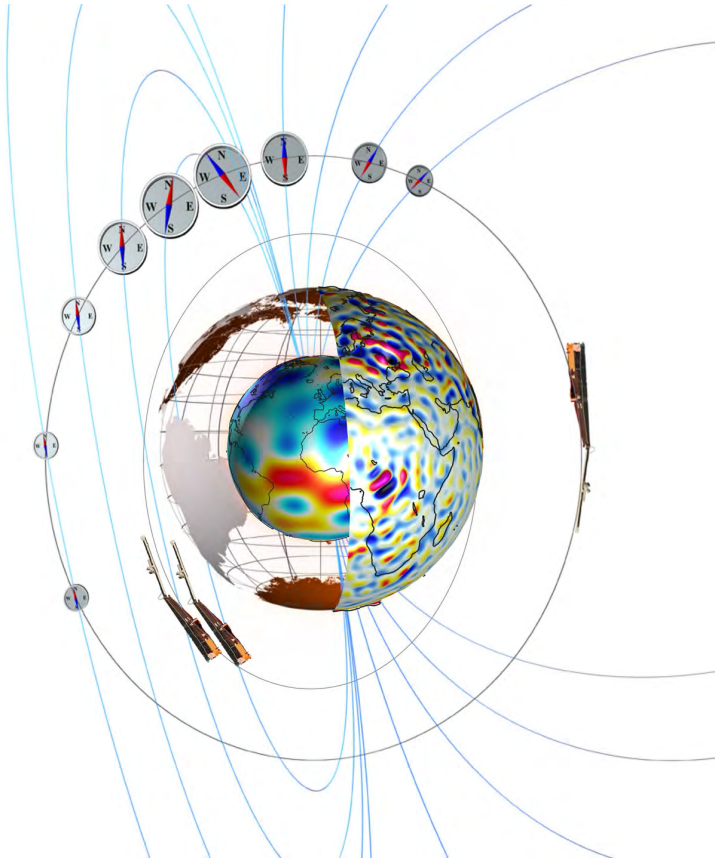
Data, Innovation, and Science Cluster

Description of Algorithm



Doc.No: SW-DS-DTU-GS-005, Rev: 3

Prepared:

Magnus Danel Hammer Date 11–02–2021
Scientist

Prepared:

William J. Brown Date 11–02–2021
Scientist

Approved:

Digitally signed by Poul Erik Holmdahl Olsen
Date: 2021.02.11 11:37:04 +01'00'

Poul Erik Holmdahl Olsen Date 11–02–2021
Project Manager

Checked:

Christopher Finlay Date 11–02–2021
Project Lead

Record of changes

Reason	Description	Rev	Date
Draft	Draft version of the DPA document	1 dA	11-12-2019
Issue 1	Final revision	1	12-05-2020
Issue 2	The following sections have been updated: -Sections 4.2, 4.3: Derivation of the GVO potential has been added and the potential expression has been updated. Definition of coordinate frames at the North and South Poles have been added.	2	27-05-2020
Issue 3	Updates after review from DISC review panel. In addition, following has been updated: -Section 4, Fig.2. -Section 4.2, eqs.(4)-(7) and eq.(11). -Section 4.4.1, eq.(13) and associated text. -Section 4.5.1, eqs.(18)-(19) and associated text replaced with new eqs.(18)-(21) and updated text. -Section 4.5.3, eq.(32). -Reference (Bendat and Piersol, 2010) has been added. -Symbols list has been updated.	3	11–02–2021

Table of Contents

Acrynoms

AT	Along-Track
CHAOS	Geomagnetic Field Model [Finlay et al., 2016]
CI	Comprehensive Inversion [Sabaka et al., 2018]
CT	Cross-Track
DFT	Descrete Fourier Transform
ECEF	Earth-Centered-Earth-Fixed
ESA	European Space Agency
EQ	Equal Area
GSM	Geocentric Solar Magnetospheric Coordinate System
IGRF	International Geomagnetic Reference Field [Alken et al., 2020]
IMF	Interplanetary Magnetic Field
LCS-1	Lithospheric Field Model [Olsen et al., 2017]
MF	Main Field
NEC	North-East-Center Coordinate System
OMNI	Solar wind magnetic field and plasma data at Earth's Bow Shock Nose
PCA	Principal Component Analysis
RC	Ring Current index
rms	Root-Mean-Square
SEU	South-East-Up Coordinate System
SH	Spherical Harmonic
SHA	Spherical Harmonic Analysis
SM	Solar Magnetospheric
Sq	Solar Quiet
SV	Secular Variation
SWA	*Swarm* satellite *Alpha*
SWB	*Swarm* satellite *Bravo*
SWC	*Swarm* satellite *Charlie*
VFM	Vector Field Magnetometer
GVO	Geomagnetic Virtual Observatory

Symbols and notation

Symbol	*Name*
$\delta\mathbf{B}_{GVO}$	GVO estimation of residual vector field
$\delta\mathbf{B}^{obs}$	Magnetic field residual vector used to derive estimates of the observed field GVOs
$\delta\mathbf{B}^{core,1month}$	Magnetic field residual vector used to derived one-monthly core field GVOs
$\delta\mathbf{B}^{core,4month}$	Magnetic field residual vector used to derived four-monthly core field GVOs
$\delta\dot{\boldsymbol{B}}$	Secular variation residuals used in PCA denoising analysis
Δd^{AT}	Along-track magnetic data differences
Δd^{EW}	East-west magnetic data differences
ΔG	Data kernels associated with data differences
δt	Time difference
δX	Residual magnetic field in northward direction
δY	Residual magnetic field in eastward direction
δZ	Residual magnetic field in downward direction
ϵ_i	Normalized error
λ_m	Magnetic latitude
μ	Mean
μ_w	Weighted mean
ν	GVO Potential coefficients
Σd	Magnetic data sums
Σd^{AT}	Along-track magnetic data sums
Σd^{EW}	East-west magnetic data sums
σ_{core}	Error estimates of Core Field GVOs
σ_{obs}	Error estimates of Observed Field GVOs
A	Area of spherical surface
a	Mean Earth radius, 6371.2km (also r_a)
B_l	Magnetic field component (l can be r, θ or ϕ)
B_r	Magnetic field component in radial direction
B_θ	Magnetic field component in meridional (southward) direction
B_ϕ	Magnetic field component in zonal/azimuthal (eastward) direction
B_y	IMF field component in dusk-dawn direction (GSM frame)
B_z	IMF field component in northward direction (GSM frame)
$\mathbf{B}^{lith}$	Lithospheric magnetic vector field
$\mathbf{B}^{ext}_{pol}$	External poloidal magnetic field
$\mathbf{B}^{int}_{pol}$	Internal poloidal magnetic field
$\mathbf{B}^{ext}_{tor}$	External toroidal magnetic field
$\mathbf{B}^{int}_{tor}$	Internal toroidal magnetic field
$\mathbf{B}^{iono}$	Ionospheric magnetic vector field
$\mathbf{B}^{mag}$	Magnetospheric magnetic vector field
$\mathbf{B}^{MF}$	Main magnetic vector field
$\mathbf{B}^{MF}_{GVO}$	IGRF magnetic vector field at GVO location
$\mathbf{B}_{GVO}$	GVO magnetic vector field
$\mathbf{B}^{obs}$	Satellite magnetic vector field measurements
$\mathbf{B}^{sh}_{pol}$	Poloidal magnetic field in a shell
$\mathbf{B}^{sh}_{tor}$	Toroidal magnetic field in a shell

Continues on next page

$\widehat{\mathbf{B}}_{corr}$	PCA de-noised magnetic field
$C_{a,b,c}$	Potential expansion coefficients
$\underline{\underline{\mathbf{C}}}_D$	GVO secular variation residuals covariance matrix
c_w	Huber turning constant
$\underline{\underline{\mathbf{D}}}$	GVO secular variation residuals matrix
$\underline{\underline{\widehat{\mathbf{D}}}}$	Corrected GVO secular variation residuals matrix
$\mathbf{d}$	Data vector
$\mathbf{d}_{CHAOS}$	Data vector of CHAOS internal field predictions
$\mathbf{d}_{GVO}$	Data vector of GVO data
$\mathbf{d}_{vec}$	Data vector of residual field vector components
$\underline{\underline{\mathbf{E}}}$	Eigenvector matrix
$\underline{\underline{\widehat{\mathbf{E}}}}$	Corrected eigenvector matrix
E_m	Merging electric field
e_i	Error between data and model prediction
$\hat{\mathbf{e}}_\mathbf{r}$	Spherical unit vector
$\hat{\mathbf{e}}_\theta$	Spherical unit vector
$\hat{\mathbf{e}}_\phi$	Spherical unit vector
$\hat{\mathbf{e}}_\mathbf{x}$	Cartesian unit vector
$\hat{\mathbf{e}}_\mathbf{y}$	Cartesian unit vector
$\hat{\mathbf{e}}_\mathbf{z}$	Cartesian unit vector
$F_{10.7}$	Solar flux index at wavelength 10.7cm
Flags_B	Flags related to the magnetic field vector measurement
Flags_q	Flags related to the attitude data
$\underline{\underline{\mathbf{G}}}$	GVO design matrix
$\underline{\underline{\mathbf{G}}}_{SH}$	Spherical harmonic design matrix
$\underline{\underline{\mathbf{G}}}_{vec}$	GVO design matrix for vector data
g_n^m	Internal Gauss coefficients
h_n^m	Internal Gauss coefficients
h_{GVO}	Altitude of GVO above r_a
$\hat{\mathbf{i}}_\mathbf{1}$	Geocentric unit vector
$\hat{\mathbf{i}}_\mathbf{2}$	Geocentric unit vector
$\hat{\mathbf{i}}_\mathbf{3}$	Geocentric unit vector
$\mathbf{J}_{tor}^{ext}$	Toroidal electrical currents external to the ionosphere
$\mathbf{J}_{tor}^{int}$	Toroidal electrical currents internal to the Earth's surface
$\mathbf{J}_{tor}^{sh}$	Toroidal electrical currents within the ionosphere
K	Number of principal components removed
Kp	Planetary geomagnetic activity index
m	Spherical harmonic order
$\mathbf{m}$	GVO model parameters vector
$\mathbf{m}_{SH}$	Spherical harmonic model vector
N_{GVO}	Number of GVO's
n	Spherical harmonic degree
$\underline{\underline{\mathbf{P}}}$	Principle component matrix
$\underline{\underline{\widehat{\mathbf{P}}}}$	Corrected principle component matrix
P_n^m	Associated Legendre functions of degree n and order m
q_n^m	External Gauss coefficients

Continues on next page

r_1	Radius of shell 1
r_2	Radius of shell 2
r_3	Radius of shell 3
r_a	Mean Earth radius, 6371.2km (also a)
r_{cyl}	GVO target cylinder search radius
$\mathbf{r}_{GVO}$	GVO position vector
rms	Root-mean-square
rms_w	Weighted root-mean-square
$S(r_1, r_3)$	Shell where magnetic measurements are taken
$\mathbf{SV}_{GVO}$	GVO secular variation vector
s_n^m	External Gauss coefficients
$t_n^{m,c}$	Toroidal expansion coefficients
$t_n^{m,s}$	Toroidal expansion coefficients
T^{sh}	Toroidal magnetic scalar potential
V^{int}	Internal magnetic scalar potential
V^{ext}	External magnetic scalar potential
V	Magnetic scalar potential
v	Solar wind velocity
w	Huber weights
$\underline{\underline{\mathbf{W}}}$	Matrix of Huber weights
X	Magnetic field component in meridional (northward) direction
Y	Magnetic field component in zonal/azimuthal (eastward)
Z	Magnetic field component in vertical downward direction

