The Cluster Virtual Observatory for ULF Waves

O.D. Constantinescu[1,2], K.H. Fornaçon[1], U. Motschmann[3], I. Richter[1], K.H. Glassmeier[1]

[1]Institute for Geophysics and Extraterrestrial Physics, TU Braunschweig, Germany
[2]Institute for Space Sciences, Bucharest, Romania
[3]Institute for Theoretical Physics, TU Braunschweig, Germany

Key Points:

- The Cluster Virtual Observatory is an on-line archive of low frequency wave parameters based on Cluster fluxgate magnetometers measurements.
- The Cluster Virtual Observatory also provides the tetrahedron geometrical configuration parameters.
- The waves and geometrical parameters are available as quick-plots as well as binary *Hierarchical Data Format* data.

Corresponding author: O.D. Constantinescu, d.constantinescu@tu-bs.de

Abstract

Since its launch in 2000, the Cluster fleet visited a vast domain of the circum-terrestrial environment, from upstream solar wind and distant tail, down to the plasmasphere, scanning in detail all magnetospheric regions during over two solar cycles. This led to an unprecedented rich data collection of multi-point measurements which will be used for years to come to decipher the mechanisms of Solar-Terrestrial interactions.

The large volume of the data gathered by Cluster requires special strategies to make efficient use of it. To address this issue we constructed a browsable database containing ULF waves and spacecraft formation parameters. Recently, we made this database accessible at `http://plasma.spacescience.ro/cluster.html`. Here we present the newly developed tool, discuss the methods used to derive the parameters and give practical examples.

1 Introduction

The Cluster mission (Escoubet et al., 1997), consisting of four identical spacecraft flying in formation around the Earth, is the first multi-spacecraft mission to study the Earth's magnetosphere and the near Earth solar wind. Simultaneous measurements allowed for the first time to separate spatial from temporal fluctuations and to investigate the three-dimensional structures in the Earth's plasma environment. Each spacecraft caries 11 state-of-the-art instruments to measure the surrounding plasma properties. One of the key quantities is the magnetic field which is measured by two instruments: a searching coil magnetometer (Cornilleau-Wehrlin et al., 1997) to measure high frequency fluctuations of the magnetic field and a flux gate magnetometer (FGM) (Balogh et al., 1997) to measure low frequency fluctuations. The spacecraft were launched in July and August 2000 on polar orbits and they will likely deliver science data at least until 2024 when the first spacecraft in the formation enters the Earth's upper atmosphere.

ULF waves are oscillations of the electromagnetic field occurring in the magnetised plasma around the Earth at frequencies in the order of mHz to Hz, e.g. Glassmeier (1995); Pilipenko (1990); Keiling et al. (2016). They play a crucial role in the transfer and distribution of the energy coming from the Sun by energising particles, triggering reconnection, storing and propagating energy, modifying distribution functions, and carrying information between distant regions of the magnetosphere and the solar wind. Various wave modes are excited, depending on the local plasma parameters as well as on the solar wind conditions, and each wave mode has its particular way of interacting with the magnetosphere.

Estimating the wave parameters is crucial for the wave mode identification and for understanding the part played by the waves in the Solar-Terrestrial interaction. Many of these parameters, such as the power spectral density, coherency, ellipticity, propagation direction can be determined using the magnetic field alone. Other parameters such as the Poynting vector and the phase relation between the magnetic field and the particle density, require additional measurements of other physical quantities.

Until the end of the mission, Cluster will gather the equivalent of over one century of single spacecraft data from each instrument. Managing this amount of data poses challenges in computing, storing, and searching for relevant events. For instance, only obtaining the 2000 to 2018 ULF waves parameters discussed in the following sections required more than two months of continuous computing time on the eight threads of a 3.1 GHz Intel processor. A searchable archive of pre-computed ULF waves parameters has the potential to enable event-based and statistical studies otherwise difficult to conduct.

The multipoint capabilities of the Cluster fleet set it apart from most other spacecraft probing the Earth's magnetosphere and the solar wind. While even merely comparing the measurements from two spacecraft is useful in many investigations, more sophisticated techniques, such as the wave telescope / k-filtering (Motschmann et al., 1996; Pinçon & Motschmann, 1998; Glassmeier et al., 2001) or the curlometer (Dunlop et al., 2002) require specific shapes and sizes of the spacecraft formation, and their accuracy depends on the geometric parameters of the spacecraft tetrahedron. Even though computing these parameters is not as resource-demanding as computing the ULF waves parameters, a tool allowing quick estimation of the Cluster tetrahedron shape and size represents a useful instrument.

The remainder of this work is organized as follows: Section 2 discusses the derivation of the ULF waves parameters and provides the relations used to compute them later. Similarly, section 3 discusses the tetrahedron geometric parameters. Section 4 is dedicated to the Cluster Virtual Observatory and describes in detail the archived data and the capabilities of the online tool. Section 5 summarises this work.

2 Derivation of the ULF waves parameters

Among the fundamental properties of the waves in magnetized plasma are the polarization parameters: the polarization degree, the ellipticity, and the orientation of the variance ellipsoid. The first step in finding these parameters is finding the principal component system of the magnetic field. Though this can be done in various ways, the most common approach is based on the analysis of the spectral matrix obtained from the Fourier components corresponding to the analysed frequency. For the magnetic field $\boldsymbol{B}$ with components B_i, $i = 1, \ldots, 3$, the spectral matrix elements are:

$$\mathcal{S}_{ij}(\omega) = \left\langle \tilde{B}_i(\omega)\tilde{B}_j^{\star}(\omega) \right\rangle \qquad i, j = 1, \ldots, 3 \tag{1}$$

where $\langle \cdots \rangle$ denotes the average and * denotes the complex conjugate. The Fourier component

$$\tilde{B}_i(\omega) = \frac{1}{2\pi} \int B_i(t) e^{-\mathrm{i}\omega t} \mathrm{d}t \tag{2}$$

is approximated by the discrete Fourier transform usually through the Fast Fourier Transform (FFT) method.

