

Earth and Space Science



METHOD

10.1029/2022EA002692

Key Points:

- We built an interactive software analysis tool designed to study space plasma turbulence and intermittency
- The tool can ingest data from a large variety of sources and includes a varied portfolio of analysis methods
- The performance of our analysis tool is demonstrated using in situ magnetic field measurements from the Cluster 3 spacecraft

Correspondence to:

C. Munteanu,
costelm@spacescience.ro

Citation:

Munteanu, C., Kovács, P., & Echim, M. (2023). An integrated nonlinear analysis (INA) software for space plasma turbulence. *Earth and Space Science*, 10, e2022EA002692. <https://doi.org/10.1029/2022EA002692>

Received 21 OCT 2022

Accepted 3 FEB 2023

Author Contributions:

Conceptualization: Costel Munteanu, Péter Kovács, Marius Echim

Data curation: Costel Munteanu

Formal analysis: Costel Munteanu

Funding acquisition: Marius Echim

Methodology: Costel Munteanu, Péter Kovács, Marius Echim

Project Administration: Marius Echim

Software: Costel Munteanu

Supervision: Péter Kovács, Marius Echim

Validation: Costel Munteanu, Péter Kovács, Marius Echim

Visualization: Costel Munteanu, Péter Kovács, Marius Echim

Writing – original draft: Costel Munteanu

© 2023 The Authors. Earth and Space Science published by Wiley Periodicals LLC on behalf of American Geophysical Union.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs License](https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

An Integrated Nonlinear Analysis (INA) Software for Space Plasma Turbulence

Costel Munteanu¹ , Péter Kovács² , and Marius Echim^{1,3} 

¹Institute of Space Science (ISS), Măgurele, Romania, ²Wigner Research Centre for Physics, Budapest, Hungary, ³The Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium

Abstract We built an integrated nonlinear analysis software (INA)-designed to study space plasma turbulence and intermittency. The MATLAB programming environment was used for the algorithmic development and implementation of methods for spectral analysis, multiscale fluctuations and multifractal analysis. The performance of INA is demonstrated using magnetic field measurements during an inbound pass through the Earth's magnetosheath region of the Cluster 3 spacecraft. We show how specific features of the power spectral density can be mapped to localized time-frequency regions in the spectrogram representation, and identify multiple intermittent events using the wavelet-based local intermittency measure. Multiscale probability density functions showed clear departures from Gaussianity, signifying the presence of intermittency. Structure functions and rank-ordered multifractal analysis revealed the multifractal nature of the analyzed signal. INA is freely distributed as a standalone executable file to any interested user, and provides an integrated, interactive, and user-friendly environment in which one can import a data set, customize key analysis parameters, apply multiple methods on the same signal and then export high-quality, publication-ready figures. These are only a few of the many distinguishing features of INA.

1. Introduction

Most space plasmas, like, for example, the solar wind, are frequently in a turbulent state characterized by multi-scale fluctuations (Alexandrova et al., 2013; Bruno & Carbone, 2013; Dudok de Wit et al., 2013). Intermittency is a fundamental characteristic of solar wind variability, associated with the emergence of discontinuities, coherent structures leading to dissipation, heating, transport, and acceleration of charged particles (Bruno, 2019; Greco et al., 2018; Matthaeus et al., 2015). Significant efforts are devoted to understand turbulence and intermittency in space plasmas (Bruno et al., 2001; Burlaga, 1991; Chang et al., 2010; Echim et al., 2007, 2021; Hnat et al., 2002; Marsch & Tu, 1994; Sorriso-Valvo et al., 2001; Teodorescu et al., 2021; Vörös et al., 2003; Wawrzaszek & Echim, 2021; Wawrzaszek et al., 2015; Yordanova et al., 2005).

Variability of space plasma parameters is often investigated from in situ measurements using either spectral or statistical approaches. In the spectral approach, one estimates the energy content and how it is distributed over frequencies. The presence of turbulence is traditionally revealed by a characteristic power-law behavior of the power spectral density (PSD). However, the information contained in PSD captures only partially the turbulence dynamics, especially for the cases when intermittent events are also present (Lion et al., 2016; Roberts et al., 2017). On the other hand, the statistics of fluctuations can provide the principal tendencies (mean, variance) and their evolution, and also more advanced, higher order, measures of variability. The multiscale probability density functions (PDFs) can be used to explore additional features of turbulence compared to spectral analysis. For instance, the departure of PDFs from a Gaussian shape and the increase of Flatness (the normalized fourth-order moment of fluctuations) toward smaller scales, are generally considered signatures of intermittency (Bruno et al., 2001; Marsch & Tu, 1994). The anomalous scaling of structure functions (SFs) is the starting point for more elaborate models of turbulence based on fractals and multifractals (e.g., Benzi et al., 1993; Sreenivasan, 1991). Conventional SF analysis is based on the statistics of the full set of fluctuations, thus, the characteristics of large amplitude fluctuations (which usually are a minority) can be easily masked by those of the dominant, small amplitude population (e.g., Echim et al., 2021). This motivated Chang and Wu (2008) to develop the rank-ordered multifractal analysis (ROMA), a method of isolating out the subdominant large amplitude fluctuations and investigate their multifractal nature.

Writing – review & editing: Costel Munteanu, Péter Kovács, Marius Echim

Spectral and statistical approaches are traditionally applied independently. The two approaches are in fact relevant for the two different schools of thought in space plasma turbulence: (a) weak plasma turbulence, which considers at the center of the turbulence paradigm different types of waves (e.g., Biskamp, 1993) and (b) strong plasma turbulence, which is based mainly on nonlinear interactions between coherent structures (e.g., Goldreich & Sridhar, 1995). Modern theories of plasma complexity consider however that the two states emerge spontaneously in space plasmas, and disentangling specific effects is a task that requires the simultaneous use of both spectral and statistical approaches (Chang, 2015).

During our studies, we often encountered a gap between software analysis tools and turbulence research. We found there is no easy, straightforward, and user-friendly way of performing at once most of the analyses mentioned above. Also, even when the software solutions are relatively easy-to-use for non-experienced programmers, they are often limited in spread and scope of the methods. The Queen Mary Science Analysis System (QSAS, <http://www.sp.ph.ic.ac.uk/csc-web/QSAS>), a software package providing an environment for the selection, manipulation, and display of space physics data, is one such solution. Another example is the Space Physics Environment Data Analysis Software (SPEDAS, <http://spedas.org/wiki>), a generalized software development platform supporting multi-mission data ingestion, analysis and visualization. These tools function mostly as mission/instrument dedicated plug-ins. QSAS, for example, has a Cluster plug-in for computing the curl, gradient, divergence, and barycentric average of a vector quantity (http://www.ninepeux.net/haaland/QSAS_plugins.html). SPEDAS includes multiple plugins for loading and analyzing Magnetospheric Multiscale (MMS) mission data (http://spedas.org/wiki/index.php?title=Magnetospheric_Multiscale). These plugins are usually created and supported by the respective mission/instrument teams, and are useful to analyze data from that specific mission/instrument. IRFU-Matlab is another useful collection of tools for analyzing space physics data (<https://sites.google.com/site/irfumatlab>), strictly linked to the MATLAB proprietary environment (<https://www.mathworks.com/products/matlab.html>); it includes routines for: single-spacecraft (minimum variance, Walén analysis, ...) and multi-spacecraft methods (curlometer, timing, ...), data reading support (Cluster, OMNI, ...), and plotting routines to generate publishing quality figures (time series, spectrograms, ...).