1 Introduction

1.1 Scope and Applicability

This document presents the Description of the Processing Algorithm (DPA) for the Swarm DISC Geomagnetic Virtual Observatories (GVOs) products for the Swarm Data, Innovation and Science Cluster (Swarm DISC) consortium, as part of the Swarm DISC Activity "New Products and Services". With reference to the Statement of Work, Ref.[AD-3], the aim is to use Swarm satellite vector magnetic field measurements to produce time series of local point field estimates referred to as a "Geomagnetic Virtual Observatory" (GVO) provided at fixed locations on a uniform global grid at satellite altitude. The primary project activities are (Ref.[AD-3]):

- Produce time series of the geomagnetic field vector over a network of GVOs representing the summed contribution of all potential field sources hereafter termed the **Observed Field GVOs**
- Produce time series of the geomagnetic field vector, and its first time derivative, over the same network of GVOs, representing the contribution of the core field extracted from the observed field in which the other sources are (as far as possible) removed, hereafter termed **Core Field GVOs and Secular Variation**

For each of the above listed activities, two sets of GVO data products were produced:

- **One-monthly time series**. These are relevant when knowledge of the geomagnetic field and its time changes on timescales shorter than 4 months but longer than 1 month are desired. A sophisticated processing scheme is needed to isolate the core field, for example in order to remove local time effects.
- **Four-monthly time series**. These are relevant when the focus is on variations with timescales longer than four months. A simpler processing scheme is used to isolate the core field.

The Swarm DISC GVO products are designed to make Swarm magnetic data easily accessible to researchers investigating various aspects of the geomagnetic field at fixed geographic locations on long time scales. In particular they are useful for studying the physics of the core dynamo process, and related phenomenon such as secular variation, geomagnetic jerks and rapid core dynamics. The observed field GVO data products also provide information that can be used to investigate magnetospheric and ionospheric magnetic signals on timescales of months and longer.

The Swarm DISC Geomagnetic Virtual Observatory scheme provides the magnetic field data products listed in Table 3.

Current or updated versions of this document and of the Swarm DISC Geomagnetic Virtual Observatories Product Definition Document, Ref.[AD-2], are available in the SVN folder: https://smart-svn.spacecenter.dk/svn/smart/SwarmDISC/DISC_Projects/ITT2_1_GVO/Deliverables

1.2 Applicable Documents

The following documents are applicable to the definitions within this document

[AD-1] SW-OF-DTU-GS-121, Proposal for Swarm DISC ITT 1.2, Swarm Geomagnetic Virtual Observatories

Product file name	Product description
VOBS_1M_2_	One-monthly time series of the vector magnetic field: the Observed Field, the Core Field and the Secular Variation, and their associated error estimates, all provided in a global grid of geomagnetic virtual observatories
VOBS_4M_2_	Four-monthly time series of the vector magnetic field: the Observed Field, the Core Field and the Secular Variation, and their associated error estimates, all provided in a global grid of geomagnetic virtual observatories

Table 3: Swarm DISC Geomagnetic Virtual Observatories product overview.

[AD-2] SW-DS-DTU-GS-2101_GVO_PDD, Swarm Geomagnetic Virtual Observatories Product Definition

[AD-3] SW-SW-DTU-GS-121, Statement of Work

2 Scientific Background

2.1 Geomagnetic Virtual Observatory (GVO) Method

The Geomagnetic Virtual Observatory method was first proposed by Mandea and Olsen [2006] as a tool for making satellite magnetic field measurements easily accessible as time series of the vector geomagnetic field at pre-specified locations. The GVO method provides a procedure where a scalar magnetic potential is fitted to satellite magnetic field observations from a chosen time window and within a local region, defined by a cylinder centred on a GVO target point. The potential is then used to compute the magnetic field at the GVO target point such that a mean magnetic field over a chosen time window at satellite altitude is determined, as illustrated in Figure 1. The GVO time series thus mimics the time series produced by ground-based magnetic observatories on timescales of months and longer. The main advantage of the GVO time series is that they can be produced at any sites of interest that are covered by satellite data, and in particular, can provide a global grid of times series derived from measurements made by similar instruments onboard satellites such as the *Swarm* trio.

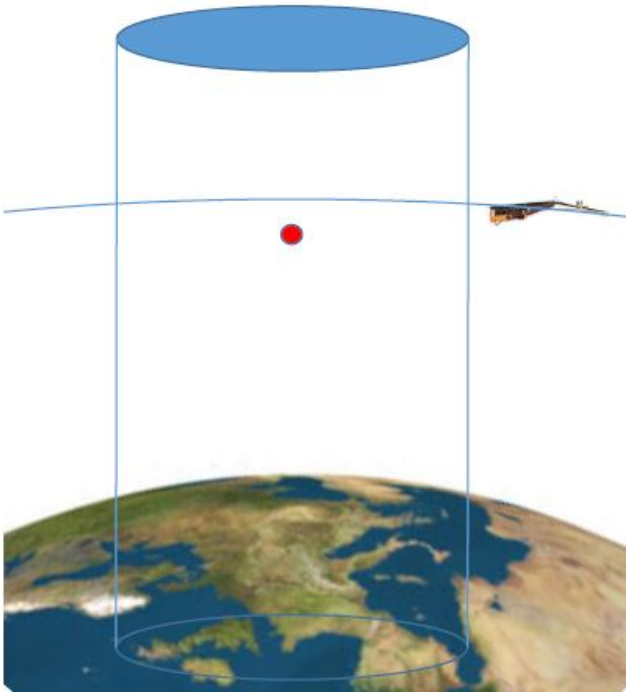


Figure 1: Illustration of the GVO concept, satellite measurements from within a target cylinder are used to infer time series at the GVO location given by a red dot.

Applications of the GVO time series include geomagnetic jerk studies [Olsen and Mandea, 2007], comparisons with spherical harmonic (SH) based geomagnetic field models [Olsen et al., 2010], core flow studies [Kloss and Finlay, 2019, Rogers et al., 2019] and data assimilations studies [Barrois et al., 2018]. The GVO method can also be used to derive estimates of magnetic field gradients [Hammer, 2018]. The GVO datasets presented here will thus provide a powerful new platform for researchers to access and explore Swarm magnetic measurements. Focusing on the core magnetic field, initial studies showed that the original GVO series were contaminated by external sources [Beggan et al., 2009, Olsen and Mandea, 2007, Shore, 2013, Domingos et al., 2019]. Recommendations for improving the original GVO concept and removing such contamination have

been proposed [Shore, 2013, Hammer, 2018]. Some of these improvements were implemented in more recent GVO series that have been used for core flow studies by Barrois et al. [2018], Kloss and Finlay [2019], Whaler [2017]. The Swarm DISC product GVOs will be computed using the most important of these recommendations and will also take advantage of recent efforts to improve the recovery of the core field signal using principal component analysis (PCA) [Cox et al., 2018].

3 Description of the Input Data

This section describes the satellite magnetic field measurements used to derive the GVO time series. The Swarm DISC GVO products use Swarm vector magnetic field measurements in the form of the Swarm Level 1b (L1b) product MAGX_LR_1B, which contains quality-screened, calibrated and corrected measurements given in physical SI units (nT) in a (North, East, Center), hereinafter NEC, reference frame. The Level 1b products are provided individually for each of the three Swarm satellites Alpha, Bravo, and Charlie on a daily basis https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/swarm. For the Swarm DISC product described in this document, data from all three Swarm satellites are used with a subsampling of 15s from low rate (1Hz) data. Measurements having Flags_B=255 or Flags_q=255, which specifies non-valid magnetometer or attitude data, are rejected as are data from known disturbed days when satellite manoeuvres took place.

Two data chains are produced:

- **Data chain (a)** extracts all available data from Swarm L1b MAGX_LR_1B using a sub-sampling of 15s
- **Data chain (b)** extracts Swarm L1b MAGX_LR_1B using a sub-sampling of 15s and applies a dark, geomagnetic quiet-time, selection criteria defined as:
  - Gross data outliers for which the vector field components deviate more than 500 nT from the predictions of the latest CHAOS field model [Finlay et al., 2016] are rejected
  - The sun is at least 10° below horizon
  - Geomagnetic activity index $K_p < 3°$
  - Time change of Ring current (RC) index $|dRC/dt| < 3\mathrm{nT/hr}^{-1}$, [Olsen et al., 2010]
  - Merging electric field at the magnetopause $E_m < 0.8\mathrm{mVm}^{-1}$, [Olsen et al., 2010]
  - Constraints on IMF requiring $B_z > 0\,\mathrm{nT}$ and $|B_y| < 10\,\mathrm{nT}$

Two-hourly means of 1 min values of the solar wind and IMF are computed based on the OMNI data-base, http://omniweb.gsfc.nasa.gov.

The GVO method described here makes use of sums and differences of the satellite magnetic field measurements written $B_l(\mathbf{r}) = \hat{\mathbf{l}} \cdot \mathbf{B}(\mathbf{r})$ where $\hat{\mathbf{l}}$ is the appropriate unit vector in a given coordinate system (e.g. in Cartesian coordinates ($l = x, y, z$)), such that Σd_l and Δd_l denotes the data sums and differences, respectively. Here we denote the north-south (NS) (i.e. the along-track) and east-west (EW) data differences by $\Delta d_l = (\Delta d_l^{\mathrm{NS}}, \Delta d_l^{\mathrm{EW}})$, and the data sums by $\Sigma d_l = (\Sigma d_l^{\mathrm{NS}}, \Sigma d_l^{\mathrm{EW}})$. The along-track data differences are calculated using the 15s differences $\Delta d_l^{\mathrm{AT}} = [B_l(\mathbf{r}, t) - B_l(\mathbf{r} + \delta\mathbf{r}, t + 15s)]$, where $\delta\mathbf{r} = (\delta r, \delta\theta, \delta\phi)$ is the change in position. A 15 s along-track difference with a satellite speed of $\approx 7.7\,\mathrm{km/s}$ corresponds to a distance of 115 km [Olsen et al., 2015]. The north-south sums are calculated as $\Sigma d_l^{NS} = [B_l(\mathbf{r}, t) + B_l(\mathbf{r}+\delta\mathbf{r}, t+15\,s)]/2$.

The east-west differences were calculated as $\Delta d_l^{\mathrm{EW}} = [B_l^{\mathrm{SWA}}(\mathbf{r}_1, t_1) - B_l^{\mathrm{SWC}}(\mathbf{r}_2, t_2)]$ having an East-West orbit separation between the *Swarm Alpha* (SWA) and *Charlie* (SWC) satellites of $\approx 1.4°$ corresponding to 155 km at the equator [Olsen et al., 2015]. The east-west sums were calculated as $\Sigma d_l^{\mathrm{EW}} = [B_l^{\mathrm{SWA}}(\mathbf{r}_1, t_1) + B_l^{\mathrm{SWC}}(\mathbf{r}_2, t_2)]/2$. For a particular orbit of *Swarm Alpha* the corresponding *Swarm Charlie* data were selected to be those closest in colatitude with the condition that $|\Delta t| = |t_1 - t_2| < 50$ s.