Arthur et al. (1976) discuss three methods which can be used to determine the polarization parameters of plasma waves. One of those was proposed by Samson (1973). This method is based on the decomposition of the spectral matrix in three terms: one corresponding to the total polarized part of the wave, one corresponding to the partial polarized part of the wave and one corresponding to the non polarized part of the wave. From this decomposition, the relative powers of the three polarization components of the wave result:

$$P_{\mathrm{T}} = \frac{\lambda_1 - \lambda_2}{\Lambda}; \qquad P_{\mathrm{P}} = 2\frac{\lambda_2 - \lambda_3}{\Lambda}; \qquad P_{\mathrm{N}} = 3\frac{\lambda_3}{\Lambda} \tag{3}$$

where $\lambda_1 \geq \lambda_2 \geq \lambda_3$ are the eigenvalues of the complex spectral matrix $\mathcal{S}$ and $\Lambda = \lambda_1 + \lambda_2 + \lambda_3$. The polarization degree is given by

$$P^2 = \frac{1}{2\Lambda} \sum_{\substack{i,j=1 \\ i<j}}^{3} (\lambda_i - \lambda_j)^2 \tag{4}$$

This method has the advantage of decomposing the data into quantities with clear physical meaning. One drawback which could become significant in the context of processing large quantities of data is the need of solving the eigenvalue problem for a complex matrix, which is more resource demanding than for a real matrix.

Santolík et al. (2003) applies singular value decomposition (SVD) to the complex spectral matrix to derive many wave parameters. The SVD is much easier on resources and in addition the proposed method is generalized to naturally include the electric field vector, providing more information.

Other methods discard either the real or the imaginary part of the spectral matrix. An example, discussed by Arthur et al. (1976) is the direct determination of the direction of the $\boldsymbol{k}$ vector from the imaginary part of the spectral matrix (Means, 1972):

$$\hat{k}_l = \pm \Im\left(\mathcal{S}_{mn}\right)/p \qquad l,m,n = 1,\ldots,3; \quad l \neq n \neq m \tag{5}$$

where $p^2 = \sum_{i,j=1}^{3} \Im\left(\mathcal{S}_{ij}\right)^2$ and the sign is given by the polarization sign.

Once the direction of the wave vector is determined, the spectral matrix is rotated in a coordinate system with the x axis along $\hat{\boldsymbol{k}}$. This is relatively easy to implement numerically and requires less CPU time than needed for the diagonalization of the spectral matrix. However, since only the minimum variance direction is provided by $\hat{\boldsymbol{k}}$, the polarization information in incomplete. Also, purely linear polarized waves cannot be treated because in this case the spectral matrix is real.

The third technique considered by Arthur et al. (1976) was proposed by McPherron et al. (1972) and, in contrast with the Means (1972) method, implies the diagonalization of the real part of the spectral matrix. The rotation matrix $\mathcal{T}$ can be derived from the eigenvectors of the real part of the spectral matrix $\Re(\mathcal{S})$:

$$\mathcal{T} = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3] \qquad \lambda_1 \geq \lambda_2 \geq \lambda_3 \tag{6}$$

The numerical implementation of the Means (1972) method is relatively straightforward and no additional assumptions about the analysed waves are necessary. The spectral matrix rotated into the principal variance system becomes:

$$\mathcal{J} = \mathcal{T}\mathcal{S}\mathcal{T}^{\dagger} \tag{7}$$

where † denotes the Hermitic conjugate (complex conjugate transpose).

The polarization degree is defined as the ratio between the wave coherent power and the total power of the wave (Fowler et al., 1967)

$$P^2 = 1 - 4\frac{\det(\mathcal{J})}{\mathrm{Tr}(\mathcal{J})^2} \tag{8}$$

For a plane wave described by

$$B_x(t) = b_x e^{\mathrm{i}\omega t} \tag{9a}$$

$$B_y(t) = b_y e^{\mathrm{i}\omega t - \pi/2} \tag{9b}$$

$$B_z(t) = b_z \tag{9c}$$

the spectral matrix $\mathcal{J}$ becomes

$$\mathcal{J} = \begin{pmatrix} b_x^2 & \mathrm{i}b_x b_y & 0 \\ -\mathrm{i}b_x b_y & b_y^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{10}$$

The ellipticity is defined as the ratio between the two axes of the polarization ellipsoid:

$$\epsilon = \tan\Big[\frac{1}{2}\arcsin\Big(\frac{2\Im(J_{xy})}{\sqrt{\mathrm{Tr}(\mathcal{J})^2 - 4\det(\mathcal{J})}}\Big)\Big] \tag{11}$$

The orientation of the maximum variance axis is given by the azimuth angle:

$$\tan(2\phi) = \frac{2\Re(J_{xy})}{J_{xx} - J_{yy}} \tag{12}$$

The variance analysis alone only provides the direction of the wave vector up to the sign. Therefore the sign of the ellipticity in relation to the propagation direction remains undetermined as well. However, for waves in magnetized plasma, the the convention is to use the magnetic field direction instead of the propagation direction as reference for the sign of ellipticity. This is because rotations to the right or to the left with respect to the magnetic field bear the physical significance. With this convention, the sign of ellipticity is determined from the imaginary part of the spectral matrix (Means, 1972).

The coherency is defined using the spectral matrix elements (Rankin & Kurtz, 1970):

$$\gamma = \frac{\left|J_{xy}J_{yx}\right|}{\left|J_{xx}J_{yy}\right|} \tag{13}$$

while, following Song & Russell (1999), the intensity of the coherent part of the wave is

$$I_{\mathrm{coh}} = \lambda_1 + \lambda_2 - 2\lambda_3 \tag{14}$$

Any wave can be decomposed into a sum of left hand polarized and right hand polarized waves (Kodera et al., 1977) with the amplitudes given by:

$$A_{\pm} = \frac{\sqrt{\lambda_1 - \lambda_3}}{2}\left(1 \pm \epsilon\right) \tag{15}$$

From these, the amplitudes of the linear polarized part of the wave and the amplitude of the circular polarized part of the wave can be determined

$$A_{\mathrm{lin}} = \sqrt{2A_{-}A_{+}} \tag{16}$$

$$A_{\mathrm{circ}} = \sqrt{\left|A_{+}^2 - A_{-}^2\right|} \tag{17}$$

Another important wave parameter determined using the magnetic field alone is the compression ratio, i.e. the ratio between the power spectral density of the oscillations orthogonal to the mean magnetic field (transverse power) and the total power spectral density. In a coordinate system with the z axis aligned with the mean magnetic field, the compression ratio is

$$C = \frac{P_x + P_y}{P_x + P_y + P_z} \tag{18}$$

where P_j is the power spectral density of the component j. The power spectral densities for the components are determined from the diagonal elements of the spectral matrix.