General purpose commercial software, like IDL (<https://www.l3harrisgeospatial.com/Software-Technology/IDL>) or MATLAB, are broader in scope and are rather expensive, but are widely used in the space physics community. MATLAB interactive apps, for example, are becoming increasingly versatile and easy to use for data science. The Signal Analyzer app (<https://www.mathworks.com/help/signal/ref/signalanalyzer-app.html>), for example, is an interactive app for visualizing, preprocessing, and analyzing signals in time, frequency, and time-frequency domains; one can access the MATLAB workspace, and resample, detrend or denoise signals without leaving the app; spectrum, spectrogram, and scalogram representations of signals can be inspected simultaneously; the Signal Analyzer tool can even be supplemented with custom preprocessing routines. However, MATLAB and IDL do not include (out of the box) the advanced turbulence and intermittency methods mentioned above, like multiscale PDFs, SFs or ROMA. Thus, most researchers devote considerable time and effort to develop their own algorithms and techniques to study turbulence and intermittency. The problem is further complicated by the ever-increasing amount of publicly available space physics data, which is provided in many different formats and file types. Thus, in addition to implementing their own analysis routines, researchers are also required to do their own data importing and preprocessing routines, and this is also a tedious and difficult task.

We note that most of the tools developed to study space records generally lack advanced analysis routines for investigating nonlinear turbulent and intermittent properties of the plasma time-series. One noteworthy exception are the open source routines distributed by Yongxiang Huang through: <https://github.com/lanlankai?tab=repositories> (see also Schmitt and Huang (2016)). This repository includes MATLAB routines for SFs and PDFs. Notably, a fast algorithm that calculates PDFs of increments without the use of box-counting with uniform bins, is provided. A fast algorithm to deal with missing data can also be found in this git-hub repository.

A recent noteworthy contribution is also the “Open-source software analysis tool to investigate space plasma turbulence and nonlinear DYNAMics” (ODYN, Teodorescu & Echim, 2020). ODYN is a Python-based tool which includes a rich collection of methods to investigate space plasma turbulence and intermittency: PSD, PDFs, multifractals, and discriminating statistics. A customizable Jupyter notebook (<https://jupyter.org>) is included with ODYN, facilitating the interaction with the analysis algorithms, and providing users an intuitive way of configuring key analysis parameters. Another distinguishing feature of ODYN is its ability to run in a “batch mode,” allowing users to apply one or all of the implemented analysis methods on multiple data files at the same time.

We designed and built an integrated software analysis tool -INA-dedicated specifically to investigate turbulence and intermittency in space plasmas. Integrated nonlinear analysis (INA) integrates a comprehensive collection of analysis methods, from lower order spectral methods to highly complex multiscale PDFs and multifractal methods. One of the main distinctive features of INA is the interactive and user-friendly environment in which one can easily import a data set, customize key analysis parameters, apply multiple analysis methods, cross-validate results, and export high-quality figures. These features make INA unique compared to similar software products dedicated to turbulence analysis, like, for example, the ODYN tool.

INA was developed in the framework of the European Community's Seventh Framework Programme project Solar system plasma Turbulence: Observations, inteRmittency and Multifractals (STORM), and is publicly available from the project website at <https://www.storm-fp7.eu>. The program was carefully tested and validated during STORM. The tool is primarily designed to ingest data from space missions like Cluster, Advanced Composition Explorer (ACE), Venus Express and Ulysses, but can also read data from other sources. INA can be used to perform a comprehensive analysis of time series, including methods for: descriptive statistics, spectral analysis, multiscale analysis of fluctuations, wavelet analysis and SFs. INA also includes the ROMA method, and this makes it one of only a few publicly available software implementing this method. In fact, to our knowledge, INA and ODYN are the only freely available tools which include the ROMA analysis method.

INA was built using MATLAB, a proprietary programming platform used by engineers and scientists. In addition to classical script-based programming, MATLAB can also be used to create standalone applications with custom, user-friendly graphical user interfaces with standard components such as buttons, check boxes, and drop-down lists. The standalone applications generated using MATLAB Compiler can be shared and used freely by any interested user, even if they don't own a MATLAB license. In addition to the executable file, the MATLAB source code for INA is also available on request. The only prerequisite needed to run the INA executable is the free MATLAB Runtime, available at <https://www.mathworks.com/products/compiler/matlab-runtime.html>.

In a recent series of papers (Deak et al., 2018, 2021; Munteanu et al., 2022; Opincariu et al., 2019; Turicu et al., 2022), we used the INA software extensively during the design, implementation and validation phases of hardware, Field-Programmable Gate Array (FPGA) implementations of various analysis methods. In Deak et al. (2018), the INA software implementation of PDFs computation, was the starting point for the FPGA implementation of the same method. Opincariu et al. (2019) added spectral analysis to our FPGA toolbox, again, starting from ideas and algorithms implemented in INA. Deak et al. (2021) used the INA mathematical kernel and algorithm for the Flatness parameter calculations, and generated an FPGA implementation of this method. In Turicu et al. (2022), a local stationarity measure is implemented in FPGA, again, based on an INA software implementation of the same measure. An FPGA implementation of a magnetic field directional discontinuity detector is presented in Munteanu et al. (2022), and, same as before, the mathematical kernel of the discontinuity detector was first developed and validated using INA, and then ported onto FPGA.

The paper is organized as follows. Section 2 gives a brief mathematical description the analysis methods currently implemented in INA. In Section 3 we provide a detailed program overview. Section 4 presents an illustrative example using magnetic field measurements from the Cluster 3 spacecraft during a magnetosheath inbound pass on 09 February 2007. Section 5 summarizes the paper.

2. Description of the Analysis Methods Implemented in INA

INA includes a comprehensive collection of analysis methods dedicated to study turbulence and intermittency: starting from descriptive statistics (histograms); advancing to spectral analysis (both Fourier-based and wavelet-based methods); then moving on to advanced methods to study multiscale fluctuations (PDFs and SFs); and, finally, the ROMA, a recently developed method to study multifractals.

2.1. Spectral and Wavelet Analyses

The Periodogram is the fastest way to estimate the PSD of a signal $x(t)$. The fast Fourier transform is used to compute the periodogram through a discrete Fourier transform of the signal. The PSD is then computed from the squared amplitudes of the Fourier transform (Bloomfield, 2000):

$$PSD(f) = \frac{1}{n} X(f)^2, \text{ where } X(f) = \sum_{t=0}^{n-1} x(t) \cdot e^{-i2\pi ft}, \quad (1)$$

where i is the imaginary unit, n is the number of samples, and f is the frequency in Hz. The Periodogram is usually represented as a log-log plot of $\text{PSD}(f)$ versus f .

The Periodogram is generally dominated by noise, which makes it difficult to estimate accurately properties like the average power or spectral index. In order to reduce this noise, Welch (1967) introduced a method to divide the signal into segments, to compute periodograms for each segment, and then take an average of the resultant periodograms to obtain a less noisy PSD estimate, with the cost of weaker frequency resolution.

Fourier analysis is also used to compute spectrograms. A spectrogram is a time-frequency representation of the signal depicting the time evolution of the PSD, and is computed using a moving-window approach. The PSD-Welch and Spectrogram methods are included in the INA Spectral Analysis module.

An alternative time-frequency representation of a signal can be obtained from wavelet-based methods. The windowed trigonometric kernel functions used when computing the Fourier-based spectrograms are designed to have the same temporal resolution, regardless of frequency. In contrast, the basis (or mother) functions of the wavelet transform can be scaled, so that their support can be adapted to the investigated frequency band: narrow functions are used for high frequencies and broader functions for low frequencies. As a result, wavelet transforms are much better at localizing short-time, high-frequency events, compared to the Fourier-based spectrograms.