4 Algorithm Description

This section describes the algorithm for deriving GVO time series for the Observed Field, and for the Core Field and Secular Variation. The Swarm DISC GVOs are delivered as time series of the spherical polar components of the magnetic field vector. Time series are constructed on an equal-distance global grid of 300 GVOs located at an altitude of 490 km above the mean Earth radius and are provided as:

1) One and four monthly time series of the observed magnetic field

2) One and four monthly time series of the core magnetic field and its secular variation

The algorithm flow used to derive the product data files is presented in Figure 2 and involves the following essential processing steps :

i Extract Swarm L1b data producing the two data chains: a) using no data selection and b) using a dark quiet time selection scheme

ii For each data point, compute predictions for the main field (using the IGRF model), the lithospheric field (using the LCS-1 model), the magnetospheric field (using the CHAOS model) and the ionospheric field (using the CI model)

iii Apply the GVO method on one-monthly windows:
-compute one-monthly Observed Field GVO time series based on data chain a)
-compute one-monthly GVO time series based on data chain a) and also subtract the predictions for the lithospheric field

iv Apply the GVO method on four-monthly windows:
-compute four-monthly Observed Field GVO time series based on data chain b)
-compute four-monthly GVO time series based on data chain b) and also subtract the predictions from models of the lithospheric, magnetospheric and ionospheric fields

v Perform a Principal Component Analysis on the one-monthly time series with the predictions for the lithospheric field subtracted, by implementing the approach introduced by Cox et al. [2018], in order to reduce contamination due to external field signals and local time sampling biases, with the aim of isolating the Core Field signal in the GVO time series

vii Perform a Spherical Harmonic Analysis (SHA) with internal, external and toroidal expansion terms, on the one-monthly PCA de-noised time series and the four-monthly time series with the predictions from models of non-core sources subtracted.

viii Compute one- and four-monthly core GVO time series of the Core Field by subtracting the predictions of the external and toroidal fields obtained from the above SHA. This aims to remove remaining contamination from external fields and horizontal fields due to in-situ electrical currents [Olsen and Mandea, 2007]. Finally estimates of the Secular Variation are obtained using annual differences.

Subsection 4.2 provides a mathematical description of how GVO estimates are obtained from satellite data. Subsection 4.3 describes the global grid of GVOs employed in this project. Subsection 4.4 describes the algorithm used to derive the Observed Field GVOs. Subsection 4.5 describes the algorithm used to derive the Core Field GVOs and associated Secular Variation time series. 4.5.1 describes the algorithm used to perform Principal Component Analysis on the observed field GVO time series, which is an intermediate processing step towards deriving the Core Field GVOs. Subsection 4.5.2 describes the algorithm for performing Spherical Harmonic Analysis on the one monthly PCA de-noised data series and four monthly data series, and the final steps before computing the Core Field GVO time series.

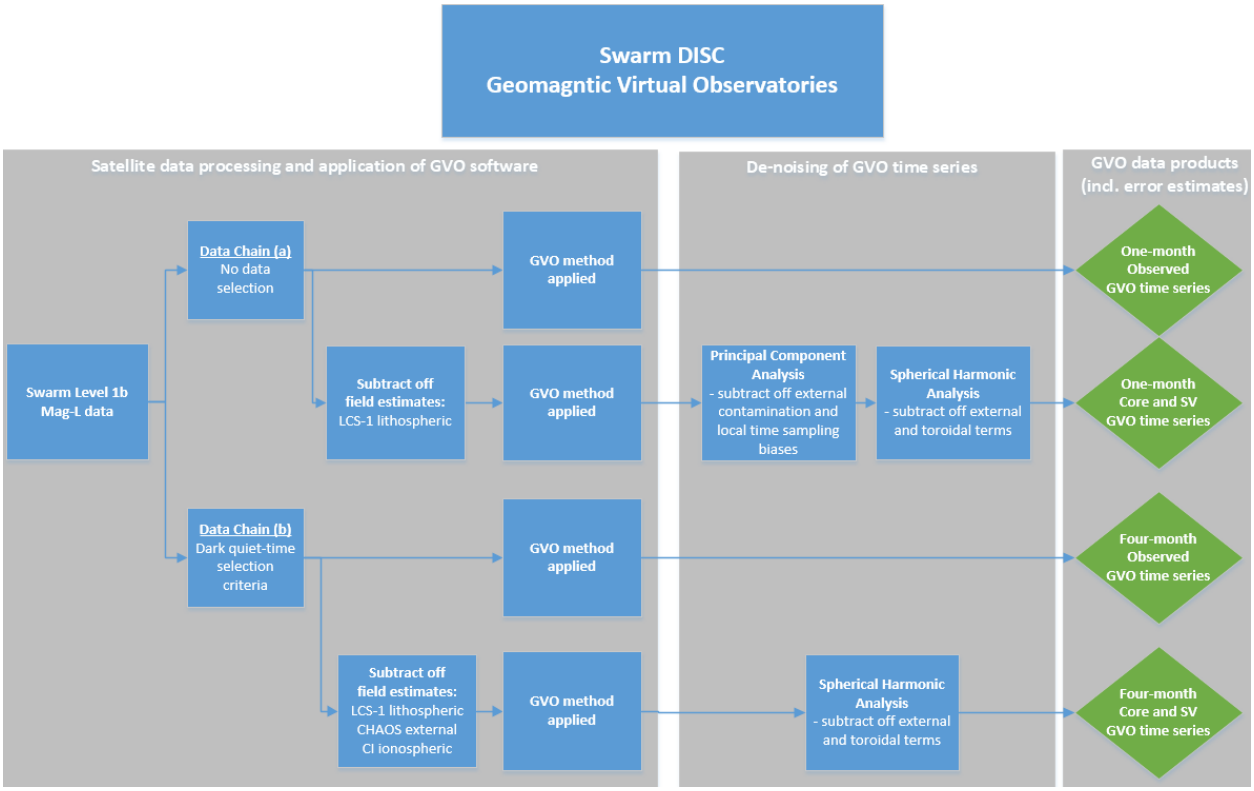


Figure 2: Overview of the Swarm DISC GVO processing flow.

4.1 Data Pre-processing

Within a specified time window of either 1 or 4 months for a given fixed GVO target location, input data as described in Section 3 from within a target cylinder having horizontal dimensions of $d_{GVO} = 700\,\text{km}$ radius, surrounding that particular GVO location, are extracted. The GVO location is provided in spherical polar coordinates as $\mathbf{r}_{GVO} = (r, \theta, \phi)$, where $r = r_a + h_{GVO}$ such that h_{GVO} specifies the height above mean Earth radius $r_a = 6371.2\,\text{km}$. The GVO altitude is selected to be $h_{GVO} = 490\,\text{km}$, such that the GVOs are located at approximately the mean orbital height of the Swarm satellites during 2013-2020.

The relevant data time windows are:

- For the one-monthly GVOs all data from within each month are used. GVO estimates are allocated the time corresponding to the middle of each month.

- For the four-monthly GVOs, all data satisfying the selection criteria within a given four month interval, chosen as January-April, May-August, and September-December, are used. The GVO estimates are allocated times corresponding to the approximate center times of the 1st March, the 1st July and the 1st November.

In order to derive the Observed Field GVO time series, we start from the geocentric spherical polar components of the vector magnetic field measured by *Swarm*, $\mathbf{B}^{obs} = (B_r, B_\theta, B_\phi)$, and subtract predictions from the latest main field model for spherical harmonic degrees $n \in [1, 13]$, $\mathbf{B}^{MF}$, which leads to the Observed Field residuals

$$\delta\mathbf{B}^{obs} = \mathbf{B}^{obs} - \mathbf{B}^{MF}. \tag{1}$$

It is these field residuals that are used in the GVO estimation scheme to derive the observed field GVO time series, with the model predictions at the target points being added back following the estimation (see below for more details).

Other variants of field residuals, in which estimates of other non-core fields are removed in addition, are computed as needed for project internal deliverable (PID) GVO time series (see Figure 2 for an illustration of the process flow). In order to derive one-monthly Core Field GVO time series in chain a) (see Sub-section 4.5.1) residuals are required with the lithospheric field also removed:

$$\delta\mathbf{B}^{core,1month} = \mathbf{B}^{obs} - \mathbf{B}^{MF} - \mathbf{B}^{lith}, \tag{2}$$

where $\mathbf{B}^{lith}$ is the predicted static internal lithospheric field for SH degrees $n \in [14, 185]$ as given by the LCS-1 model [Olsen et al., 2017]. To derive four-monthly core field GVO time series in chain b) (see Sub-section 4.5) models of the magnetospheric and ionospheric fields are also be removed:

$$\delta\mathbf{B}^{core,4month} = \mathbf{B}^{obs} - \mathbf{B}^{MF} - \mathbf{B}^{lit} - \mathbf{B}^{mag} - \mathbf{B}^{iono}, \tag{3}$$

where $\mathbf{B}^{mag}$ is the predicted large-scale magnetospheric field and its Earth induced counterpart field, as given by the CHAOS model [Olsen et al., 2014], and $\mathbf{B}^{iono}$ is the predicted ionospheric field and its Earth induced field as given by the CI model [Sabaka et al., 2018].

The reason for subtracting time-dependent estimates of the main field from the data before carrying out the GVO estimation, and then afterwards adding them back, is that it acts as a pre-whitening of the data, enabling Huber weights required in robust estimation scheme to be more effectively determined.

Note that when considering a pre-specified GVO time window of 1 or 4 months, any information on time variations with periods shorter than 1 or 4 months respectively is, of course, lost.

To summarize, the algorithm flow for selecting and pre-processing of data consists of:

- Load in MAGx_LR_1B files for a given satellite using a sub-sampling of 15s.
- Reject measurements from known disturbed days e.g. when satellite manoeuvres took place
- Reject gross outliers departing more than 500 nT from the latest CHAOS model
- Reject data having Flags_B=255 and Flags_q=255
- Produce two data sets: data chain a) having no data selection and data chain b): applying the selection criteria listed in Section 3

4.2 Method of GVO Estimation

The magnetic field measurements are assumed to be made in a source free region such that the residual magnetic field is a Laplacian potential field, which fulfils the quasi-stationary approximation. In the following we will use the general notation $\delta\mathbf{B}$ for the residual fields of eqs.(1),(2),(3) and refer to the position of the Geomagnetic Virtual Observatory as the "target location".