All the parameters discussed above are derived solely from the magnetic field. However, the time varying magnetic field of the waves propagating into the plasma is coupled with the electric field, and in addition the waves also disturb and are influenced by the plasma particle distributions. To analyse the coupling between these physical quantities one must generalize the spectral matrix Eq (1) to the cross-spectral matrix of two time series, $u(t)$ and $v(t)$:

$$\mathcal{G}_{uv}(\omega) = \langle \tilde{u}(\omega)\tilde{v}^{\star}(\omega)\rangle \tag{19}$$

From the cross-spectral matrix, a number of fundamental quantities characterizing the relation between the $u(t)$ and $v(t)$ time series can be derived. One of the most important is the phase shift between u and v, for instance between the magnetic field module and the particle density:

$$\varphi_{uv} = \arctan(\mathcal{G}_{uv}) \tag{20}$$

The coherency between $u(t)$ and $v(t)$ gives the similarity degree of two time series. The more stable the phase difference and the amplitude ratio of two signals are, the more coherent they are.

$$\gamma_{uv}(\omega) = \frac{|\mathcal{G}_{uv}|^2}{\mathcal{G}_{uu}\mathcal{G}_{vv}} \tag{21}$$

For a single time series, the coherency Eq. (13) is a particular case Eq. (21).

The co-spectrum, equal to the real part of the extra-diagonal elements, $\Re(\mathcal{G}_{uv})$, represents the in-phase and opposite-phase part of the signals. The quad-spectrum, equal to the imaginary part of the extra-diagonal elements, $\Im(\mathcal{G}_{uv})$, represents phase-quadrature ($\pm\pi/2$ phase difference, advanced/retarded) part of the signals.

Finally, a key parameter of the ULF waves is the Poynting vector which provides the electromagnetic energy flux:

$$\mathbf{S}(\omega) = \frac{1}{2\mu_0}\Re(\tilde{\mathbf{E}} \times \tilde{\mathbf{B}}^{\star}) \tag{22}$$

Knowledge of the Poynting vector supports the identification of stationary waves, and of energy source and sink regions.

3 Derivation of the geometric parameters

All multipoint data analysing techniques depend on the geometric properties of the sensor network. The minimum number of measuring points required for deriving three dimensional quantities and to differentiate between spatial and temporal fluctuations is four, the number of spacecraft in the Cluster fleet. However, not any configuration of four spacecraft is appropriate for three dimensional analysis. The dimensionality is reduced to 2 if all spacecraft are contained in a plane and to 1 if they are aligned along a line. To characterize the configurations lying in between these extremes one may use the elongation and planarity parameters derived from the volumetric tensor which is defined as (Robert, Roux, et al., 1998):

$$R_{ij} = \frac{1}{N}\sum_{\alpha=1}^{N} r_i^{\alpha} r_j^{\alpha} \tag{23}$$

where r_k^{α} is the component k of the position vector $\boldsymbol{r}^{\alpha}$ of spacecraft α relative to the baricenter of the formation, $\sum_{\alpha=1}^{N} \boldsymbol{r}^{\alpha} = 0$.

The characteristic dimensions of the tetrahedron are given by the eigenvalues of the volumetric tensor $\mathcal{R}$, while the eigenvectors give the orientation of the volumetric ellipsoid. For Cluster, $N = 4$ and the volume of the tetrahedron is given by the determinant of the volumetric tensor $V = (8/3)\sqrt{|\mathcal{R}|}$. Therefore when $|\mathcal{R}| = 0$, the formation is either flattened into a 2D plane or elongated into a 1D line.

If $a_1 \geq a_2 \geq a_3$ are the eigenvalues, and $\boldsymbol{R}_1$, $\boldsymbol{R}_2$, $\boldsymbol{R}_3$ are the corresponding eigenvectors of the volumetric tensor $\mathcal{R}$, the elongation is defined as

$$e = 1 - \sqrt{\frac{a_2}{a_1}} \tag{24}$$

with the direction given by the eigenvector $\boldsymbol{R}_1$ corresponding to the maximum eigenvalue. When the elongation approaches 0, the spacecraft tend to be equally distanced from each other, when the elongation approaches 1, the tetrahedron shape degenerates to a string.

The planarity is defined as

$$p = 1 - \sqrt{\frac{a_3}{a_2}} \tag{25}$$

with the normal to the flattening given by the eigenvector $\boldsymbol{R}_3$ corresponding to the minimum eigenvalue. When the planarity approaches 1, all spacecraft are contained in a plane.

The pair (e, p) gives the shape of the formation and determines its dimensionality which is of key importance for multi-spacecraft analysis methods.

The size of the constellation also matters. The characteristic size of the tetrahedron is the largest dimension of the volumetric ellipsoid

$$L = 2\sqrt{a_1} \tag{26}$$

If the distances between the spacecraft are larger than the correlation length of the phenomenon being studied, then the measurements become just a collection of unrelated single point measurements and can be combined only in a statistical sense. Within the correlation length, if the distances between spacecraft are larger than the wavelength / scale of the phenomenon being investigated, spatial aliasing will occur (Sahraoui et al., 2010). If the distances between the spacecraft become too small compared with the scale of the phenomenon being studied, then the measurements become identical and the constellation becomes equivalent with a single point measurement.

4 The online virtual observatory

The almost two solar cycles covered by Cluster multipoint magnetospheric measurements are a valuable resource offering opportunities for unprecedented statistical studies, as well as for case studies of common phenomena in extreme or unusual conditions. For the study of plasma waves over individual time intervals one could use dedicated analysing tools such as PRASSADCO (Cornilleau-Wehrlin et al., 2005) to compute and visualize the key waves parameters. However, the sheer size of the data collected by Cluster over two decades poses new challenges in managing and filtering the data. This is why a database containing high level data is essential to make these studies possible. The more resource-demanding the computation of the high level data is, the more valuable the database becomes. The parameters of the ULF waves discussed in Sec. 2 are among the best candidates for such database.