The continuous wavelet transform of a time series $x(t)$ is Daubechies (1992):

$$W(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \cdot \psi\left(\frac{t-b}{a}\right) dt, \quad (2)$$

where a is the scale, b is the translation parameter, ψ is the wavelet (mother) function, and $W(a, b)$ are the wavelet coefficients. A scalogram, the wavelet equivalent of a spectrogram, is a 3D representation of (the logarithm of) $|W(a, b)|^2$, usually plotted as a function of time and frequency (with b as the equivalent of time, and a as the equivalent of reciprocal frequency, i.e., $a \sim 1/f$).

The Local Intermittency Measure (LIM, Farge, 1992) is a frequently used wavelet-based method defined as:

$$LIM(a, b) = |W(a, b)|^2 / \langle |W(a, b)|^2 \rangle_b, \quad (3)$$

where $\langle |W(a, b)|^2 \rangle_b$ denotes the average over b of all (squared) wavelet coefficients corresponding to scale a . Local intermittency measure quantifies the deviations between the local and the average power of the signal $x(t)$ at each scale a . In other words, since $\langle |W(a, b)|^2 \rangle_b$ gives the mean energy content (power) of signal x at scale a , LIM values close to 1 refer to lack of intermittency in the underlying period, with fluctuation energy close to the average. Conversely, values of LIM > 1 identify those portions of the signal $x(t)$ that have more power than the average. Consequently, regions with LIM > 1 can be used to locate intermittent events in time and scale. A dedicated Wavelet Analysis module is included in INA.

2.2. Multiscale Study of Fluctuations

An incremental measure for the variability of a physical variable x is constructed from time differences computed at different scales τ :

$$\Delta x(t, \tau) = x(t + \tau) - x(t). \quad (4)$$

The PDFs at each scale τ are estimated from the normalized histograms of $\Delta x(t, \tau)$. These are sometimes referred to as unscaled PDFs. Multiscale PDFs can be computed from the Probability Density module of INA, which includes also the one parameter rescaling (OPR) method (Hnat et al., 2002), which aims to find a single parameter able to rescale/collapse the PDFs on a single (master) curve. The assumption behind OPR is that the multiscale probability densities can be described by a symmetric universal shape. To verify this assumption, one generates a log-log plot of the peak values of the PDFs versus τ . An OPR scaling exponent is then determined from the slope of this plot.

Multiscale PDFs provide valuable data on the intermittency of a fluctuating phenomenon. The conventional method to evaluate intermittency is to study the scaling behavior of the moments of order q of the time increments, known as the structure functions (SFs; e.g., Chang & Wu, 2008):

$$SF_q(\tau) = \langle |\Delta x(t, \tau)|^q \rangle = \int_0^{\Delta x_{\max}} |\Delta x(t, \tau)|^q \cdot P(\Delta x, \tau) d(\Delta x), \quad (5)$$

where $\langle \dots \rangle$ represents the ensemble average, Δx_{\max} is the largest value of Δx defined in Equation 4, and $P(\Delta x, \tau)$ is the PDF at scale τ . A set of scaling exponents ζ_q are then estimated from the (log-log) slopes of $SF_q(\tau)$ versus τ . We note that SFs defined as in Equation 5 bear many similarities with generalized variograms. For details regarding the generalized variogram analysis see, for example, Section 3.6.2 in the book by Rodríguez-Iturbe and Rinaldo (2001). When ζ_q varies linearly with the moment order q , the fluctuating variable is considered monofractal or self-similar. From the linearity of ζ_q curve we can also infer the “color of noise” of the investigated signal. For instance, if, besides the linearity, $\zeta_1 = 0.5$, the signal corresponds to a random walk noise (also called Brownian noise), while for $\zeta_1 = 0$ we are most probably faced with a white noise. We encourage users to explore such cases by using the synthetic data generator implemented in INA (see Section 3.1). When the linear relation between ζ_q and q is not satisfied, with ζ_q being a convex function of q , the fluctuating phenomenon is considered multifractal. Structure function analysis has a dedicated module in INA.

A quantitative measure of intermittency is the Flatness parameter $F(\tau)$, computed as (e.g., Bruno & Carbone, 2013):

$$F(\tau) = \frac{SF_4(\tau)}{SF_2^2(\tau)}. \quad (6)$$

A given time series is said to be intermittent if $F(\tau)$ continually increases toward smaller scales. Another time series can be said to be more intermittent if $F(\tau)$ grows faster. Gaussian fluctuations constitute a special case of non-intermittent signals, for which $F(\tau)$ is equal to 3. Computations of the Flatness parameter can be performed from the Probability Density module.

2.3. Rank-Ordered Multifractal Analysis (ROMA)

Because the conventional SF analysis reveals the statistics of the full set of fluctuations (which is dominated by that of the small amplitudes), see Equation 5, the physical interpretation of the multifractal nature cannot be easily understood by simply examining the curvatures of the deviations from the linearity between ζ_q and q . This means that we have to search for a way of isolating out the subdominant (large-amplitude) fluctuations and then perform the statistical analyses for each of the isolated populations. Such grouping of fluctuations must depend on their amplitude, but cannot depend merely on the raw values of the fluctuations because the amplitude ranges will be different for different scales.

These ideas led Chang and Wu (2008) to propose a new multifractal analysis technique, called ROMA, which is able to isolate the subdominant fluctuations and investigate their multifractal nature (see also Chang et al. (2010), Chang (2015), and Echim et al. (2021)).

Consider the physical variable $x(t)$, from which the scale dependent fluctuations $\Delta x(t, \tau)$ are computed as described in Equation 4. If the phenomenon represented by $x(t)$ is monofractal, then the PDFs, $P(\Delta x, \tau)$, would fully collapse onto one scaling function P_s , as follows:

$$P(\Delta x, \tau) \cdot (\tau/\tau_0)^s = P_s(\Delta x \cdot (\tau/\tau_0)^{-s}), \quad (7)$$

where s is the scaling exponent and τ_0 is a reference scale (see also Consolini and De Michelis (2011)). In practice, the PDFs only partially collapse, that is, parts of the PDFs collapse onto the master curve P_s while the rest remains unscaled. In this case it is considered that the dynamical process is multifractal and the scaling factor s depends on the scaled-sizes of fluctuations $s = s(Y)$, with $Y \equiv \Delta x \cdot (\tau/\tau_0)^{-s}$. Chang and Wu (2008) designed a way of grouping the fluctuations based on the values of the local fractal invariant Y . For a range ΔY within (Y_1, Y_2) , range-limited SFs are formed as follows:

$$SF_q^{(r)}(\Delta x, \tau) = \int_{a_1}^{a_2} (\Delta x)^q P(\Delta x, \tau) d(\Delta x) = \tau^{sq} \int_{Y_1}^{Y_2} Y^q P_s(Y) dY, \quad (8)$$