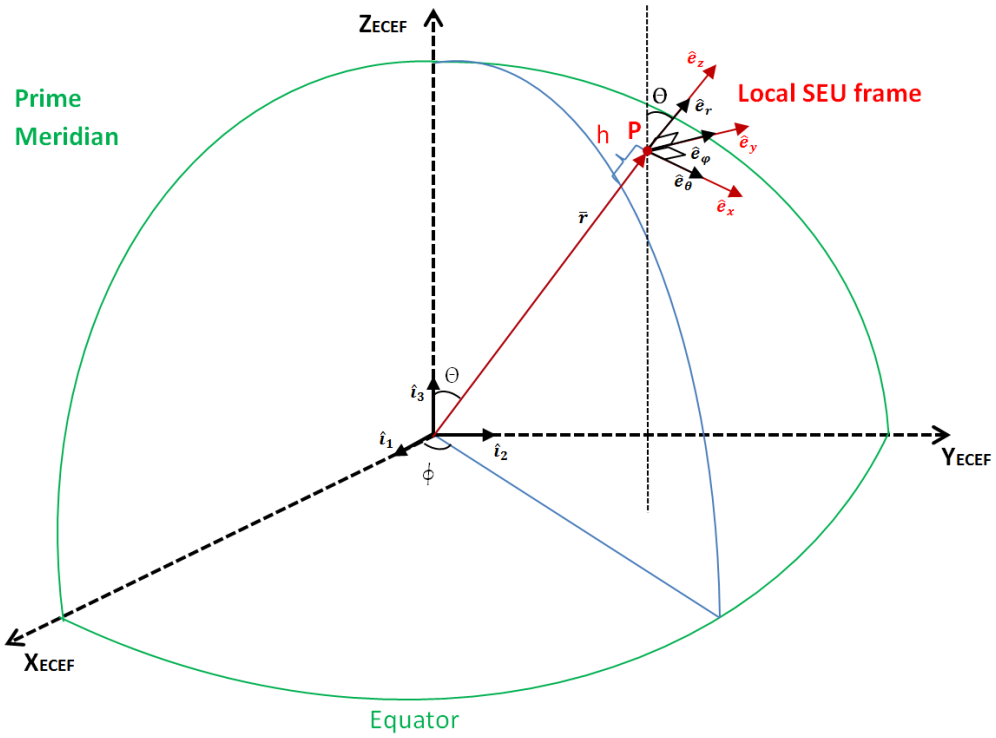


Figure 3: Illustration of the geocentric coordinate systems and the local SEU topocentric Cartesian coordinate system (in red).

The residual magnetic field (including its sums and differences) and the associated locations within a specific target cylinder are transformed from the spherical coordinate system to a right-handed local topocentric Cartesian frame connected to (and constant within) each target cylinder, where x points towards geographic south, y points towards east and z points upwards, here termed the SEU (South,East,Up) frame. Figure 3 illustrates the geocentric spherical and local topocentric frames. Note that the unit vectors of the local Cartesian frame, $(\hat{\mathbf{e}}_\mathbf{x}, \hat{\mathbf{e}}_\mathbf{y}, \hat{\mathbf{e}}_\mathbf{z})$, coincide with the spherical unit vectors $(\hat{\mathbf{e}}_\mathbf{r}, \hat{\mathbf{e}}_\theta, \hat{\mathbf{e}}_\phi)$ at the target location.

The magnetic scalar potential, V, associated with the residual magnetic field in a source-free region must satisfy Laplace's equation $\nabla^2 V = 0$. The solution to Laplace's equation in Cartesian

coordinates can be written as a sum of Harmonic polynomials [e.g. Backus et al., 1996]

$$\begin{aligned} V(x,y,z) &= \sum_{l=1}^{L} C_{abc} x^a y^b z^c \\ &= C_{100}x + C_{010}y + C_{001}z + C_{200}x^2 + C_{020}y^2 + C_{002}z^2 \\ &+ C_{110}xy + C_{101}xz + C_{011}yz + C_{300}x^3 + C_{030}y^3 + C_{003}z^3 \\ &+ C_{210}x^2y + C_{201}x^2z + C_{120}y^2x + C_{021}y^2z + C_{102}z^2x + C_{012}z^2y \\ &+ C_{111}xyz + \cdots , \end{aligned} \tag{4}$$

where $l = a + b + c$, and C_{abc} are the expansion coefficients, a, b, c are non-negative integers, and L is the expansion order. Here, following previous notation for GVOs, we rename the expansion coefficients as ν and use a subscript to indicate the associated polynomial. For this product, with a GVO cylinder radius of 700 km, it was found sufficient to expand the magnetic scalar potential to cubic order $L = 3$. With independent linear spatial dependencies in the Cartesian coordinates x, y, z, the cubic potential can then be written using 19 parameters

$$\begin{aligned} V(x,y,z) &= C_{100}x + C_{010}y + C_{001}z + C_{200}x^2 + C_{020}y^2 + C_{002}z^2 + C_{110}xy \\ &+ C_{101}xz + C_{011}yz + C_{300}x^3 + C_{030}y^3 + C_{003}z^3 + C_{210}x^2y \\ &+ C_{201}x^2z + C_{120}y^2x + C_{021}y^2z + C_{102}z^2x + C_{012}z^2y + C_{111}xyz. \end{aligned} \tag{5}$$

To be a valid potential field two additional criteria must apply: 1) $\nabla \cdot \delta\mathbf{B} = 0$ and 2) $\nabla \times \delta\mathbf{B} = 0$.

1. $\nabla \cdot \delta\mathbf{B} = 0$ criterion
   This requires that

$$\begin{aligned} 0 &= \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} \\ &= -2C_{200} - 6C_{300}x - 2C_{210}y - 2C_{201}z - 2C_{020} - 6C_{030}y \\ &- 2C_{120}x - 2C_{021}z - 2C_{002} - 6C_{003}z - 2C_{102}x - 2C_{012}y \\ &= -(C_{200} + C_{020} + C_{002}) - x(3C_{300} + C_{120} + C_{102}) \\ &- y(3C_{030} + C_{210} + C_{012}) - z(3C_{003} + C_{201} + C_{021}). \end{aligned} \tag{6}$$

   Each term in the parenthesis of the last line must equal zero in order for this to be true. This means that

$$\begin{aligned} C_{002} &= -(C_{200} + C_{020}), \qquad C_{300} = -\frac{1}{3}(C_{102} + C_{120}) \\ C_{030} &= -\frac{1}{3}(C_{210} + C_{012}), \qquad C_{003} = -\frac{1}{3}(C_{201} + C_{021}). \end{aligned}$$

   The cubic potential series is thereby reduced by 4 parameters to 15 parameters

$$\begin{aligned} V(x,y,z) &= C_{100}x + C_{010}y + C_{001}z + C_{200}x^2 + C_{020}y^2 \\ &- (C_{200} + C_{020})z^2 + C_{110}xy + C_{101}xz + C_{011}yz \\ &- \frac{1}{3}(C_{102} + C_{120})x^3 - \frac{1}{3}(C_{210} + C_{012})y^3 - \frac{1}{3}(C_{201} + C_{021})z^3 \\ &+ C_{210}x^2y + C_{201}x^2z + C_{120}y^2x + C_{021}y^2z + C_{102}z^2x \\ &+ C_{012}z^2y + C_{111}xyz. \end{aligned} \tag{7}$$

2. $\nabla \times \delta\mathbf{B} = 0$ criterion
   This criterion requires that

$$\nabla \times \delta\mathbf{B} = \begin{pmatrix} \hat{\mathbf{e}}_\mathbf{x} & \hat{\mathbf{e}}_\mathbf{y} & \hat{\mathbf{e}}_\mathbf{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ B_x & B_y & B_z \end{pmatrix} = 0$$

$$\nabla \times \delta\mathbf{B} = \hat{\mathbf{e}}_\mathbf{x}\left[\frac{\partial}{\partial y}\nabla V_z - \frac{\partial}{\partial z}\nabla V_y\right] - \hat{\mathbf{e}}_\mathbf{y}\left[\frac{\partial}{\partial x}\nabla V_z - \frac{\partial}{\partial z}\nabla V_x\right] + \hat{\mathbf{e}}_\mathbf{z}\left[\frac{\partial}{\partial x}\nabla V_y - \frac{\partial}{\partial y}\nabla V_x\right] = 0. \tag{8}$$

   The potential eq.(7) fulfills this criterion.

In deriving GVO estimates we apply this expression for the potential to estimate a model of the residual magnetic field $\delta\mathbf{B}$. Using the potential eq.(7), centered at a given GVO target position, the residual magnetic field, $\delta\mathbf{B} = (\delta X, \delta Y, \delta Z)$ at positions $r_1, r_2, ..., r_n$, can be linked to the potential coefficients $\mathbf{m} = [C_{100}, C_{010}, C_{001}, C_{200}, C_{020}, C_{110}, C_{101}, C_{011}, C_{210}, C_{201}, C_{120}, C_{021}, C_{102}, C_{012}, C_{111}]^T$, via the spatial derivatives of the potential $\delta\mathbf{B} = -\nabla V$ such that we have the following linear system

$$\begin{pmatrix} \delta X(r_1) \\ \vdots \\ \delta X(r_n) \\ \delta Y(r_1) \\ \vdots \\ \delta Y(r_n) \\ \delta Z(r_1) \\ \vdots \\ \delta Z(r_n) \end{pmatrix} = \begin{bmatrix} -1 & 0 & 0 & -2x_1 & 0 & -y_1 & -z_1 & 0 & -2x_1y_1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots \\ -1 & 0 & 0 & -2x_n & 0 & -y_n & -z_n & 0 & -2x_ny_n & \dots \\ 0 & -1 & 0 & 0 & -2y_1 & -x_1 & 0 & -z_1 & y_1^2 - x_1^2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots \\ 0 & -1 & 0 & 0 & -2y_n & -x_n & 0 & -z_n & y_n^2 - x_n^2 & \dots \\ 0 & 0 & -1 & 2z_1 & 2z_1 & 0 & -x_1 & -y_1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots \\ 0 & 0 & -1 & 2z_n & 2z_n & 0 & -2x_n & -y_n & 0 & \dots \end{bmatrix}$$

$$\begin{bmatrix} \dots & -2x_1z_1 & x_1^2 - y_1^2 & 0 & x_1^2 - z_1^2 & 0 & -y_1z_1 \\ \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & -2x_nz_n & x_n^2 - y_n^2 & 0 & x_n^2 - z_n^2 & 0 & -y_nz_n \\ \dots & 0 & -2x_1y_1 & -2y_1z_1 & 0 & y_1^2 - z_1^2 & -x_1z_1 \\ \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & 0 & -2x_ny_n & -2y_nz_n & 0 & y_n^2 - z_n^2 & -x_nz_n \\ \dots & z_1^2 - x_1^2 & 0 & z_1^2 - y_1^2 & -2x_1z_1 & -2y_1z_1 & y_1y_1 \\ \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & z_n^2 - x_n^2 & 0 & z_n^2 - y_n^2 & -2x_nz_n & -2y_nz_n & y_ny_n \end{bmatrix} \begin{pmatrix} C_{100} \\ C_{010} \\ C_{001} \\ C_{200} \\ C_{020} \\ C_{110} \\ C_{101} \\ C_{011} \\ C_{210} \\ C_{201} \\ C_{120} \\ C_{021} \\ C_{102} \\ C_{012} \\ C_{111} \end{pmatrix}.$$

This can be stated more succinctly in the form of a classical forward problem

$$\mathbf{d}^{vec} = \underline{\underline{\mathbf{G}}}^{vec}\mathbf{m}, \tag{9}$$

which links the model vector of GVO potential coefficients $\mathbf{m}$ with a predicted data vector $\mathbf{d}^{vec}$ that contains the residual field vector components in the SEU system, $\delta\mathbf{B} = (\delta X, \delta Y, \delta Z)$ at the positions of the satellite observations, and $\underline{\underline{\mathbf{G}}}^{vec}$ is a design matrix of the form given above.