To take full advantage of the multipoint measurements offered by Cluster, one must employ specialised multipoint techniques, such as the curlometer (Robert, Dunlop, et al., 1998; Dunlop et al., 2002) or the wave telescope (Glassmeier et al., 2001) which work only for specific geometric configurations of the Cluster tetrahedron. A database allowing quick retrieval of these parameters can substantially ease the search for (or confirmation of) suitable configurations.

The magnetic field data used to compute the ULF waves parameters is obtained by calibrating locally stored level 0 FGM data using daily calibration parameters produced by the TUBS team. Other data such as the electric field and the particle density are obtained from the ESA *Cluster Science Archive* (`https://cosmos.esa.int/web/csa`).

4.1 ULF waves parameters

Using measurements of the magnetic field $\boldsymbol{B}$, electric field $\boldsymbol{E}$ and the spacecraft potential (to derive the electron density n (Lybekk et al., 2012)), 47 parameters characterising the ULF waves are computed, saved in the database and plotted to enable quick browsing. A list of the computed parameters can be found in the *Supporting Information*.

Before computing the ULF waves parameters, the magnetic field data is preprocessed to identify data gaps and bad data, and resampled to one vector per second rate.

When available, high resolution (25 or 450 vect/s) electric field data are downloaded from the CSA. These are also resampled to one vector per second and synchronized with the magnetic field data. For the time intervals without high resolution electric field data available, low resolution (one vector per 4 seconds) are used. In this case the resulting frequency domain is more restricted. The electron density is obtained from the spacecraft potential following Lybekk et al. (2012). The advantage over using the density delivered by the dedicated particle instruments is the much higher temporal resolution and data availability.

The computed parameters are valid only under certain conditions. For instance, all the parameters involving the magnetic field lose their validity when the magnetic field variations decrease towards the noise level. A binary mask based on the power spectral density is constructed to mark the invalid time-frequency regions. Similarly, binary masks are constructed based on the PSDs of the electric field and of the Poynting vector, on the polarization degree, on the spectral matrix eigenvalues ratio, on the coherency between the magnetic field and the density variations, and on the angle between the wave vector and the mean magnetic field direction. Not all masks are applied to all parameters, for instance, the last mask ($\widehat{\boldsymbol{kB}}$) is only used to mark as bad the computed ellipticity when the wave vector is nearly parallel to the mean magnetic field. The masks are saved together with the computed parameters and are used to remove the bad values from the plots.

4.1.1 Plots archive

The heart of the ULF Observatory is the browsable archive of plots representing the wave parameters as daily spectrograms. While all parameters enumerated in Sec. 4.1 are saved in the data archive, only a subset of them are part of the plots archive. The plots are organized into six sets of seven parameters each. To facilitate visual analysis, the first panel for all plot sets shows the time series of the magnetic field together with condensed information about the orbit. In addition, the magnetospheric regions crossed by the spacecraft and the location of the magnetic foot points when the spacecraft orbit intersects closed field lines are encoded by horizontal coloured bars. The information about the magnetospheric regions and foot points is obtained from the Goddard SSCWeb interface `http://sscweb.gsfc.nasa.gov` using IGRF internal and Tsyganenko 89C external model and is only meant as a rough guide. Many parameters appear in multiple sets, therefore the total number of distinct parameters in the plots archive is 29. We briefly describe below the default parameter set in the plots archive. The other five sets are shown and discussed in the *Supporting Information*.

The basic parameters set is illustrated in Fig. 1. The top panel shows the time series for the three components of the magnetic field in GSE coordinates. The plotted magnetic field is high-pass filtered to exclude periods larger than one hour. To differentiate between the three components, an offset of 5 nT is added to the x component and subtracted from the y component. The three insets on the left show the projections of the spacecraft orbit on the GSE (x,z), (y,z) and (x,y) planes with the starting point marked by the magenta circle. The coloured bar at the top of the panel indicates the magnetospheric regions crossed by the spacecraft according to Tsyganenko 89C (Tsyganenko, 1989) with $K_p = 3$, in this case dayside magnetosheath during the entire interval. Because of departures of the actual K_p index from the constant index used, and because of the dynamics of the magnetospheric boundaries, one can clearly see from the data that in fact the spacecraft finds itself in the solar wind at the beginning of the interval and the bowshock moves back and forth several times until the spacecraft moves deeper into the magnetosheath. If the spacecraft reaches further into the magnetosphere into the closed field lines region, then the foot points in the northern and southern hemisphere would be indicated by two coloured bars stacked