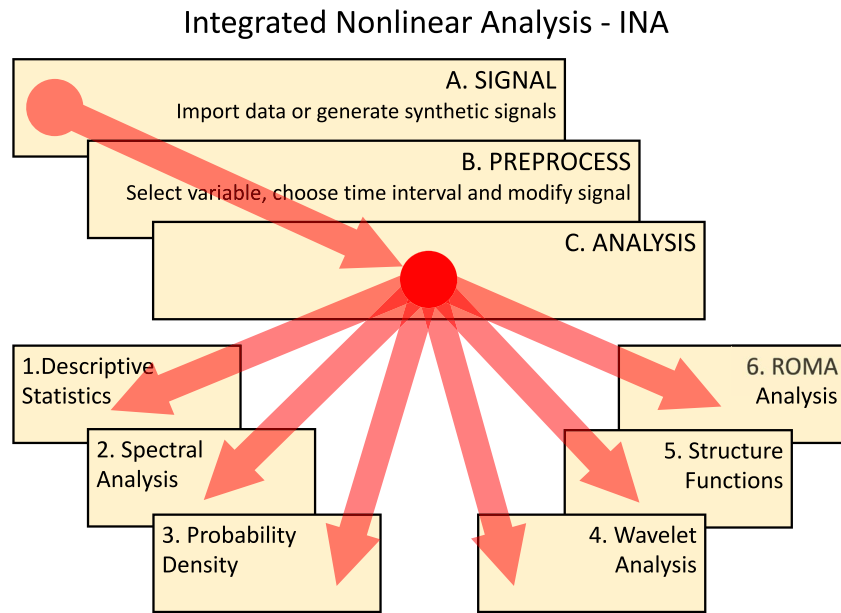


Figure 1. Diagram representation of the integrated nonlinear analysis toolbox. The program is structured into three layers (A, B, and C), and six analysis modules. The user is guided to follow a linear progression from data import to the analysis layer, which then acts as a “hub” connecting all of the analysis modules.

where $a_1 = Y_1 \cdot (\tau/\tau_0)^s$ and $a_2 = Y_2 \cdot (\tau/\tau_0)^s$. As in the conventional SF analysis, we may now search for the value of s such that the scaling property of the range-limited SF is: $SF_q^{(r)}(\Delta x, \tau) = \tau^{sq}$. If such a value of s exists, then we have found one region of the multifractal spectrum for which the PDFs in the range ΔY within Y_1 and Y_2 collapse onto one scaled PDF. Repeating this for all contiguous ranges of ΔY will generate the rank-ordered multifractal spectrum $s(Y)$.

The physical interpretation of the ROMA spectrum is based on the fact that s is equivalent to a local Hurst exponent for the subset of fluctuations in the given range ΔY . The Hurst exponent reveals the degree of persistency of a signal (e.g., Section 3.10 in Hergarten, 2002). Persistency means that a time series has a long-term tendency for positive variations (with respect to the mean value) to be followed by other positive variations, and vice-versa. At the other end, anti-persistency means that a time series has a long-term tendency for positive variations to be followed by negative variations, and so on. A signal is said to be persistent if it has a Hurst exponent between 0.5 and 1, and anti-persistent if its Hurst exponent is between 0 and 0.5. A Hurst exponent equal to 0.5 means that the signal fluctuations are completely uncorrelated, showing no long-term tendency.

3. Program Overview

Figure 1 depicts a diagram representation of INA. The program is structured into three layers: (A) SIGNAL, (B) PREPROCESS, and (C) ANALYSIS, and includes six data analysis modules: (1) Descriptive Statistics, (2) Spectral Analysis, (3) Probability Density, (4) Wavelet Analysis, (5) SFs, and (6) ROMA Analysis. INA adopts a strategy which helps the user to follow an increasingly complex logic, from data import to data analysis. Figure 1 also illustrates the branching of layer (C), which can be viewed as a hub that connects all of the analysis modules implemented in INA.

Figure 2 shows a screenshot of the INA HOME screen highlighting six areas, each providing a user interface for a specific scope. Area 1 allows access to HELP and PREFERENCES. The two export functionalities (SCREENSHOT and EXPORT) are also accessible from this panel. A set of buttons corresponding to the layers and modules mentioned in the first paragraph, is always displayed on the left part of the main user interface (Area 2). Using these buttons one can freely navigate between layers/modules and apply various analysis methods. The HOME and RESTART buttons are always displayed on the lower left corner of the main interface (Area 3), and offer two options to open the HOME screen: (a) without erasing any previous settings (imported signal, time selection, preprocessing option, etc.) and (b) by erasing all settings and RESTART the program. Two additional

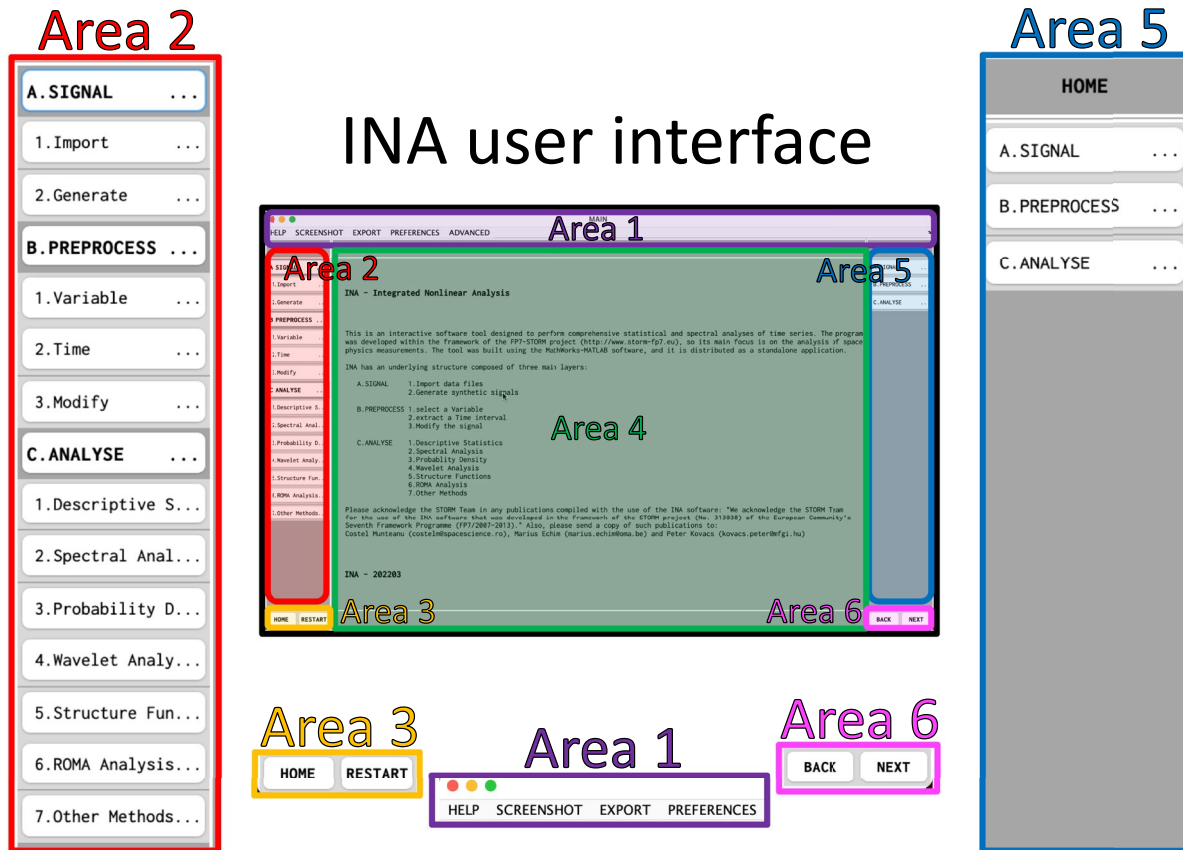


Figure 2. (central panel) Screenshot of the integrated nonlinear analysis HOME screen highlighting six areas of the graphical user interface. Area 1 contains links to HELP, SCREENSHOT, EXPORT, and PREFERENCES. Area 2 includes a static set of buttons accessible at any moment. Area 3 includes the HOME and RESTART buttons. Area 4 is used to depict analysis results. Area 5 includes user interface controls which change depending on the layer/module selected by the user. Area 6 contains the BACK and NEXT buttons (see the text for more details).

navigation buttons are always displayed on the lower right corner of the main interface (Area 6), labeled NEXT and BACK, and allow a user to advance (or go back) to the next (previous) layer/module, relative to the default linear progression from layer (A) to layer (C) and from analysis module (1) to module (6).