In order to find the GVO estimates at the target point, rather than directly using the residual field as data we instead used sums and differences of the vector components of residual magnetic field such that the data vector is $\mathbf{d} = \left[\Delta d_x^{vec}, \Delta d_y^{vec}, \Delta d_z^{vec}, \Sigma d_x^{vec}, \Sigma d_y^{vec}, \Sigma d_z^{vec}\right]^T$, where Δ and Σ denotes the differences and sums of the computed residual field as described in Section 3. The relevant design matrix linking the sums and differences to the model parameters is constructed as $\underline{\underline{\mathbf{G}}} = \left[\Delta G_x^{vec}; \Delta G_y^{vec}; \Delta G_z^{vec}; \Sigma G_x^{vec}; \Sigma G_y^{vec}; \Sigma G_z^{vec}\right]$ where $\Delta G_k^{vec} = [G_k^{vec}(\mathbf{r}_1) - G_k^{vec}(\mathbf{r}_2)]$ and $\Sigma G_k^{vec} = [G_k^{vec}(\mathbf{r}_1) + G_k^{vec}(\mathbf{r}_2)]/2$ where $k = (x, y, z)$.

Using this definition of $\mathbf{d}$ and $\underline{\underline{\mathbf{G}}}$ involving the sums and differences of the vector data, the GVO model parameters are then determined using a robust least-squares inversion solution

$$\mathbf{m} = (\underline{\underline{\mathbf{G}}}^T \underline{\underline{\mathbf{W}}}\, \underline{\underline{\mathbf{G}}})^{-1} \underline{\underline{\mathbf{G}}}^T \underline{\underline{\mathbf{W}}} \mathbf{d}, \tag{10}$$

where $\underline{\underline{\mathbf{W}}}$ is a diagonal weight matrix, consisting of robust (Huber) weights for each entry in the data vector [e.g. Constable, 1988, and equation (15) below], and an additional down-weighting factor of 1/2 for satellites Alpha and Charlie that accounts for the fact that these two satellites fly side-by-side so do not provide completely independent measurements.

In practise, we implement the weighted least squares inversion using the Matlab *lscov* tool https://se.mathworks.com/help/matlab/ref/lscov.html with the robust weights determined using Matlab's robust multilinear regression tool *robustfit*, https://se.mathworks.com/help/stats/robustfit.html. A limit of 30 data points are used as a minimum number of data required for the inversion. If below this limit no solution was computed.

Using the determined potential, an estimate for residual magnetic field for the GVO target point position $(x = 0, y = 0, z = 0)$ which is a time-average over the considered time window, is then computed as

$$\delta \mathbf{B}_{GVO}(x, y, z) = -\nabla V(0, 0, 0) = -\begin{pmatrix} C_{100} \\ C_{010} \\ C_{001} \end{pmatrix}. \tag{11}$$

At the GVO target point the local Cartesian field components can be simply rotated to the spherical components as $\delta B_{GVO,r} = \delta B_{GVO,z}$, $\delta B_{GVO,\theta} = \delta B_{GVO,x}$ and $\delta B_{GVO,\phi} = \delta B_{GVO,y}$. The above procedure is then repeated for each time window at each target location to obtain time series of estimates of the residual vector field at the GVO target locations.

The final step in the GVO field estimation involves adding back the main field model prediction for the target point and epoch under consideration i.e. adding back $\mathbf{B}_{GVO}^{MF}(t)$, for SH degrees $n \in [1, 13]$. This step is carried out separately for each time window such that we finally obtain the GVO vector field time series

$$\mathbf{B}_{GVO}(t) = \delta \mathbf{B}_{GVO}(t) + \mathbf{B}_{GVO}^{MF}(t). \tag{12}$$

The estimated GVO magnetic field are thus provided in spherical polar (r, θ, ϕ) vector components (i.e. $(-C, -N, E)$ components in North-East-Center, NEC frame).

4.3 GVOs Global Grid Description

The GVO time series are provided in a global approximately equal area grid based on the sphere partitioning algorithm of Leopardi [2006]. Selecting a number of GVO grid points, with an associated target cylinder search radius r_{cyl} chosen such as to avoid target cylinder overlaps, involves a trade-off; decreasing the number of target points and increasing the search radius allows for more data within each GVO determination but at the same time lowers the spatial resolution.

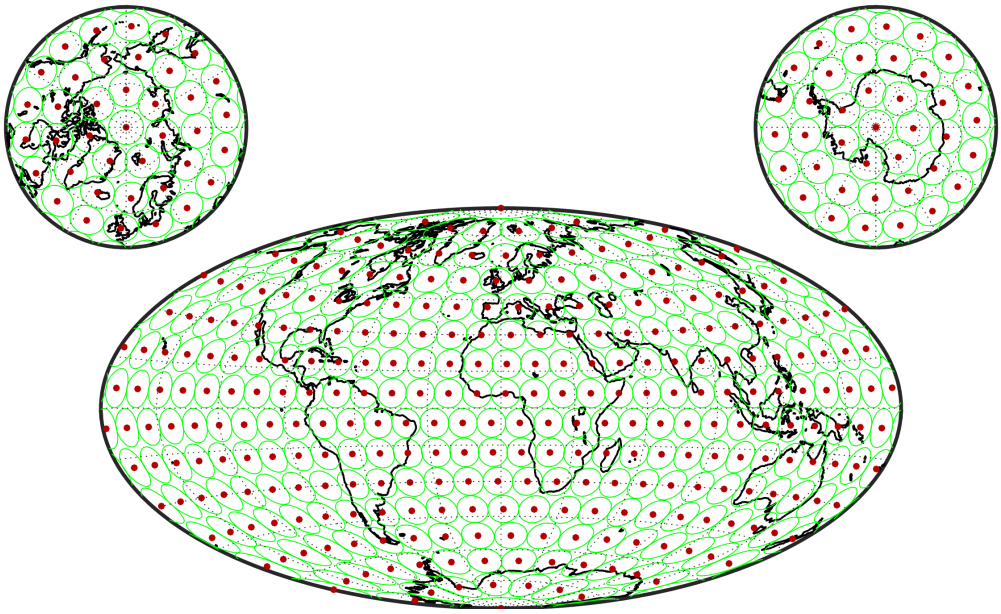


Figure 4: Distribution of the 300 VO's (red dots) and associated cylinder bins (in green) using a Hammer projection.

Preliminary tests with *Swarm* data [Hammer, 2018] suggested that 300 GVO grid locations represented a suitable balance. If higher spatial resolution is required, longer time windows than used here are necessary in order to obtain stable GVO estimates. The surface area dS covered by each GVO target cylinder is the total surface area A divided by the number of GVO's $N_{GVO} = 300$, i.e. $dS \sim A/N_{GVO} = 4\pi r_a^2/N_{GVO}$, where r_a is the Earth's mean radius. Equating this area with the area of a circle surrounding the GVO, πr_{cyl}^2, gives a target cylinder search radius of $r_{cyl} = \sqrt{4r_a^2/N} \approx 700$km. Thus the distance between any two GVO's are ≈ 1500km. This corresponds roughly to SH degree $n = 14$ at a spherical radius r of $(490 + 6371) = 6861$ km, since the SH degree n is associated a horizontal wavelength $\lambda_n \sim 2\pi r/\sqrt{n(n+1)}$). In the *Swarm* DISC product described here the selected target cylinder search radius 700km is half the inter-grid distance d. Figure 4 illustrates the locations of the 300 globally distributed GVO's and the footprint of the data target cylinders for each GVO. The grid used to generate the GVO target points (see Section 4.3), contains GVOs at the North and South Poles. At these positions the (r, θ, ϕ) frame is defined by letting θ be towards the Greenwich meridian, r points upwards and ϕ completes the right-handed coordinate system. When adding back the main field at the North/South Pole, the average of the main field values evaluated 0.1° in latitude from the North/South Pole at longitudes 0° and 180° are used.

4.4 Observed Field GVOs

Classification: *the Observed Field GVOs consist of locally derived field time series, provided on a global grid, computed from a potential fit to satellite observations including all sources of the geomagnetic field.*

The Observed Field GVO time series are derived from the sums and differences of the residual field computed using eq.(1) and then GVO estimation using the algorithm given in eqns.(10)-(12). Two observed field GVO time series are produced:

1) One-monthly observed field GVOs are computed from data chain a) using eqns.(10)-(12)

2) Four-monthly observed field GVOs are computed from data chain b) using eqns.(10)-(12)

4.4.1 Observed Field GVOs Error Estimates

The error estimates σ_{obs} for the Observed Field GVO time series are assumed to be time-independent and spatially uncorrelated. They are calculated separately for each GVO times series i.e. for each field component (r, θ, ϕ) at each GVO location. These are based on a robust total mean square error [e.g. Bendat and Piersol, 2010] between the input data d_i and the GVO predictions d_i^{pre} for a given series

$$\sigma_{obs} = \sqrt{\frac{\sum_i w_i (e_i - \mu_w)^2}{\sum_i w_i} + \mu_w^2}, \tag{13}$$

where the index i runs over all data contributing to a given series, $e_i = d_i - d_i^{pre}$ are residuals and

$$\mu_w = \frac{\sum_i w_i e_i}{\sum_i w_i}, \tag{14}$$

is the robust residual mean and and the weights w_i are calculated iteratively assuming a long-tailed Huber distribution [Constable, 1988]

$$w_i = \begin{cases} 1 & \text{if } \epsilon_i \leq c_w \\ c_w/\epsilon_i & \text{if } \epsilon_i > c_w, \end{cases} \tag{15}$$

with $\epsilon_i = abs(e_i)/std(e)$ and $c_w = 1.5$ is the chosen breakpoint for the Huber distribution.

4.4.2 Observed Field GVOs Processing Overview

The processing flow of the algorithm for computing Observed Field GVO time series consists of :

- Load in MAGx_LR_1B files for a given satellite using a sub-sampling of 15s.
- Measurements from known disturbed days e.g. when satellite manoeuvres took place are rejected.
- Check input data quality and reject suspicious values (departures of more than 500 nT from the latest version of the CHAOS model)
- Reject data having Flags_B=255 and Flags_q=255, resulting in data chain a)
- Compute predictions for the main field (using a main model) at data locations and times
- Produce one and four monthly data residual sets for the GVO location of interest from chain a) and using eq.(1)
- Derive one-monthly and four-monthly GVO estimates using eqns.(10)-(12)
- Derive error estimates for each GVO time series, treating each component separately, using (13)
- Repeat for all locations in the global GVO grid described in subsection 4.3

Product file name	Product description
VOBS_1M_2_	Contains one-monthly time series of the observed field in a global grid of GVOs
VOBS_4M_2_	Contains four-monthly time series of the observed field in a global grid of GVOs

Table 4: Output files collecting the Observed Field GVO data produced using the algorithm described in Sub-section 4.4.

4.4.3 Observed Field GVOs Output Files

Table 4 presents an overview of the output data files containing the Observed Field GVOs. The files VOBS_1M_2_ and VOBS_4M_2_ collect all the Swarm DISC product data files, see Table 7.

4.5 Core Field GVOs and Secular Variation

Classification: *the Core Field GVOs consist of global time series computed from GVOs with estimates of non-core fields removed as far as possible.*

This subsection describes the algorithm for computing the core field and secular variation (SV) GVO time series. One- and four-monthly GVO data files are first produced, after which the one-monthly GVOs are de-noised by a Principal Component Analysis. Next, an epoch-by-epoch spherical harmonic analysis is used to obtain estimates of the external and toroidal magnetic fields (i.e. non-internal parts) which are then removed in order obtain the core field GVO time series. Annual differences are then be computed in order to obtain the GVO core field secular variation time series.