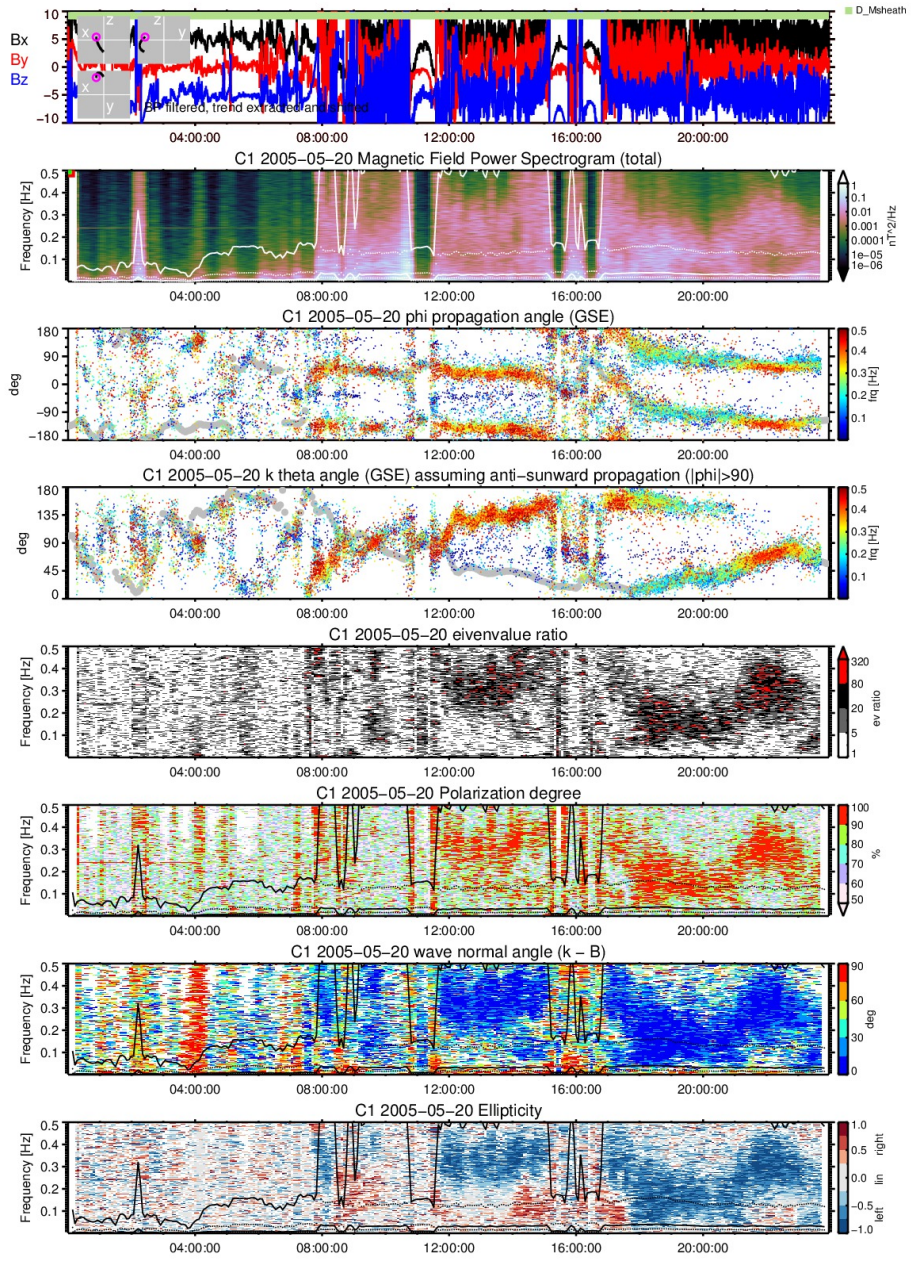

Figure 1. The *basic parameters* plot set for C1 on 20th May 2005. From top to bottom: GSE components of the magnetic field, HP filtered. The insets show the orbit in GSE, the top bar shows the magnetospheric region; Sum of the PSD of the magnetic field components. The white lines show the gyrofrequencies of H, He, and O; Azimuth and elevation directions of the wave vector in GSE; Colour encodes the frequency, grey thick line shows the mean field direction; Ratio between intermediate and minimum eigenvalue; Polarization degree; Angle between wave vector and mean field; Ellipticity.

at the bottom of the panel, not present in this figure, but visible in Fig. S10 showing the web interface.

The first parameter in the set is the total power spectral density of the magnetic field fluctuations, defined as the sum of the power spectral densities of the three GSE components computed – as all other parameters – using 1 s resolution data and a sliding window W of 2048 s with a step Δt of 256 s. The PSD per Hz per s in the time-frequency cell ($\nu - \Delta\nu/2$ to $\nu + \Delta\nu/2, t - \Delta t/2$ to $t + \Delta t/2$) for the component j is computed as

$$P_j(\nu, t) = 2W\left|\tilde{B}_j(\nu)\right|^2\Big|_{t-\frac{W}{2}}^{t+\frac{W}{2}} \tag{27}$$

where W is the window length in seconds, $\Delta\nu = 1/W$ is the FFT frequency resolution in Hz, Δt is the sliding step in seconds and $\tilde{B}_j(\nu)$ is the Fourier transform Eq. (2) of the magnetic field evaluated in the time interval $(t - W/2, t + W/2)$. The total wave energy contained in the $(\Delta\nu, \Delta t)$ cell is equal with the PSD multiplied by the cell area, $\Delta\nu\Delta t$: $E_{\text{cell}}(\nu, t) = 2\Delta t|\tilde{B}_j(\nu, t)|^2$.

The gyrofrequencies of the H, He and O computed from the magnetic field smoothed using a boxcar average of 1024 s are plotted on top of the spectrogram with white solid, dotted and again solid lines, respectively. The position of the spacecraft relative to the GSE (x, y) (equatorial) and (x, z) (noon-midnight meridian plane) planes is marked at the top of this panel by coloured bars as follows: Red if the angle between the position vector and the equatorial plane is less than 10°. Yellow if the angle between the position vector and the noon-midnight meridian plane is less than 10° and the x coordinate is positive (dayside). Green if the angle between the position vector and the noon-midnight meridian plane is less than 10° and the x coordinate is negative (nightside). None of these conditions occurred during 20th May 2005 therefore no coloured bars are present in Fig. 1 PSD panel. Fig. S10 shows an example when the spacecraft orbit intersects both planes.

The next two panels show the orientation of the wave vector $\boldsymbol{k}$ Eq. (5) as given by the azimuth φ and by the elevation θ angles in GSE coordinates. The corresponding frequencies are colour coded. In both panels the mean magnetic field direction is represented with a thick grey line in the background. Because of the uncertainty in the sign of $\boldsymbol{k}$, frequent jumps between opposite directions of the $\boldsymbol{k}$ vector can be noticed in the azimuth angle. We could have limited the representation to half-space without losing any information, however, in order to show the mean field direction together with the wave vector we choose to represent the entire space. Nevertheless, we changed the elevation angle to correspond to anti-sunward propagation ($|\varphi| \leq 90$). With this convention inside the magnetosheath the waves propagate anti-parallel to the magnetic field at the beginning and parallel to the magnetic field towards the end. Outside the magnetosheath the propagation direction is less well defined. During this day the wave propagation direction does not depend on frequency.