All the results generated by INA are displayed in Area 4. Note that in Figure 2, Area 4 shows the HOME screen. Most of the user interactions with INA take place in Area 5. This panel displays various user interface controls (buttons, editable text, dropdown menus, etc.), depending on the specific layer/module selected by the user, as described in the following sections.

3.1. Importing Data

The INA layer titled SIGNAL offers various options to import data for further analysis. This layer is organized in two sections: Import and Generate. The Import section includes several possible data sources: Measurement, File, and Other. Import-Measurement offers options to ingest in situ satellite observations and ground-based geomagnetic indices. Import-File allows one to import data formats like Common Data Format (CDF, <https://cdf.gsfc.nasa.gov>), TXT or MAT (binary MATLAB files, <https://www.mathworks.com/help/matlab/workspace.html>). Through Import-File-MAT, user generated binary files previously exported from MATLAB can be imported in INA. Using Import-File-CDF, most CDF files downloaded from CDAWeb can be imported in INA; this includes the very popular OMNI data sets, and also data from Solar Orbiter or Parker Solar Probe (<https://cdaweb.gsfc.nasa.gov>). Import-File-TXT allows a user to import custom generated text files formatted according to AMDA templates, in fact, most files downloaded from AMDA can be imported in INA using this functionality (<http://amda.irap.omp.eu>). Sample files that can be used as templates to create custom binary or text files for INA,

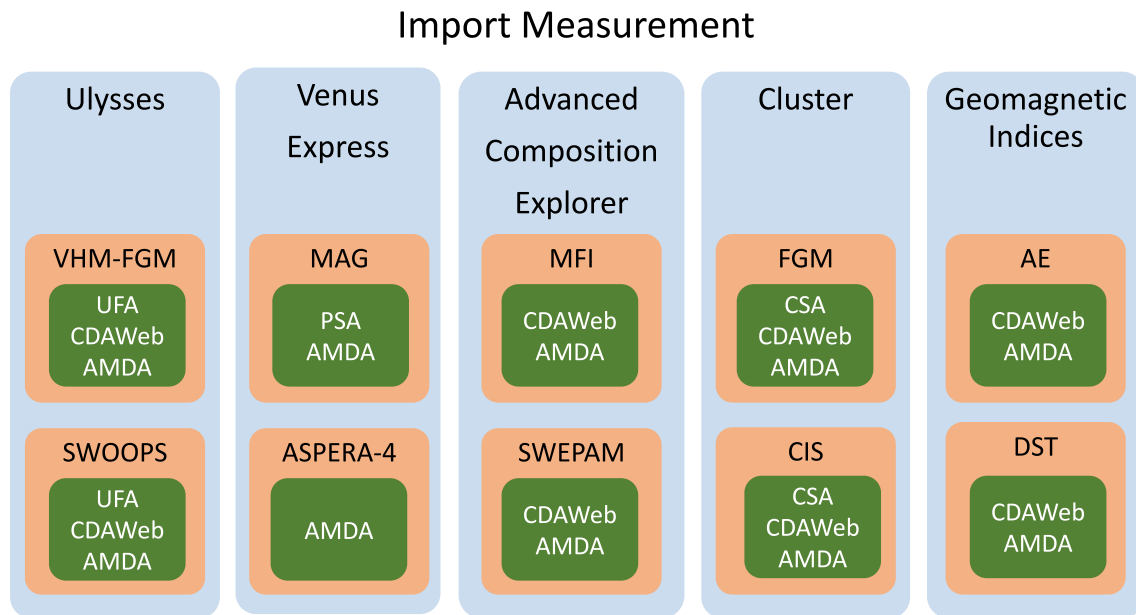


Figure 3. Diagram depicting the structure of the Import-Measurement section of integrated nonlinear analysis. From left to right, there are five branches divided into three color-coded levels: level 1—spacecraft name (blue); level 2—instrument acronym (orange); and level 3—data source (green).

are delivered to the user upon request (see Data Availability Statement). Import-Other includes links to custom data formats requested by users. Note that the program is still in development, and user requests for custom file formats can be added here.

Figure 3 depicts a diagram representation of the Import-Measurement section, which includes five branches: Ulysses, Venus Express (VEX), ACE, Cluster and Geomagnetic Indices; for each branch are defined three levels of content: satellite name, instrument name (acronym), and the source where the respective data can be downloaded. Note that in order to analyze data from the sources depicted in Figure 3, the user has to manually download the file from the respective source, and then import it into INA using the corresponding branch. The targeted spacecraft are: ACE (<https://www.nasa.gov/ace>), Ulysses, VEX and Cluster (<https://sci.esa.int/web/home/-/51459-missions>); the targeted instruments are VHM-FGM and SWOOPS, the magnetic field and plasma instruments onboard Ulysses; MAG and ASPERA-4, the magnetic field and plasma instruments onboard VEX; MFI and SWEPPAM, the magnetic field and plasma instruments onboard ACE; FGM and CIS, the magnetic field and plasma instruments onboard Cluster. INA was optimized and extensively tested to ingest data from these spacecraft.

In addition to in situ spacecraft measurements, INA is also designed to import ground-based measurements. Geomagnetic indices (branch 5 in Figure 3) are important elements in studies related to space weather and/or space climate. The current version of INA includes modules for analysis of Disturbance storm time (Dst) and Auroral Electrojet (AE) indices (<https://www.ngdc.noaa.gov/stp/geomag/indices.html>).

In Section 4 we present a case study to demonstrate the functionalities of INA. The case study uses a data file containing magnetic field measurements from the fluxgate magnetometer (FGM) onboard the Cluster 3 spacecraft, manually downloaded from the Cluster Science Archive (<https://www.cosmos.esa.int/web/csa>), the official data repository for the Cluster mission. In Figure 3 the acronyms denote the following data sources: UFA stands for Ulysses Final Archive (<http://ufa.esac.esa.int/ufo>), and PSA stands for Planetary Science Archive (<https://archives.esac.esa.int/psa>). We also include two alternative data sources: CDAWeb and AMDA.

In addition to external data, INA can also generate its own, customized synthetic signals. The program includes a comprehensive collection of synthetic signals which can be used to test the various analysis methods. For example, custom-generated sinusoidal signals can be used to test spectral analysis methods. Predefined synthetic signals with specific statistical properties are also included: random noise, signals with nonstationary features, etc. When generating custom synthetic signals, the user can set the signal length, add sinusoids with custom frequency, add random Gaussian noise, etc.