The GVO field time series underlying these products are the following two internal data files (PIDs):

1) PID_1M GVO time series: one-monthly GVOs computed from sums and differences of the field residuals using eq.(2) based on data chain a) (i.e. without data selection criteria)

2) PID_4M GVO time series: four-monthly GVOs are computed from sums and differences of the field residuals based on data chain b) (i.e. with dark geomagnetically quiet-time criteria applied)

The data file PID_1M serves as input in Sub-section 4.5.1 while the file PID_4M serves as input in Sub-section 4.5.2.

4.5.1 Principal Component Analysis of GVOs

Here we describe the algorithm for performing a Principal Component Analysis (PCA) on the input PID_1M GVO time series which serves as an intermediate step towards computing the core field GVO time series in Subsection 4.5.2. This method and its implementation in MagPySV, the open-source Python package for de-noising magnetic data used to do the processing here, is described in Cox et al. [2018]. It is based on earlier work by Wardinski and Holme [2011] and Brown et al. [2013], who used the PCA method to de-noise ground observatory data one observatory at a time, rather than de-noising series from several locations simultaneously, as in this work.

PCA method

The PCA method implemented here works with the residuals between "observed" GVO Secular Variation (SV) and that predicted by an internal magnetic field model. The key premise is that the residuals provide information about contaminating signals, such as magnetospheric and ionospheric magnetic fields and local time sampling biases, that are present in the data but not the internal model, and that PCA of the residuals covariance matrix leads to a proxy for these contaminating signals that can be removed from the data. We hereafter refer to these contaminating signals as "noise" with respect to the core field that is the focus of attention here. Here, we approximate the SV using annual differences of the GVO magnetic field time series, $\boldsymbol{B}_{GVO}(t)$, such that the difference between a sample at time t_n and the sample from the same month of the previous year (t_{n-12}) gives the SV six months between the two dates

$$\frac{\partial \boldsymbol{B}_{GVO}}{\partial t}(t_{n-6}) = \boldsymbol{B}_{GVO}(t_n) - \boldsymbol{B}_{GVO}(t_{n-12}). \tag{16}$$

The SV residuals are then calculated as the difference between the input GVO SV and that predicted by the internal field model evaluated up to spherical harmonic degree 13 at the same times and locations,

$$\delta\dot{\boldsymbol{B}} = \frac{\partial \boldsymbol{B}_{GVO}}{\partial t} - \frac{\partial \boldsymbol{B}_{MF}}{\partial t}. \tag{17}$$

To calculate these residuals, $\partial \boldsymbol{B}_{MF}/\partial t$ is calculated from a smoothly time-varying internal field model e.g. CHAOS-7, [Finlay et al., 2019] or the comprehensive-chain Swarm core field model MCO_SHA_2C [Sabaka et al., 2018]. In this project, the CHAOS-7.2 model was used. Coherent signals in the $\delta\dot{B}_r$, $\delta\dot{B}_\theta$ and $\delta\dot{B}_\phi$ SV residuals at M GVOs (note that not all GVOs are considered together, they are considered in groups according to their magnetic latitude, see below). The SV residuals, $\delta\dot{\boldsymbol{B}}$, are collected into a data matrix for all GVOs being considered,

$$\underline{\underline{\boldsymbol{D}}} = \left[\delta\dot{\boldsymbol{B}}_1 \quad \cdots \quad \delta\dot{\boldsymbol{B}}_M\right], \tag{18}$$

which can be described by a $3M \times 3M$ covariance matrix (assumed constant through time),

$$\underline{\underline{\boldsymbol{C}}}_D = \frac{\underline{\underline{\boldsymbol{D}}}^T \underline{\underline{\boldsymbol{D}}}}{3M}, \tag{19}$$

which can be decomposed into $3M$ eigenvectors ($\underline{\underline{\boldsymbol{E}}}$) and their associated eigenvalues. The eigenvectors are used to rotate the residuals into directions contributing the most (and the least) to the residuals by,

$$\underline{\underline{\boldsymbol{P}}} = \underline{\underline{\boldsymbol{D}}}\,\underline{\underline{\boldsymbol{E}}}. \tag{20}$$

Data projected into each eigendirection ($\underline{\underline{\boldsymbol{P}}}$, called a principal component, PC) is a linear combination of the original variables (i.e. the $\delta\dot{B}_r$, $\delta\dot{B}_\theta$ and $\delta\dot{B}_\phi$ SV residuals at all M GVOs considered). The largest eigenvalue corresponds to the PC with the largest contribution to the residuals (i.e. the unmodelled signals), and the smallest eigenvalue corresponds to the PC with the lowest contribution to the residuals. Depending on the eigenvalue spectrum, and the content of residuals projected into each PC, a number of PCs may be identified as proxy signals for magnetospheric and ionospheric fields and local time sampling biases. Care must be taken to correctly identify the source of each PC before deciding it should be removed, because signals of internal origin that are too rapid to be captured by slowly varying core field models may also appear in the residuals.

Once PCs representing noise sources have been identified, we can remove the associated column vectors to get a modified $\underline{\underline{\widehat{P}}}$ and $\underline{\underline{\widehat{E}}}$, and use these modified matrices to rotate our remaining PCs back into a de-noised data residual matrix with,

$$\underline{\underline{\widehat{D}}} = \underline{\underline{\widehat{P}}}\,\underline{\underline{\widehat{E}}}^{T}. \tag{21}$$

The SV data are then re-formed by adding back the internal field model SV values to the de-noised residuals. This yields SV series with reduced external field and sampling bias contamination, with noise having been removed from each SV component in varying amounts, based on how much it contributed to the K dominant PCs that were removed.

For ground magnetic data, Cox et al. [2018] showed that considering several observatories simultaneously in the PCA permits easy characterisation of external signals at groups of nearby observatories, allowing the user to see patterns in the noise. Also, PCA de-noising is most effective when considering observatories at similar magnetic latitudes because they experience similar external field noise sources at the same time, which shows up clearly in the dominant PCs. On that basis, we used the AACGM-v2 Python package [Shepherd, 2014, Burrell et al., 2020] to calculate the magnetic latitude of each GVO and split the 300 GVO locations into five regions of magnetic latitude (given in Table 5): Polar North, Polar South, Auroral North, Auroral South and Low-mid-magnetic latitudes. Note that the AACGM-v2 model is undefined for some locations at low geographic latitudes; as these locations are all at low magnetic latitude, we assigned these GVOs to the 'low-mid-' magnetic latitude region. Defining these regions was an iterative process; since the dominant PC (or PCs) should be consistent for all GVOs considered in the same analysis, we ran the PCA using approximate boundaries, examined the geographic contributions of the dominant PCs to ensure that they were consistent across all locations, and then refined the boundaries if any GVOs appeared to belong to a different magnetic region.

We established the content of the dominant PCs using their geographic signature, correlations to annual differences of various external magnetic field proxies at the same cadence (e.g. Dst, Polar Cap North/South, Em, AE) and the Discrete Fourier Transforms (DFTs) of projected residuals.

We selected as noise sources the first K PCs from each region based on how many PCs we could confidently identify as either external magnetic fields or a local time sampling bias (see Table 5). For one monthly GVOs in each of the five regions, Table 5 lists the magnetic latitude boundaries, number of PCs removed during de-noising, the percentage variance of the SV residuals that was accounted for in each removed PC and the total percentage variance removed during PCA de-noising. Note that we attempted PCA on the four-monthly GVOs (internal file PID_4M) but these time series contain more data gaps, particularly at higher latitudes, and contained little identifiable external signal due to the data selection criteria. By design, the 4 month local time sampling bias is not present in these series. Therefore, we found little benefit to performing PCA denoising on the four-monthly GVOs and so the de-noising method is applied only to the one-monthly GVO times series (internal file PID_1M).

Finally, the de-noised SV are numerically integrated to produce de-noised magnetic field $\widehat{\mathbf{B}}_{corr}$ time series, again using annual differences such that

$$\widehat{\mathbf{B}}_{corr}(t_{n+12}) = \mathbf{B}_{GVO}(t_n) + \frac{\partial \widehat{\mathbf{B}}_{GVO}}{\partial t}(t_{n+6}). \tag{22}$$

Note that in order to begin this calculation, $\widehat{\mathbf{B}}_{corr}$ must be levelled by providing an initial twelve time samples of field. Here, for $n = 1, ..., 12$, we use $\widehat{\mathbf{B}}_{corr}(t_n) = \mathbf{B}_{GVO}(t_n)$, so that the de-noised field series truly begins twelve months after the first value in the original Observed Field series. We investigated using the mean, median, and Huber-weighted mean levels of observed or modelled field

Region	**Magnetic latitude,** λ_m	K	**Variance** k**-th PC** (%)	**Total (%)**
Polar North	$70° \leq \lambda_m < 90°$	4	34.8, 30.1, 14.5, 5.7	85.1%
Auroral North	$50° \leq \lambda_m < 70°$	3	69.2, 7.8, 6.3	83.3%
Low-mid latitudes	$-50° \leq \lambda_m < 50°$	4	70.8, 9.4, 5.1, 3.3	88.6%
Auroral South	$-70° \leq \lambda_m < -50°$	3	62.9, 10.4, 6.0	79.3%
Polar South	$-90° \leq \lambda_m < -70°$	4	31.7, 22.8, 15.0, 13.3	82.8%

Table 5: Magnetic latitude boundaries for each of the five regions de-noised separately using PCA. The number (K) of PCs removed during each analysis, the percentage variance accounted for by each of the K PCs and the total percentage variance removed.

for times $t_{n=1,12}$, and found that there was negligible difference from using the original Observed Field samples. In all of these cases, the de-noised SV can be exactly reproduced from $\widehat{\mathbf{B}}_{corr}$.

4.5.2 Spherical Harmonic Analysis of GVOs

The magnetic field time series produced by the GVO method assumes a potential field description. This implies that no electrical currents within the measurement region exist. However, when estimating a global grid of GVOs this assumption might locally be invalid due to the presence of electrical currents within the ionospheric region where the satellites are flying. This will especially be the case at high latitudes towards the auroral regions, where due to space-time aliasing, non-potential fields can leak into GVO estimates [Olsen and Mandea, 2007].

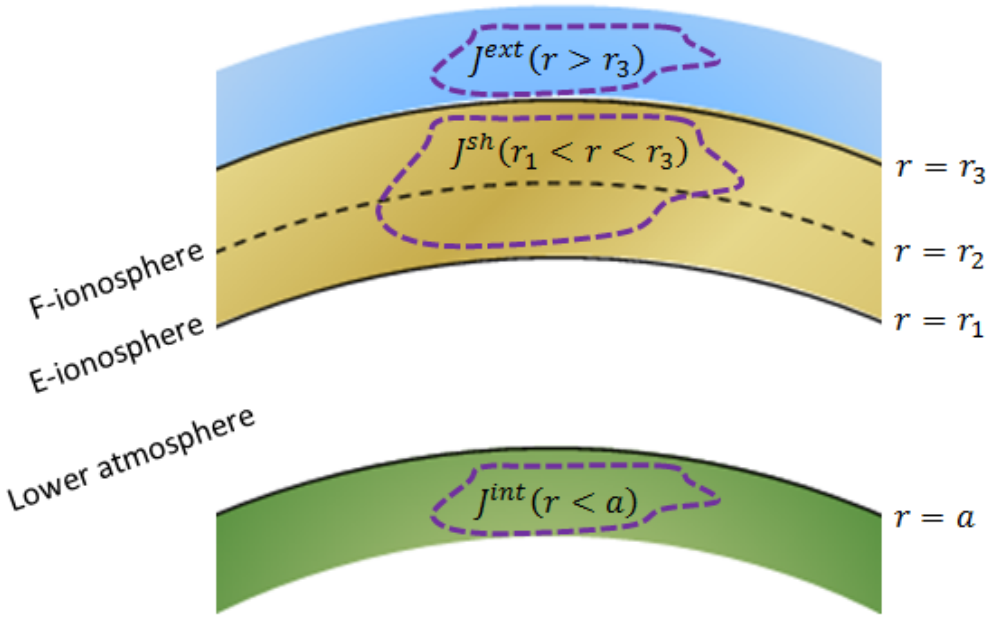


Figure 5: Illustration of the electrical currents and their location in connection to magnetic field observations. After Olsen [1997] and Sabaka et al. [2010].