To remove from the plot low power fluctuations we discarded the time-frequency cells corresponding to total PSD less than $10^{-3}\,\text{nT}^2\,\text{Hz}^{-1}\,\text{s}^{-1}$. We applied this power mask to all parameters derived from the magnetic field, except the other power spectral densities. Moreover, to reduce the noise in the plot we discarded the time-frequency cells corresponding to polarization degree below 70%. Finally, because a low eigenvalue ratio leads to large errors in the determined wave vector direction we also discarded the time-frequency cells corresponding to eigenvalue ratios less than 5.

The ratio between the intermediate and the minimum eigenvalues of the spectral matrix Eq. (1) plotted in the next panel provides useful information about how well the wave propagation direction is defined. To properly represent the range of eigenvalue ratios we use a logarithmic base 4 representation. All cells below the threshold ratio

(equal to 5) are masked out (white) and all values above 80 are plotted with red. Other masks applied: PSD.

The polarization degree Eq. (8) is shown in the next panel. Only the PSD mask is applied here. Note that the high polarization domain resembles the large eigenvalue ratio domain in the previous panel. Here and in the next two panels the gyrofrequencies of H, He and O are plotted with black lines on top of the spectrogram.

The penultimate panel shows the wave normal angle, i.e. the angle between the wave vector and the mean magnetic field computed over the window length W. Because of the sign uncertainty we reduced the angles to the $[0°, 90°]$ interval. For this panel we applied the masks for PSD, polarization degree and eigenvalue ratios with the thresholds mentioned above. One can see that during this day most of the waves with high polarization and power in the magnetosheath propagate either parallel or anti-parallel to the mean magnetic field. In front of the magnetosheath the waves do not seem to have a preferred direction with the exception of a short time interval around 04 UT when the wave vector is orthogonal to the magnetic field for all frequencies.

The last panel in this set shows the waves ellipticity Eq. (11). For this panel we applied the same masks as for the wave normal angle. In addition we masked out the time-frequency cells with near orthogonal propagation ($\alpha_{kB} \geq 80°$) because for orthogonal propagation the ellipticity is undefined. Most of the parallel propagating waves are circularly polarized to the left with some exceptions in the lower frequency range close to the bowshock crossings when the polarization is to the right. In front of the bowshock no preferred ellipticity is observed with the exception of the waves around 04 UT showing linear polarization.

4.1.2 Data archive

The daily 330×1025 time-frequency matrix for each parameter is saved using the *Hierarchical Data Format* (HDF) standard (Poinot, 2010), together with a low resolution plot for quick reference. The corresponding time and frequency vectors, details on how the parameter was computed, the reference system, and the units used are saved as well in the same HDF file. The total size of the archived parameters is over 2 TB. This high level database can be used to automatically search for events satisfying certain imposed conditions for case or statistical studies.

A subset of the low resolution plots meant for quick visual inspection of the database content are presented in the *Supporting Information*. To avoid obscuring potential significant features, no masks are applied, therefore one should exercise caution when interpreting these plots.

4.2 Geometric parameters

The geometric parameters characterize the shape and the size of the Cluster tetrahedron. They are essential for determining the applicability and for the error evaluation of multi-point analysis techniques. In Sec. 3 we introduced the elongation and the planarity which condense the shape information for a tetrahedron in an intuitive fashion. Since both can only take values between 0 and 1, any tetrahedron shape corresponds to one point in the unit square of the elongation-planarity diagram. Table 1 summarises the possible shapes in the (e,p) space.

For a long term overview of the geometric parameters of the Cluster tetrahedron we produced plots such as the one shown in Fig. 2. Each plot consists of four panels, each panel showing the mean inter-spacecraft distance during one year. The shape of the formation is encoded in the colour of the plot line, with a colour-key in the upper right corner. The colour-key is a representation of the (e,p) space, with the origin

Table 1. Shapes in the planarity-ellipticity domain. Adapted from Robert, Roux, et al. (1998).

p \ e	0	low	intermediate	large	1
1	circle	ellipse			
large	ellipsoid	pancake	elongated pancake	knife blade	line
intermediate		thick pancake	potato	flat cigar	
low		pseudo sphere	short cigar	cigar	
0	sphere	rugby ball			

$(0, 0)$ at the lower left corner, and maximum value $(1, 1)$ at the upper right corner. The red colour near origin indicates a nearly regular tetrahedron. The yellow colour (large elongation, low planarity) indicates cigar shapes. The green colour indicates shapes resembling a long, flat knife blade. The blue colour indicates nearly circular flattened pancake shaped formations. The grey colour in the centre indicates irregular "potato" shapes. The function chosen to map the (e,p) space to the (r,g,b) space is bijective, therefore one can estimate the shape of the tetrahedron at a given moment in time from the colour of the corresponding line in the plot.

Yearly, monthly and daily plots are also provided. Detailed descriptions and example plots are included in the *Supporting Information*.

4.3 The web interface

The Cluster virtual observatory for ULF waves is organized into three main sections. The *ULF quickplots* (Fig. S10, `http://plasma.spacescience.ro/waves/index.html`) section offers access to the plots archive discussed in sec. 4.1.1. The spacecraft, date, and plot set can be selected using the buttons above the image. Below the image, there are links to high resolution JPG or pdf formats of the plots, to a catalogue containing the plots for the entire year, and to a text file with details about how the plots were produced. This text file is embedded both in the JPG image header and in the pdf file.

The *ULF data* (Fig. S11, `http://plasma.spacescience.ro/waves/spectra-data/`) section offers access to the ULF waves parameters data archive discussed in sec. 4.1.2. The spacecraft, year, month and parameter are selected using the buttons at the top of the page. The "go" button displays the low resolution images of the selected parameter for the selected month together with links to the HDF files containing the archived data.

The *Tetrahedron geometry* (Fig. S12, `http://plasma.spacescience.ro/waves/geometry.html`) section offers access to the plots of the geometric parameters of the Cluster fleet discussed in sec. 4.2. From the buttons at the top of the page, the desired plot can be selected and displayed. Links to high resolution JPG and pdf files as well as to a yearly catalogue are provided at the bottom of the page.