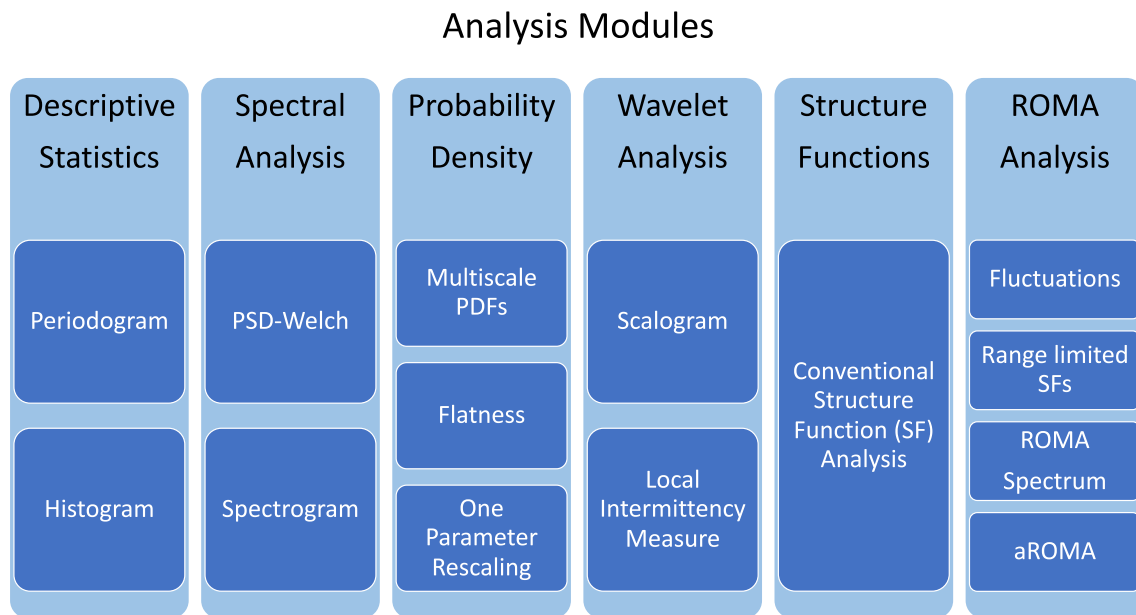


Figure 4. Diagram depicting the data analysis modules currently implemented in integrated nonlinear analysis (INA). From left to right, the INA analysis modules are: (a) Descriptive Statistics, (b) Spectral Analysis, (c) Probability Density, (d) Wavelet Analysis, (e) Structure Functions, and (f) rank-ordered multifractal analysis.

3.2. Preliminary Preprocessing

The second stage of INA data analysis cycle, after importing the data, is to apply one or several (optional) preprocessing procedures. These procedures are included in the preprocessing layer (B), comprising three sections: Variable, Time, and Modify. The Variable section lists all the variables available in the imported data file. In order to pursue the analysis, the user has to select a variable and then define a time interval. The time selection section offers two options: (a) interactively, using the cursor or (b) manually, by setting the time limits.

Layer (B) also includes a section labeled Modify. Most often, one would like to compare results from the analysis of two different signals or of two different time intervals from the same signal. In order to facilitate this comparison, the user often needs to standardize the two signals/intervals. Such a task is achieved using the option “standardize.” Mathematically, a signal y is standardized using the transformation defined by: $y_s = (y - \langle y \rangle) / \sigma_y$, where $\langle y \rangle$ is the mean value and σ_y is the standard deviation. This section also includes two options to handle data gaps: (a) fill-in with NaNs (Not a Number) and (b) linear interpolation. Note that for all branches of the Import-Measurement section of INA depicted in Figure 3, the data gaps are flagged by the data providers, and linear interpolation across data gaps is done by default. The option labeled “fill-in with NaNs,” allows one to keep the data gaps. As expected, this option cannot work with spectral analysis methods (because they require strict uniform time sampling), but it can work with statistical methods (multiscale PDFs and SFs). In some cases, for example, when dealing with very large data gaps, linear interpolation can introduce spurious/unrealistic data points, thus a statistical analysis performed without linear interpolation would be desirable.

3.3. Data Analysis Modules

The ANALYSIS layer acts as a hub connecting all the analysis modules implemented in INA. In the sections described previously, the user is guided along a rather linear pathway. This linearity breaks down in the data analysis layer, and one can choose to follow different pathways, depending on the specific interests. This branching of INA, together with the NEXT/BACK functionalities, allows one to apply the full set of analysis methods on the same signal, thus obtaining a complete picture of its spectral and statistical properties. The six analysis modules currently implemented in INA are illustrated in Figure 4.

The module titled Descriptive Statistics (analysis module 1) includes two methods: Periodogram and Histogram. This module can be used as a first degree estimate of some zero order spectral and statistical properties of the time series. Periodogram gives a nonparametric estimate of the PSD of the input signal (see Section 2.1). Histogram

displays a bar plot of the elements of the input signal, sorted into a number of equally spaced bins along the x -axis. INA implementations for Periodogram and Histogram are based on standard MATLAB functions.

Spectral Analysis (module 2) contains the methods titled PSD-Welch and Spectrogram. PSD-Welch estimates the PSD using the Welch method (see Section 2.1). INA PSD-Welch is based on the standard “pwelch” function, included in MATLAB Signal Processing Toolbox. In addition to the signal segmentation described in Section 2.1, PSD-Welch also offers the option to multiply each segment with a window function (e.g., Blackman, Hamming or Tukey window). Spectrogram method estimates the PSD for a series of time segments, thus it provides the PSD as a function of time. The result is presented as a three dimensional color plot (time-frequency-PSD) where the values of PSD are color coded. INA-Spectrogram is based on the standard “spectrogram” function included in MATLAB Signal Processing Toolbox. In INA PSD-Welch and Spectrogram, one can customize analysis parameters like: the window function, the segment length and the overlap between adjacent segments.

INA PSD-Welch offers an option called Slope Analysis. This functionality provides an interactive computation of PSD slopes from the log-log representation of PSD versus frequency. This can be accomplished with the help of three drop-down menus: (a) *fit*, where the user can choose the fitting method (linear or Levenberg-Marquardt fit), (b) *int*, where the user can choose to simultaneously fit one, two, or three different frequency ranges of the same PSD, and (c) *met*, where the user can choose how the frequency ranges are selected: using either the mouse or by typing explicit limits for the frequency interval. For future reference, the spectral indices (i.e., slopes) for the PSD-Welch results depicted in Figure 5, were obtained using linear fits (*fit* = “linfit”) over two frequency ranges (*int* = “two”).

Analysis module 3, titled Probability Density, includes three statistical approaches: PDF, Flatness, and OPR. The PDF method computes the multiscale PDFs, the flatness parameter F is defined in Equation 6 and gives a quantitative measure of intermittency, and OPR is the OPR method (see Section 2.2). The time differences $\Delta x(t, \tau)$ computed as described in Equation 4, requires the user to choose a range of time scales τ . By default, the smallest possible time scale is given by the time resolution of the signal, while the largest scale is limited by its length. The range of values for τ can be customized by the user. Estimating the PDFs involves the computation of normalized histograms for $\Delta x(t, \tau)$. The number of bins required by the histogram method can be adjusted by the user (using fewer bins can lead to an incorrect estimation of the shape of the PDFs, while using a very large number of bins increases the variability of the results). This INA module uses the standard “histogram” function from MATLAB.

Wavelet Analysis (module 4) can be used to compute scalograms and the associated LIM. The scalogram is a 3D representation of the logarithm of the (squared) wavelet coefficients, and LIM can be regarded as a normalized scalogram (see more in Section 2.1). Wavelet analysis involves the choosing of a wavelet mother function. For example, Torrence and Compo (1998) argue that the wavelet function should reflect the type of features present in a time series: for time series with sharp jumps, one should choose a function such as the Haar, while for smoothly varying time series one should choose a smoother wavelet function. The current version of the INA Wavelet Analysis module includes three wavelet families: *db* are the Daubechies' extremal phase wavelets, with *db1* being the Haar wavelet; *sym* are the Symlets family of wavelets, and *coif* stands for Coiflets. Depending on user feedback, we plan to expand this list (see, e.g., <https://www.mathworks.com/help/wavelet/gs/introduction-to-the-wavelet-families.html>). INA computations of wavelet coefficients uses the standard “cwt” function of MATLAB Wavelet toolbox.

