed_days

The function remove the top 5 international disturbed days from each month

Find the list of disturbed days for each month on https://kp.gfz-potsdam.de/en/

```
import data_processing_tools as dpt
dpt.remove_disturbed_days(dataframe = df_name)
```

Example of SV calculate using normal data and removing the top 5 disturbed days from each month.
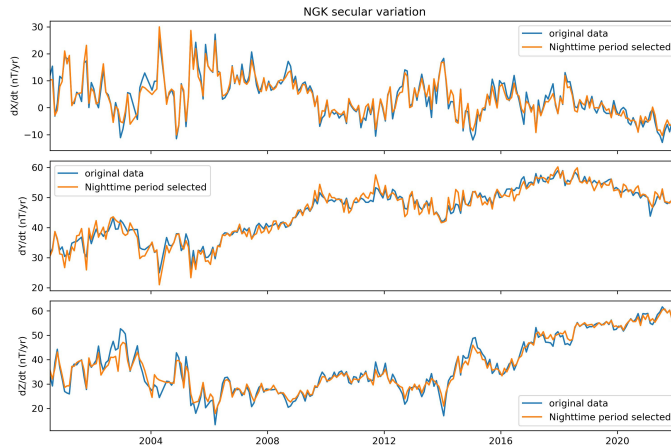

night_time_selection

The function select the nighttime period from the geomagnetic data (default from 22pm to 2 am LT)

```
import data_processing_tools as dpt
dpt.night_time_selection(station = 'XXX', dataframe = df_name)
```

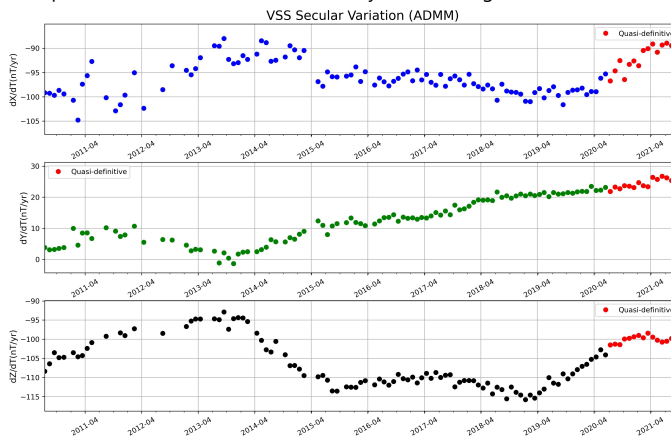Example of SV calculate using normal data and selecting only nighttime period.

calculate_sv

Calculate geomagnetic secular variation using monthly means or annual means

```
import data_processing_tools as dpt
dpt.calculate_sv(dataframe = df_name, method = 'ADMM')
```

Example of SV calculate from VSS monthly means using MOSFiT.

chaos_model_prediction

Predict core fiel, crustal field and magnetospheric field (GSM and SM) from CHAOS-7 model predictions in a hourly rate

find the model realease on http://www.spacecenter.dk/files/magnetic-models/CHAOS-7/

References Finlay, C.C., Kloss, C., Olsen, N., Hammer, M. Toeffner-Clausen, L., Grayver, A and Kuvshinov, A. (2020), The CHAOS-7 geomagnetic field model and observed changes in the South Atlantic Anomaly, Earth Planets and Space 72, doi:10.1186/s40623-020-01252-9 [.pdf]

Finlay, C.C., Kloss, C., Olsen, N., Hammer, M. and Toeffner-Clausen, L., (2019) DTU Candidate models for IGRF-13. Technical Note submitted to IGRF-13 task force, 1st October 2019 [.pdf]

Example of how to use MOSFiT chaos_model_prediction. The station (3 letter IAGA code) must be in the MOSFiT imos database. All INTERMAGNET observatories are included in the database automatically. If you are interest in predict the field for other observatory or location, use the 'IMO' MOSFiT class to add the location in the database. See utilities_tools section for an explanation about how to include the location.

```
import data_processing_tools as dpt
dpt.chaos_model_prediction(station = 'XXX', starttime = 'yyyy-mm-dd', endtime =
'yyyy-mm-dd', n_core = 20, n_crust = 110, n_gsm = 2, n_sm = 2)
```

external_field_correction_chaos_model

Subtract the magnetospheric field (GSM and SM) from CHAOS-7 model predictions from the observatory data

find the model realease on http://www.spacecenter.dk/files/magnetic-models/CHAOS-7/

References Finlay, C.C., Kloss, C., Olsen, N., Hammer, M. Toeffner-Clausen, L., Grayver, A and Kuvshinov, A. (2020), The CHAOS-7 geomagnetic field model and observed changes in the South Atlantic Anomaly, Earth Planets and Space 72, doi:10.1186/s40623-020-01252-9 [.pdf]

Finlay, C.C., Kloss, C., Olsen, N., Hammer, M. and Toeffner-Clausen, L., (2019) DTU Candidate models for IGRF-13. Technical Note submitted to IGRF-13 task force, 1st October 2019 [.pdf]

Example of how to use MOSFiT external_field_correction_chaos_model.

```
import data_processing_tools as dpt
dpt.external_field_correction_chaos_model(station = 'XXX', starttime = 'yyyy-mm-
dd', endtime = 'yyyy-mm-dd',df_station = None, df_chaos = None, n_core = 20,
n_crust = 110, n_gsm = 2, n_sm = 2)
```