5 Summary and conclusions

The large volume of data accumulated by the Cluster mission in the last two decades is both a valuable resource and a challenge to digest in an efficient manner. The many parameters characterizing the low frequency plasma waves are a good example of high level data whose availability have the potential to significantly reduce the effort needed to select relevant events or to conduct large statistical studies. If multipoint analyses are performed, a possibility to quickly find appropriate spacecraft

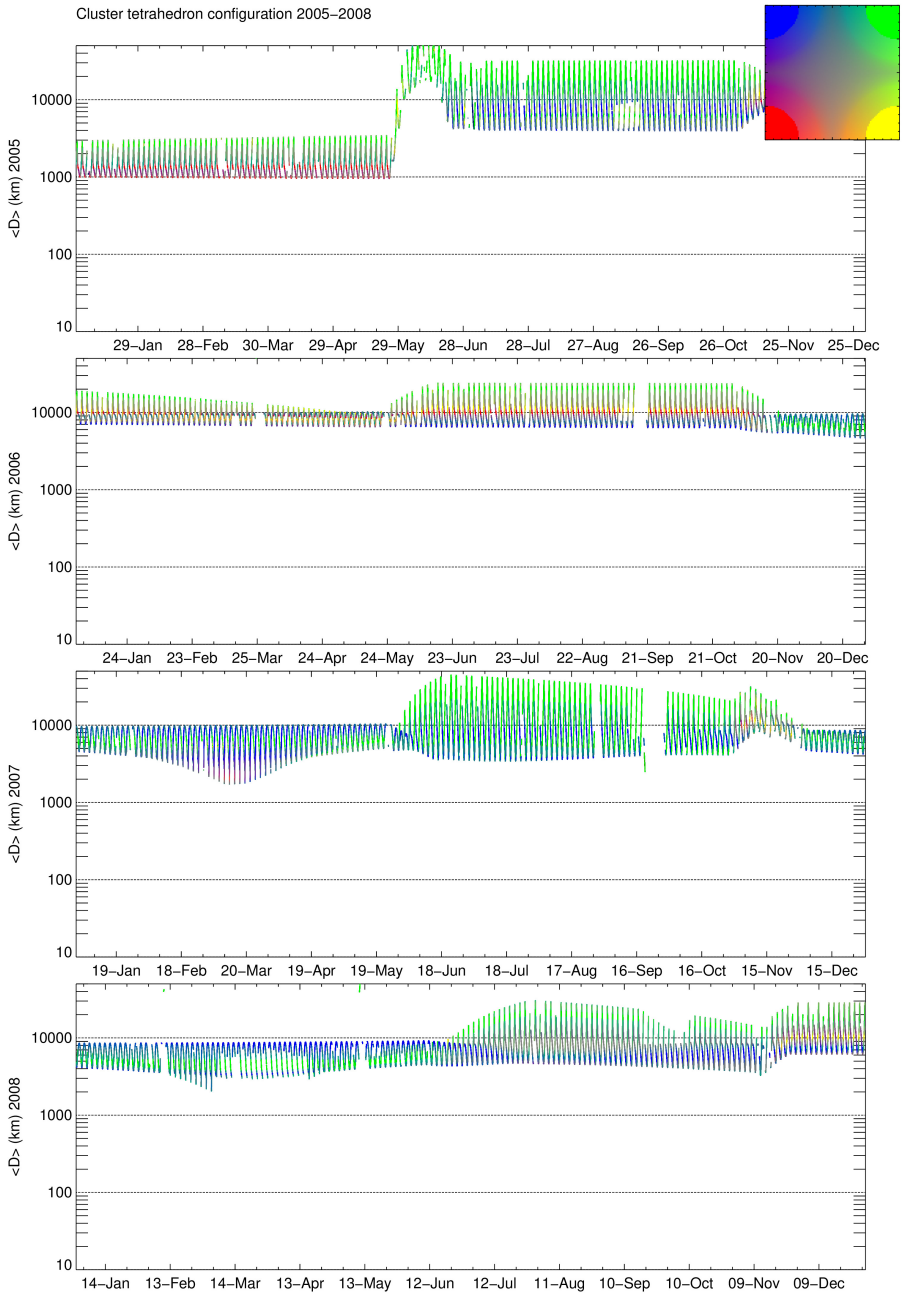

Figure 2. Geometric parameters of the Cluster tetrahedron between 2005 and 2008. The y axis shows the mean inter-spacecraft distance and the colour encodes the tetrahedron shape in the (e,p) domain as given by the colour legend in the upper right.

configurations is highly desirable as well. The Cluster Virtual Observatory (CVO) offers both the high level ULF waves parameters and an optimized interface to graphical representations of the spacecraft configuration. A browsable database of daily plots of the ULF waves parameters allows for a rapid search for significant events and provides publication quality images. The main data used to build the CVO database is the magnetic field delivered by the Cluster FGM instruments. Additionally, the spacecraft potential and the electric field from the EFW instrument are used to compute specific parameters. For a rough positioning within the magnetospheric regions, data from the Goddard SSCWeb is utilised. The Cluster Virtual Observatory can be accessed without restrictions at `http://plasma.spacescience.ro/cluster.html`.

Acknowledgments

This work was financially supported by the Deutsches Zentrum für Luft- und Raumfahrt under contract 50OC1803. The *ULF quick plots* and *ULF data* sections of the CVO until 2018 were supported by the Romanian Space Agency STAR program, Contract 149/20.07.2017. Datasets for this work are available from the ESA *Cluster Science Archive* (`https://cosmos.esa.int/web/csa`) and from the Goddard SSCWeb interface (`http://sscweb.gsfc.nasa.gov`).

References

Arthur, C. W., McPherron, R. L., & Means, J. D. (1976). A comparative study of three techniques for using the spectral matrix in wave analysis. *Radio Science*, *11*(10), 833–845. doi: 10.1029/RS011i010p00833

Balogh, A., Dunlop, M. W., Cowley, S. W. H., Southwood, D. J., Thomlinson, J. G., Glassmeier, K. H., ... Kivelson, M. G. (1997, January). The Cluster Magnetic Field Investigation. *Space Science Reviews*, *79*, 65-91.