SF (analysis module 5) can be used to perform a conventional SF analysis of the signal. The user can customize the range of time scales involved in SFs calculation, as described in Equation 5. The moment order q is another customizable parameter. In practice, SFs provide meaningful results only for $q > 0$, as they invariably diverge for $q < 0$; thus, the lower limit for q should be a positive number. The upper limit, q_{\max} , depends on the length of the signal. As a rule of thumb, Dudok de Wit et al. (2013) recommends the use of $q_{\max} = \log(N) - 1$, where N is the number of samples in the data set (see also Teodorescu et al. (2021)).

Rank Ordered Multifractal Analysis (ROMA, INA module 6) includes four methods: Fluctuations, Range-limited SFs, ROMA Spectrum, and aROMA. ROMA is a complex analysis method, thus the output, that is, the ROMA multifractal spectrum, must be supplemented by multiple preliminary tests and analyses in order to understand and validate the results. Analysis steps like the ones grouped under the labels “Fluctuations” and “Range-limited SFs” are used for such preliminary purposes. The functions ROMA Spectrum and aROMA on the other hand provide the ROMA spectrum itself using two slightly different approaches. A brief mathematical background on ROMA is provided in Section 2.3.

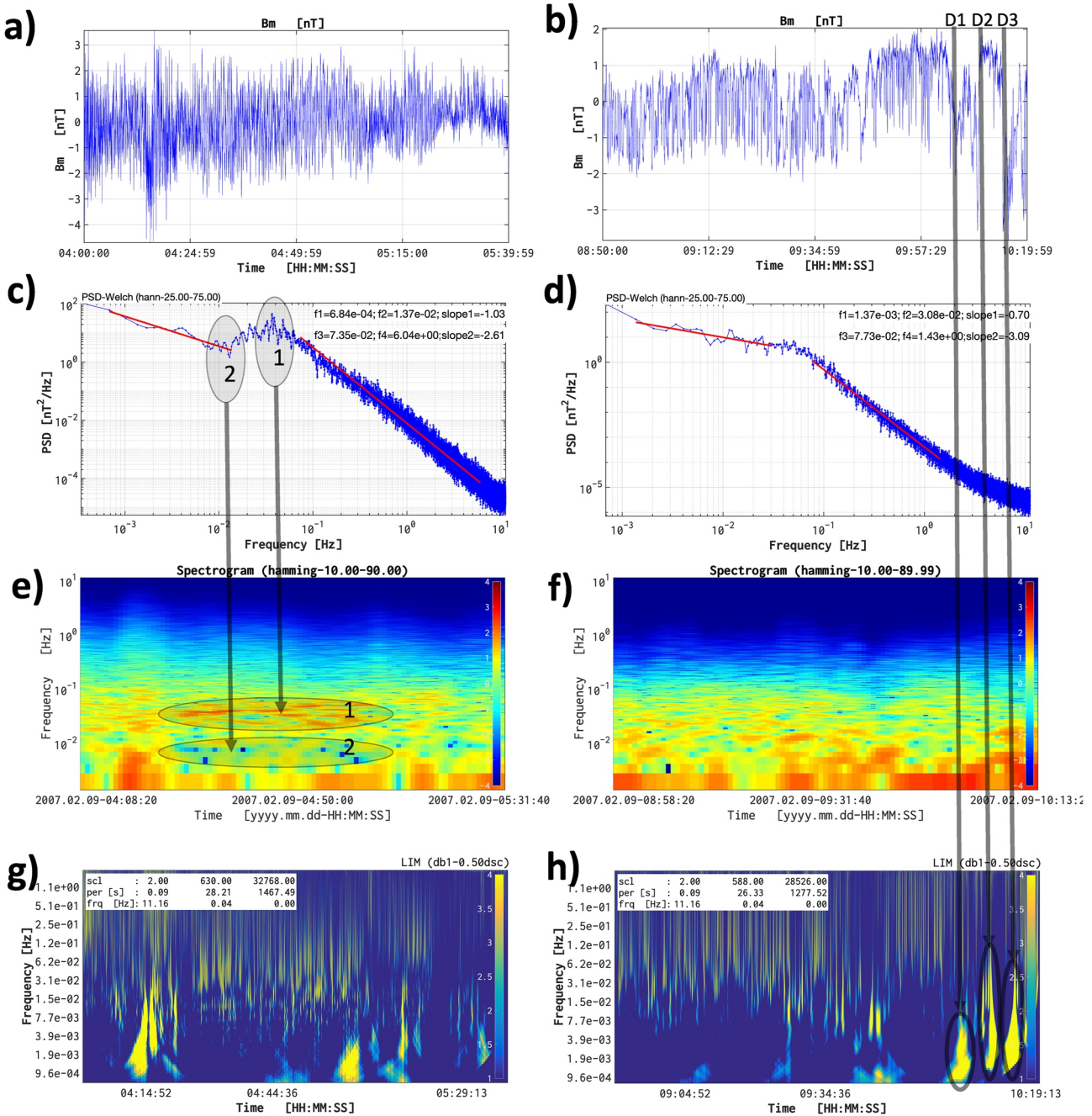


Figure 5. Spectral analysis using integrated nonlinear analysis of the magnetic field intensity (B_m) measured by Cluster 3 in the magnetosheath on 9 February 2007. Column 1: (a) time series representation of (B_m) for an interval close to the bowshock. (c) PSD-Welch, (e) Fourier Spectrogram, and (g) wavelet local intermittency measure. Column 2 shows the corresponding results for an interval close to the magnetopause. Panels (c and d) depict a series of red lines superposed on the power spectral density (PSD)-Welch results (which are depicted in blue); these are linear fits of the corresponding PSD values, and were generated using the “Slope analysis” functionality; frequency ranges and slopes for the linear fits are displayed in the top-right corner of these panels.

3.4. Exporting Results From INA

There are two main ways to export the graphical results obtained with INA. The easiest and most straightforward option is to take a screenshot. INA offers a dedicated screenshot option, accessible through the corresponding button in area 1 (see Figure 2). INA uses a default naming convention for the screenshots, which are saved as jpg files in the working directory. Of course, screenshots can also be taken using the built-in options of the operating

system, or other third party applications. Figure 2, for example, was generated using the built-in screenshot functionality of the macOS operating system.

The second, and recommended option to generate publication-quality figures is the EXPORT functionality, also accessible from area 1 illustrated in Figure 2. Note that each individual plot in Figures 5–8 was generated using this option. The button labeled EXPORT opens a new window (not illustrated here), which includes a series of editable text fields, drop-down menus and buttons, and can be used to set the file name, figure type and export folder. Each export window also includes a series of checkboxes which can be used to select which plots and data sets to be exported. The content of this window depends on the layer/module from which it was called from.

4. An Illustrative Example

In this section we provide an example which demonstrates most of INA capabilities. We analyze the magnetic field intensity measured by Cluster 3 spacecraft during a magnetosheath inbound pass on 09 February 2007. We use high (22 Hz) time resolution data from the fluxgate magnetometer (FGM; Balogh et al., 2001) onboard Cluster 3. Two time intervals, each 1.5 hr long, were selected, one in the vicinity of the bowshock and the second one close to the magnetopause. We deployed on these data the full suit of analysis methods implemented in INA. For brevity, we will sometimes denote the bowshock measurements as interval 1, and the magnetopause ones as interval 2. Figures 5–8, show the data analysis results. Figure 5 presents the results of spectral analysis, Figure 6 shows the multiscale analysis of fluctuations, Figure 7 illustrates the results of SF analysis, and Figure 8 shows results from the ROMA analysis. Each one of these figures has a two-column layout: the left column shows the results corresponding to the bowshock interval, and the right-column depicts the magnetopause-interval results.