Example of SV calculate from VSS monthly means using MOSFiT magnetospheric field correction from CHAOS predictions.
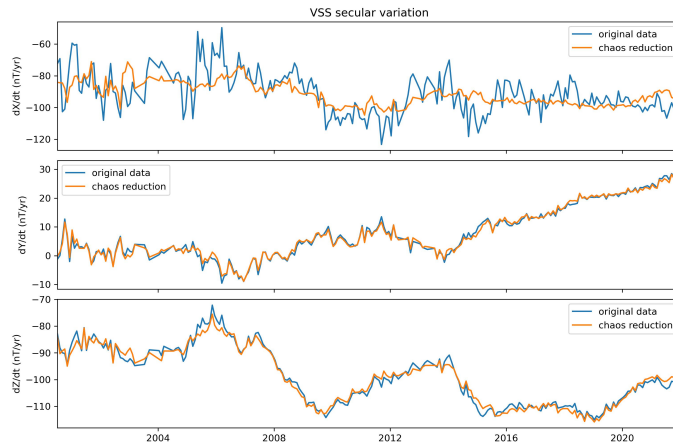
jerk_detection_window

Automatically fits two straight line segments for an user selected window. Determine the geomagnetic jerk occurrence time (t0), amplitude (A) and coefficiente of determination (R2).

The function uses the occurrence time as input in the plot_tdep_map to plot the SA changes for the jerk detection.

```python
import data_processing_tools as dpt
dpt.jerk_detection_window(station = 'NGK', window_start = '2012-04',  window_end =
'2017-08',  starttime = '2010-01-01',  endtime = '2021-06-30', df_station = None,
df_chaos = None, files_path = None, plot_detection = True, chaos_correction =
True, plot_chaos_prediction = False, convert_hdz_to_xyz = False, save_plots =
False)
```

Geomagnetic jerk detection
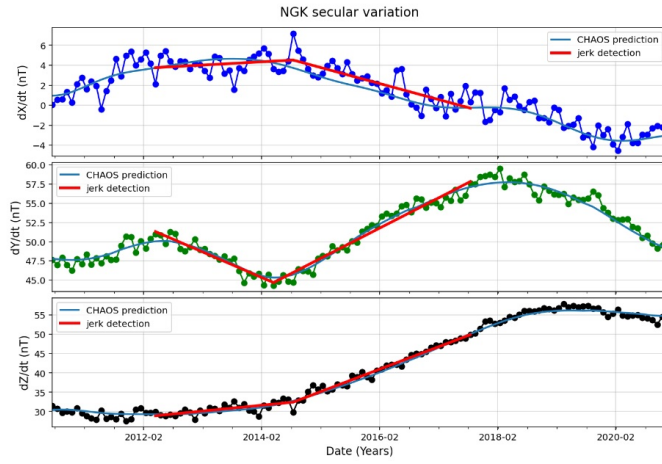
Statistics

```
NGK Jerk statistics for the X component.

Jerk occurence time -t0-: 2014.62
Jerk amplitute: -1.93
R^2: 0.68

****************************************************************

NGK Jerk statistics for the Y component.

Jerk occurence time -t0-: 2014.28
Jerk amplitute: 7.26
R^2: 0.95

****************************************************************

NGK Jerk statistics for the Z component.

Jerk occurence time -t0-: 2014.62
Jerk amplitute: 4.32
R^2: 0.98

****************************************************************
```
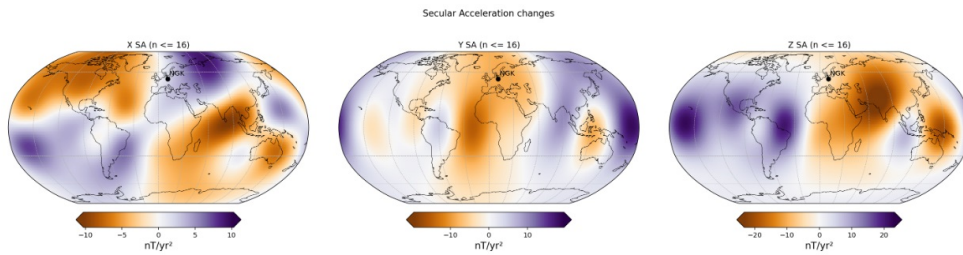
Secular acceleration changes for the jerk occurrence time

utilities_tools usage

Here I describe some functions that let easier the use of the package.

IMO

Class used to check the IMOs database, IMO coordinates and to delete or add a new IMO in the database.

Very useful if you want to add a specific location in the world, station or observatory. Adding the location you can use all package features if your data format is IAGA-2002.

Checking existent database

```
import utilities_tools as utt
utt.IMO.database()
```

Imos	Latitude	Longitude	Elevation
AAA	43.250	76.920	1300.0
ABG	18.638	72.872	6.0
ABK	68.358	18.823	380.0
AIA	-65.245	-64.258	10.0
API	-13.815	-171.781	2.0
...	...	...	...
TAN	-18.917	47.552	1375.0
TEO	19.747	-99.182	2280.0
TIK	71.580	129.000	57.0

The coordinates can be checked by:

```
import utilities_tools as utt
utt.IMO.latitude('VSS')
utt.IMO.longitude('VSS')
utt.IMO.elevation('VSS')
```

Adding a new IMO

```
import utilities_tools as utt
utt.IMO.add('XXX', 123 , 123, 123)
```

Deleting an IMO

```
import utilities_tools as utt
utt.IMO.delete('XXX')
```

70

Checking the existince of an especific IMO on the database

```
import utilities_tools as utt
utt.IMO.check_existence('XXX')
```

hdz_to_xyz_conversion

Sometimes quasi-definitive IAGA-2002 data are submited containing HDZ components (also some data in the past). This Function search for periods with HDZ components reported and convert to XYZ.

Checking the existince of an especific IMO on the database

```
import utilities_tools as utt
utt.hdz_to_xyz_conversion('XXX', dataframe = df_name, files_path =
'path//to//files')
```