Cornilleau-Wehrlin, N., Chauveau, P., Louis, S., Meyer, A., Nappa, J. M., Perraut, S., ... Louarn, P. (1997, January). The Cluster Spatio-Temporal Analysis of Field Fluctuations (STAFF) Experiment. *Space Science Reviews*, *79*, 107-136.

Cornilleau-Wehrlin, N., Alleyne, H. S., Yearby, K. H., de la Porte de Vaux, B., Meyer, A., Santolík, O., ... Cao, J. (2005). The staff-dwp wave instrument on the dsp equatorial spacecraft: description and first results. *Annales Geophysicae*, *23*(8), 2785–2801. doi: 10.5194/angeo-23-2785-2005

Dunlop, M. W., Balogh, A., Glassmeier, K. H., & Robert, P. (2002, November). Four-point Cluster application of magnetic field analysis tools: The Curlometer. *J. Geophys. Res. Space Phys.*, *107*(A11), 1384. doi: 10.1029/2001JA005088

Escoubet, C. P., Schmidt, R., & Goldstein, M. L. (1997). Cluster: Science and Mission Overview. *Space Sci. Rev.*, *79*, 11-32.

Fowler, R., Kotick, B., & Elliott, R. (1967). Polarization Analysis of Natural and Artificially Induced Geomagnetic Microplulsations. *J. Geophys. Res.*, *72*, 2871-2883.

Glassmeier, K.-H. (1995). ULF pulsations. In H. Volland (Ed.), *Handbook of atmospheric electrodynamics* (p. 463-502). CRC Press.

Glassmeier, K.-H., Motschmann, U., Dunlop, M., Balogh, A., Acuña, M. H., Carr, C., ... Buchert, S. (2001). Cluster as a wave telescope - first results from the fluxgate magnetometer. *Ann. Geophys.*, *19*, 1439-1447. (Correction, Annales Geophysicae, 21, 1071, 2003)

Keiling, A., Lee, D., & Nakariakov, V. (Eds.). (2016). *Low frequency waves in space plasmas*. American Geophysical Union. doi: 10.1002/9781119055006

Kodera, K., Gendrin, R., & de Villedary, C. (1977, Mar). Complex representation of a polarized signal and its application to the analysis of ULF waves. *J. Geophys. Res.*, *82*(7), 1245. doi: 10.1029/JA082i007p01245

514 Lybekk, B., Pedersen, A., Haaland, S., Svenes, K., Fazakerley, A. N., Masson, A.,
515 ... Trotignon, J.-G. (2012, January). Solar cycle variations of the Cluster space-
516 craft potential and its use for electron density estimations. *Journal of Geophysical*
517 *Research (Space Physics)*, *117*, A01217. doi: 10.1029/2011JA016969
518 McPherron, R. L., Russell, C. T., & Coleman, P. J., Jr. (1972, July). Fluctuating
519 Magnetic Fields in the Magnetosphere. II: ULF Waves. *Space Sci. Rev.*, *13*, 411-
520 454. doi: 10.1007/BF00219165
521 Means, J. D. (1972, October). Use of the three-dimensional covariance matrix in
522 analyzing the polarization properties of plane waves. *J. Geophys. Res.*, *77*, 5551-
523 5559. doi: 10.1029/JA077i028p05551
524 Motschmann, U., Woodward, T. I., Glassmeier, K. H., Southwood, D. J., & Pinçon,
525 J. L. (1996, March). Wavelength and direction filtering by magnetic measurements
526 at satellite arrays: Generalized minimum variance analysis. *J. Geophys. Res.*, *101*,
527 4961-4966.
528 Pilipenko, V. A. (1990, December). ULF waves on the ground and in space.
529 *Journal of Atmospheric and Terrestrial Physics*, *52*, 1193-1209. doi: 10.1016/
530 0021-9169(90)90087-4
531 Pinçon, J., & Motschmann, U. (1998). Multi-Spacecraft Filtering: General Frame-
532 work. In G. Paschmann & P. Daly (Eds.), *Analysis methods for multi-spacecraft*
533 *data* (p. 65-78). Bern: ISSI.
534 Poinot, M. (2010, September). Five Good Reasons to Use the Hierarchical Data For-
535 mat. *Computing in Science and Engineering*, *12*(5), 84-90. doi: 10.1109/MCSE
536 .2010.107
537 Rankin, D., & Kurtz, R. (1970, Oct). Statistical study of micropulsation polariza-
538 tions. *J. Geophys. Res.*, *75*(28), 5444-5458. doi: 10.1029/JA075i028p05444
539 Robert, P., Dunlop, M. W., Roux, A., & Chanteur, G. (1998, January). Accuracy of
540 Current Density Determination. *ISSI Scientific Reports Series*, *1*, 395-418.
541 Robert, P., Roux, A., Harvey, C. C., Dunlop, M. W., Daly, P. W., & Glassmeier,
542 K.-H. (1998, January). Tetrahedron Geometric Factors. *ISSI Scientific Reports*
543 *Series*, *1*, 323-348.
544 Sahraoui, F., Belmont, G., Goldstein, M. L., & Rezeau, L. (2010, April). Limitations
545 of multispacecraft data techniques in measuring wave number spectra of space
546 plasma turbulence. *Journal of Geophysical Research (Space Physics)*, *115*(A4),
547 A04206. doi: 10.1029/2009JA014724
548 Samson, J. C. (1973, December). Descriptions of the Polarization States of Vec-
549 tor Processes: Applications to ULF Magnetic Fields. *Geophysical Journal Interna-*
550 *tional*, *34*, 403-419. doi: 10.1111/j.1365-246X.1973.tb02404.x
551 Santolík, O., Parrot, M., & Lefeuvre, F. (2003, February). Singular value decompo-
552 sition methods for wave propagation analysis. *Radio Science*, *38*(1), 1010. doi: 10
553 .1029/2000RS002523
554 Song, P., & Russell, C. T. (1999, January). Time Series Data Analyses in Space
555 Physics. *Space Sci. Rev.*, *87*, 387-463. doi: 10.1023/A:1005035800454
556 Tsyganenko, N. A. (1989, January). A magnetospheric magnetic field model with
557 a warped tail current sheet. *Planet. Space Sci.*, *37*(1), 5-20. doi: 10.1016/0032
558 -0633(89)90066-4