The situation dealt with in this project is depicted in Figure 5 showing how the satellite magnetic measurements are assumed to be taken in a shell, $S(r_1, r_3)$, situated in the ionospheric F-region at altitudes where in-situ ionospheric electrical currents, J^{sh}, may be present especially at high latitudes [Olsen, 1997, Sabaka et al., 2010]. Internal electrical currents, J^{int}, and external currents,

J^{ext}, are below and above the measurements shell, respectively.

We assume that the electrical currents in the F-region are poloidal in nature (more specifically purely radial), thus they produce toroidal magnetic fields within the shell, $S(r_1, r_3)$. In this situation a non-potential field formulation, called the Mie-representation, which deviates from that of the usual potential field description should to be considered [Backus et al., 1996, Olsen, 1997, Sabaka et al., 2010]. It involves the magnetic field in the current carrying (sampling) shell $S(r_1, r_3)$ being written as [Backus et al., 1996, Olsen, 1997, Sabaka et al., 2010]

$$\mathbf{B} = \mathbf{B}^{int}_{pol} + \mathbf{B}^{ext}_{pol} + \mathbf{B}^{sh}_{pol} + \mathbf{B}^{sh}_{tor}, \tag{23}$$

where the superscripts int, ext, sh denotes the internal, external and shell parts, respectively. The subscripts pol and tor denote the poloidal and toroidal parts, respectively. $\mathbf{B}^{int}_{pol}$ and $\mathbf{B}^{ext}_{pol}$ denotes the internal and external poloidal fields, respectively, while $\mathbf{B}^{sh}_{pol}$ and $\mathbf{B}^{sh}_{tor}$ denote the non-potential fields due to currents in the shell. The internal, $\mathbf{B}^{int}_{pol}$, and external, $\mathbf{B}^{ext}_{pol}$, magnetic fields are generated by internal, $\mathbf{J}^{int}_{tor}$, and external, $\mathbf{J}^{ext}_{tor}$, toroidal currents, respectively.

Here, we assume the ionospheric F-region field contribution is due to radial currents with no toroidal component, hence the poloidal field due to currents in the shell is neglected [Backus et al., 1996, Olsen, 1997, Sabaka et al., 2010]. Furthermore a thin-shell approximation is used whereby $h/(r_2 \to 0)$, where $h = r_3 - r_1$ is the thickness of the shell, is considered to be thin as compared to the radius, r_3. In this case the expression for the magnetic field further simplifies to [Olsen and Mandea, 2007]

$$\mathbf{B} = \mathbf{B}^{int}_{pol} + \mathbf{B}^{ext}_{pol} + \mathbf{B}^{sh}_{tor}. \tag{24}$$

It can be shown that these assumptions and knowledge of the vector field on the surface at $r = r_2$, allows for the determination of spherical harmonic expansion coefficients (i.e. the Gauss coefficients along with the toroidal coefficients). In this case the magnetic field is then written in terms of poloidal, V^{int}, V^{ext} and toroidal, T^{sh}, scalar potentials [Olsen and Mandea, 2007]

$$\mathbf{B} = -\nabla V^{int} - \nabla V^{ext} + \nabla \times \hat{r} T^{sh}, \tag{25}$$

where each of the potentials are expanded up to some maximum SH degree N can be written by the expansions

$$V^{int} = r_a \sum_{n=1}^{N} \sum_{m=0}^{n} \left[g_n^m(t) \mathrm{cos} m\phi + h_n^m(t) \mathrm{sin} m\phi \right] \left(\frac{r_a}{r} \right)^{n+1} P_n^m \tag{26}$$

$$V^{ext} = r_a \sum_{n=1}^{N} \sum_{m=0}^{n} \left[q_n^m(t) \mathrm{cos} m\phi + s_n^m(t) \mathrm{sin} m\phi \right] \left(\frac{r_a}{a} \right)^{n} P_n^m \tag{27}$$

$$T^{sh} = r_a \sum_{n=1}^{N} \sum_{m=0}^{n} \left[t_n^{m,c}(t) \mathrm{cos} m\phi + t_n^{m,s}(t) \mathrm{sin} m\phi \right] P_n^m, \tag{28}$$

where $r_a = 6371.2$km is the reference value for the Earth's mean spherical radius, n and m are here the spherical harmonic degree and order, respectively, and P_n^m are the associated Schmidt semi-normalized Legendre functions. In the three expansions, $\{g_n^m, h_n^m\}$ are the internal coefficients, $\{q_n^m, s_n^m\}$ are the external coefficients and $\{t_n^{m,c}, t_n^{m,s}\}$ are the expansion coefficients associated with the toroidal scalar potential.

Predictions for the spherical polar components of the geomagnetic field at the GVO locations are linearly related to the above expansion coefficients such that a forward problem can be

written

$$\mathbf{d}_{GVO} = \underline{\underline{G}}_{SH}\mathbf{m}_{SH}, \tag{29}$$

where the data are given by $\mathbf{d}_{GVO} = [B_r(\mathbf{r}_1), ...B_r(\mathbf{r}_{N_{GVO}}), B_\theta(\mathbf{r}_1), ...B_\theta(\mathbf{r}_{N_{GVO}}), B_\phi(\mathbf{r}_1), ...B_\phi(\mathbf{r}_{N_{GVO}})]^T$, where N_{GVO} is the number GVOs, and these are related to the expansion coefficients $\mathbf{m}_{SH} = [g_n^m, h_n^m, q_n^m, s_n^m, t_n^{m,c}, t_n^{m,s}]^T$ via a design matrix $\underline{\underline{G}}_{SH}$. This is constructed from spatial derivatives of equations (26),(27),(28). The internal, external and toroidal expansion were truncated at SH degree $N = 13$ and the model coefficients were determined epoch by epoch using a simple least squares solution

$$\mathbf{m}_{SH} = (\underline{\underline{\mathbf{G}}}_{SH}^T \underline{\underline{\mathbf{G}}}_{SH})^{-1}\underline{\underline{\mathbf{G}}}_{SH}^T\mathbf{d}_{GVO}. \tag{30}$$

The SH analysis is performed epoch by epoch on the one-monthly PCA de-noised data and separately on the four-monthly GVO time series from which estimates of the lithospheric, magnetospheric and ionospheric fields have been removed (i.e. eq.(3)). At epochs where an insufficient number of GVOs are available to ensure a stable solution, the external and toroidal coefficients are determined by a linear interpolation between nearby epochs.

Following the spherical harmonic analysis the external and toroidal field estimates are removed epoch by epoch from the one- and four monthly time series to produce final Core Field GVO time series. The Secular Variation of the core field at a particular GVO location for a given time window at epoch t, is computed using annual differences between core field values at time $t + 0.5\,\mathrm{yr}$ and at time $t - 0.5\,\mathrm{yr}$

$$\mathbf{SV}_{GVO}(t) = \mathbf{B}_{GVO}(t + 0.5\,yr) - \mathbf{B}_{GVO}(t - 0.5\,yr). \tag{31}$$

4.5.3 Core Field GVOs and Secular Variation Error Estimates

The error estimates σ for the core field GVO time series are also assumed to be time-independent and spatially uncorrelated. These error estimates are computed separately for each field component at each GVO using the residuals $\mathbf{e}$, between the core GVO data (i.e. $\mathbf{d}_{GVO} = \mathbf{B}_{GVO}$) and the corresponding field predictions of the time-dependent internal part of the CHAOS model for SH degree $n \in [1, 20]$, $\mathbf{d}_{CHAOS}$, that is $\mathbf{e} = \mathbf{d}_{GVO} - \mathbf{d}_{CHAOS}$. The error estimates are computed using the total mean squared of the residuals, being the square root of the mean square deviation plus the residual mean squared, i.e.

$$\sigma_{core} = \sqrt{\frac{\sum_i (e_i - \mu)^2}{M} + \mu^2}, \tag{32}$$

where $i = 1, ..., M$ denotes the ith data element, and M is the number of data in a given series and μ is the residual mean, with respect to the CHAOS model time-dependent internal field, for a given spherical polar component of the Core Field GVO time series at a given GVO location.

The error estimates of the Secular Variation GVO time series are computed in a similar manner as described above for the Core Field but using residuals between the SV GVO data, SV_{GVO}, and the SV predictions of the CHAOS model time-dependent internal field model for SH degree $n \in [1, 20]$.

4.5.4 Core Field GVOs and Secular Variation Processing Overview

The algorithm flow for computing the Core Field and Secular Variation GVO time series consists of four main modules:

Module 1: Data selection

Module 2: Compute internal and external field model predictions

Module 3: Derive PID GVO time series for use in Modules 4 and 5

Module 4: Perform PCA on one-monthly PID GVO time series

Module 5: Perform SHA and compute Core Field and Secular Variation GVOs

Module 1: Data selection

The following list describes the single processing steps in the module 1:

- Load in MAGx_LR_1B files for a given satellite using a sub-sampling of 15s.
- Measurements from known disturbed days e.g. when satellite manoeuvres took place are rejected.
- Check input data quality and reject suspicious values (departing from the predictions of the CHAOS field model by more than 500 nT)
- Reject data having Flags_B=255 and Flags_q=255
- Produce two data sets: data chain a) having no data selection and data chain b): applying the selection criteria listed in Section 3

Module 2: Compute internal and external field model predictions

The following list describes the single processing steps in the module 2:

- Vector field predictions of the IGRF main field for SH degrees $n \in [1, 13]$
- Vector field predictions of the LCS-1 lithospheric field for SH degrees $n \in [14, 185]$
- Vector field predictions of the CHAOS large-scale magnetospheric field and its induced field
- Vector field predictions of the CI ionospheric field and its induced field

Module 3: Derive PID GVO time series for use in Modules 4 and 5

The following list describes the single processing steps in the module 3:

- Produce one-monthly data residual sets for all locations in the GVO global grid using the data from chain a) (no data selection) and using eq.(2)
- Derive the one-monthly PID_1M GVO time series for all locations in the global GVO grid using eqns.(10)-(12), this serves as input data for Module 4
- Produce four-monthly data residual set for all locations in the GVO global grid using data from chain b) (with dark quiet-time selection criteria) and using eq.(3)
- Derive the four-monthly PID_4M GVO time series for all locations in the global GVO grid using eqns.(10)-(12), this serves as input data for Module 5

Module 4: Perform PCA on one-monthly PID GVO time series

The processing flow of the algorithm for de-noising the GVO time series by performing Principal Component Analysis consists of four main steps:

1. Read GVO time series files
   - Read the input files PID_1M
   - Reorder the data into time series for each GVO location
2. Compute the residuals between the GVO SV and an internal field model (e.g. CHAOS-7)
   - Calculate the SV of GVO magnetic field time series as annual differences (see eq.(16))
   - Compute the SV prediction using an internal field model for SH degrees $n = 1 - 13$
   - Calculate the SV residuals by subtracting the model SV values from the GVO values
3. Perform PCA on the residuals to remove magnetospheric and ionospheric field contamination and local time sampling biases
   - Perform PCA on the SV residuals according to the algorithm (**??**) and Cox et al. [2018]
   - Use the eigenvalue spectrum of the residuals covariance matrix, correlations with external magnetic field proxies and the Discrete Fourier Transforms (DFTs) of the principal components (PCs) to validate which PCs should be removed in the previous step
4. Re-compute the de-noised magnetic field from the de-noised secular variation time series
   - Compute the de-noised SV by adding the model SV values back on to the de-noised SV residuals
   - Compute the de-noised magnetic field from the de-noised SV according to (22)

Module 5: Perform SHA and compute Core Field and Secular Variation GVOs

The following list describes the single processing steps in the module 4:

- Perform a SHA epoch-by-epoch on the one-monthly PID_1M_PCA_MF time series and on the four-monthly PID_4M time series, determining internal, external and toroidal expansion coefficients
- Compute one- and four monthly Core Field GVO time series by subtracting off external and toroidal field estimates from the PID_1M_PCA_MF and PID_4M data files
- Compute one- and four monthly Secular Variation GVO estimates as annual differences of the Core Field GVOs using eq.(30)
- Compute error estimates for the one-monthly and four-monthly Core Field and Secular Variation time series, considering each time series separately and using (32)

4.5.5 Core Field GVOs and Secular Variation Output Files

Table 6 presents an overview of the output data files collecting the Core Field and Secular Variation GVO data. The internal data files PID_1M serve as input for the Principal Component Anaylysis described in Sub-section 4.5.1. The internal data files PID_4M and PID_1M_PCA_MF serve as inputs in Sub-section 4.5.2. The files VOBS_1M_2_ and VOBS_4M_2_ are overall product data files collected in the Swarm DISC One-monthly and Four-monthly GVO product files, see Table 7.

Product file name	Product description
PID_1M	One-monthly GVO time series having subtracted LCS-1 lithopsheric field estimates
PID_4M	Four-monthly GVO time series having subtracted LCS-1, CHAOS and CI model non-core field estimates
PID_1M_PCA_SV	One-monthly time series of the PCA de-noised secular variation on the global grid of GVOs
PID_1M_PCA_MF	One-monthly time series of the magnetic field calculated from the PCA de-noised SV
VOBS_1M_2_	One-monthly time series of the core field vector and secular variation field vector in a global grid of GVOs
VOBS_4M_2_	Four-monthly time series of the core field vector and secular variation field vector in a global grid of GVOs

Table 6: Product file list for the derived Core Field and Secular Variation GVOs

5 Swarm DISC Product Output Files

The Swarm DISC Geomagnetic Virtual Observatories data products comprise two cdf files summarised in Table 7; one cdf file collecting all the one-monthly GVO time series and one cdf file collecting all the four-monthly GVO time series. Detailed descriptions of the cdf output variables including type, format description, and units can be found in the Swarm Geomagnetic Virtual Observatories Product Definition document, Ref.[AD-2].

Product file name	Product description
VOBS_1M_2_	One-monthly time series of the vector magnetic field: the Observed Field, the Core Field and the Secular Variation, and their associated error estimates, all provided in a global grid of geomagnetic virtual observatories
VOBS_4M_2_	Four-monthly time series of the vector magnetic field: the Observed Field, the Core Field and the Secular Variation, and their associated error estimates, all provided in a global grid of geomagnetic virtual observatories

Table 7: Swarm DISC Product data files summary list.

References

P. Alken, E. Thébault, C. D. Beggan, H. Amit, J. Aubert, J. Baerenzung, T. Bondar, W. Brown, S. Califf, A. Chambodut, A. Chulliat, G. Cox, C. C. Finlay, A. Fournier, N. Gillet, A.Grayver, M. D. Hammer, M. Holschneider, L. Huder, G. Hulot, T. Jager, C. Kloss, M. Korte, W. Kuang, A. Kuvshinov, B. Langlais, J.-M. Léger, V. Lesur, P. W. Livermore, F. J. Lowes, S. Macmillan, W. Magnes, M. Mandea, S. Marsal, J. Matzka, M. C. Metman, T. Minami, A. Morschhauser, J. E. Mound, M. N. , S. Nakano, N. Olsen, F. J. Pavón-Carrasco, V. G. Petrov, G. Ropp, M. Rother, T. Sabaka, S. Sanchez, D. Saturnino, N. R. Schnepf, X. Shen, C. Stolle, A. Tangborn, L. Tøffner-Clausen, H. Toh, J. M. Torta, J. Varner, F. Vervelidou, P. Vigneron, I. Wardinski, J. Wicht, A. Woods, Z. Z. Y. Yang, and B. Zhou. International Geomagnetic Reference Field: the thirteenth generation. *Earth, Planets and Space*, in press, 2020. doi:10.1186/s40623-020-01288-x.

G. Backus, R. Parker, and C. Constable. *Foundations of Geomagnetism*. Cambridge Univ. Press, New York, 1996.

O. Barrois, M. D. Hammer, C. C. Finlay, Y. Martin, and N. Gillet. Assimilation of ground and satellite magnetic measurements: inference of core surface magnetic and velocity field changes. *Geophys. J. Int.*, 215:695–712, 2018.

C. D. Beggan, K. A. Whaler, and S. Macmillan. Biased residuals of core flow models from satellite-derived virtual observatories. *Geophys. J. Int.*, 177(2):463–475, 2009.

J. Bendat and A. Piersol. *Random Data, Analysis and Measurement Procedures*. Wiley, New Jersey, 2010.

W. Brown, J. Mound, and P. Livermore. Jerks abound: An analysis of geomagnetic observatory data from 1957 to 2008. *Physics of the Earth and Planetary Interiors*, 223(0):62 – 76, 2013. ISSN 0031-9201. doi: 10.1016/j.pepi.2013.06.001. URL http://www.sciencedirect.com/science/article/pii/S0031920113000708. SI:13th SEDI conference.

A. Burrell, C. van der Meeren, and K. M. Laundal. aburrell/aacgmv2: Version 2.6.0. Jan 2020. doi: 10.5281/zenodo.3598705.

C. Constable. Parameter estimation in non-Gaussian noise. *Geophys. J. Int.*, 94(1):131–142, 1988.

G. Cox, W. Brown, L. Billingham, and R. Holme. Magpysv: A python package for processing and denoising geomagnetic observatory data. *Geochem. Geophys. Geosyst.*, 19(9):3347–3363, 2018.

J. Domingos, M. A. Pais, D. Jault, and M. Mandea. Temporal resolution of internal magnetic field modes from satellite data. *Earth, Planets and Space*, 71(1):1–17, 2019.

C. Finlay, C. Kloss, N. Olsen, M. Hammer, and L. Tøffner-Clausen. DTU candidate field models for IGRF-13, 2019. URL http://www.spacecenter.dk/files/magnetic-models/CHAOS-7/DTU_IGRF13.pdf.

C. C. Finlay, N. Olsen, S. Kotsiaros, N. Gillet, and L. Tøffner-Clausen. Recent geomagnetic secular variation from Swarm and ground observatories as estimated in the CHAOS-6 geomagnetic field model. *Earth, Planets and Space*, 68(1):1–18, 2016.

M. D. Hammer. *Local Estimation of the Earth's Core Magnetic Field*. PhD thesis, Technical University of Denmark, 2018.

C. Kloss and C. C. Finlay. Time-dependent low latitude core flow and geomagnetic field acceleration pulses. *Geophys. J. Int.*, 217:140–168, 2019.

P. Leopardi. A partition of the unit sphere into regions of equal area and small diameter. *Electronic Transactions on Numerical Analysis*, 25(12):309–327, 2006.

M. Mandea and N. Olsen. A new approach to directly determine the secular variation from magnetic satellite observations. *Geophys. Res. Lett.*, 33(15), 2006.

N. Olsen. Ionospheric F region currents at middle and low latitudes estimated from Magsat data. *J. Geophys. Res.: Space Physics*, 102(A3):4563–4576, 1997.

N. Olsen and M. Mandea. Investigation of a secular variation impulse using satellite data: The 2003 geomagnetic jerk. *Earth Planet. Sci. Lett.*, 255(1):94–105, 2007.

N. Olsen, M. Mandea, T. J. Sabaka, and L. Tøffner-Clausen. CHAOS-2 - a geomagnetic field model derived from one decade of continuous satellite data. *Geophys. J. Int.*

N. Olsen, M. Mandea, T. J. Sabaka, and L. Tøffner-Clausen. The CHAOS-3 geomagnetic field model and candidates for the 11th generation IGRF. *Earth, planets and space*, 62(10):719–727, 2010.

N. Olsen, H. Lühr, C. C. Finlay, T. J. Sabaka, I. Michaelis, J. Rauberg, and L. Tøffner-Clausen. The CHAOS-4 geomagnetic field model. *Geophys. J. Int.*, 197(2):815–827, 2014.

N. Olsen, G. Hulot, V. Lesur, C. C. Finlay, C. Beggan, A. Chulliat, T. J. Sabaka, R. Floberghagen, E. Friis-Christensen, R. Haagmans, et al. The Swarm Initial Field Model for the 2014 geomagnetic field. *Geophys. Res. Lett.*, 42(4):1092–1098, 2015.

N. Olsen, D. Ravat, C. C. Finlay, and L. K. Kother. LCS-1: a high-resolution global model of the lithospheric magnetic field derived from CHAMP and Swarm satellite observations. *Geophys. J. Int.*, 211(3):1461–1477, 2017.

H. F. Rogers, C. D. Beggan, and K. A. Whaler. Investigation of regional variation in core flow models using spherical slepian functions. *Earth, Planets and Space*, 71(1):19, 2019.

T. J. Sabaka, G. Hulot, and N. Olsen. Mathematical properties relevant to geomagnetic field modeling. In *Handbook of Geomathematics*, pages 503–538. Springer, 2010.

T. J. Sabaka, L. Tøffner-Clausen, N. Olsen, and C. C. Finlay. A Comprehensive Model of the Earth's Magnetic Field Determined From 4 Years of Swarm Satellite Observations. *Earth, Planets and Space*, 70(1):130, 2018.

S. G. Shepherd. Altitude-adjusted corrected geomagnetic coordinates: Definition and functional approximations. *J. Geophys. Res.: Space Physics*, 119(9):7501–7521, 2014.

R. M. Shore. *An improved description of Earth's external magnetic fields and their source regions using satellite data*. PhD thesis, The University of Edinburgh, 2013.

I. Wardinski and R. Holme. Signal from noise in geomagnetic field modelling: denoising data for secular variation studies. *Geophys. J. Int.*, 185(2):653–662, 2011.

K. A. Whaler. Probing the core surface flow with satellite data. 2017. IAGA Joint Assembly Session A02 - Earth's core dynamics and planetary dynamos (DIV I).