The two selected intervals are illustrated in Figures 5a and 5b. Note that the two signals were standardized, that is, we remove the mean and divide by the standard deviation, in order to facilitate their intercomparison (see Section 3.2). After standardization both series are centered around 0 nT and have a standard deviation equal to 1 nT.

The results of the PSD-Welch method are depicted in Figures 5c and 5d, and indicate the presence of two spectral regions characterized by different power laws, for both intervals. These spectral regions, and the corresponding slope values, that is, spectral indices, were determined using the INA PSD-Welch “Slope Analysis” (see Section 3.3). At low frequencies, the near-bowshock power spectrum scales as $f^{-\alpha}$, with the spectral index α equal to one; a slightly shallower spectrum ($f^{-0.7}$) is found for the near-magnetopause interval. At higher frequencies, the spectrum scales approximately as $f^{-8/3}$ near the bowshock; a slightly steeper scaling ($f^{-3.09}$) is observed at the magnetopause. Similar findings related to magnetosheath turbulence spectrum are also reported by Huang et al. (2017) (see also Teodorescu and Echim (2020)). Damping of kinetic Alfvén waves through Landau resonances is invoked as a possible cause of the steepening of the high-frequency spectrum (Howes et al., 2008).

Figures 5e and 5f depict the Fourier spectrograms for the two intervals. These results suggest that both signals are nonstationary, that is, their spectra changes in time. Nonstationarity should cast doubt, or even invalidate the results obtained using the PSD-Welch analysis. Two features dominate the PSD of the bowshock magnetic field depicted in Figure 5c: (a) a frequency band of increased power between 10^{-2} and 10^{-1} Hz and (b) a frequency band of decreased power around 10^{-2} Hz. The Fourier spectrogram depicted in Figure 5e clearly shows that these two features are in fact localized in time. The Spectrogram shown in Figure 5e suggests that the time-frequency bands highlighted here are responsible for the corresponding features observed in the PSD presented in Figure 5c.

Wavelet analysis was also applied on these two intervals. Figures 5g and 5h show the LIM, that is, normalized wavelet scalograms, for the two intervals. The results indicate that both time series contain strong intermittent events, identified by localized bright yellow regions. Note that the LIM color scale varies from 1 to 4, meaning that the brightest features correspond to regions where the local PSD is (at least) four times larger than the background power. Multiple intermittent events can be found in both intervals, see, for example, the three strong discontinuities observed in interval 2 between 09:57 and 10:19 UT, labeled in Figure 5b as (D1), (D2), and (D3). One notes that these rapid signal variations correspond to the three localized bright yellow regions highlighted in Figure 5h.

Figure 6 depicts the results of the multiscale statistical analysis of the two intervals, adding supplementary information to the results provided by the spectral analysis above. A typical study involving multiscale PDFs implies

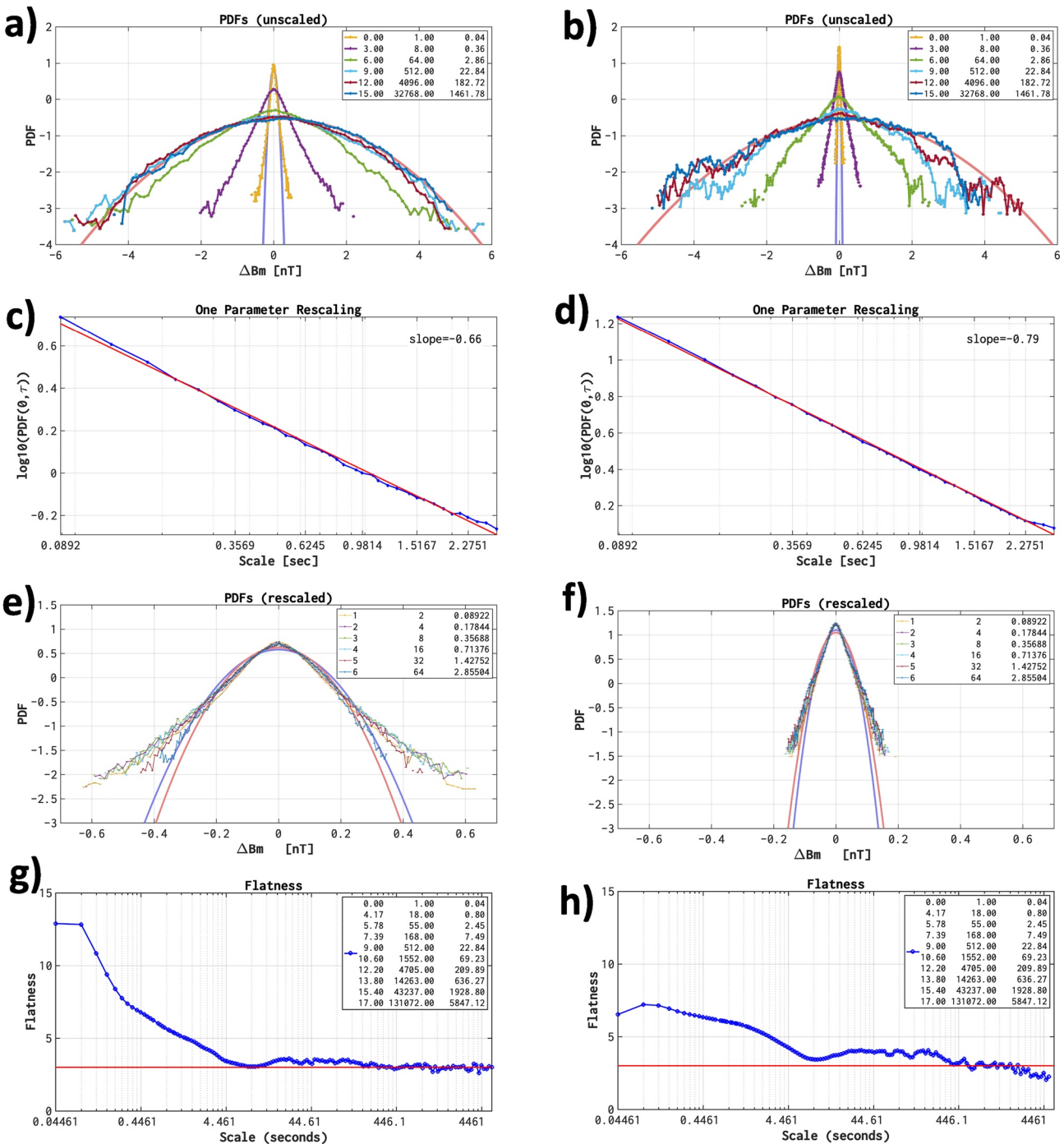


Figure 6. Multiscale study of fluctuations using integrated nonlinear analysis. Column 1 shows results for the bowshock interval: (a) unscaled probability density functions (PDFs) for $\Delta B_m(t, \tau)$ (see Equation 4, with x substituted by B_m), (c) One Parameter Rescaling, (e) rescaled PDFs, and (f) Flatness, defined as in Equation 6. Column 2 shows the corresponding results for the magnetopause interval. The legends show the scales τ in powers of two (first column), in number of points (second column) and in seconds (third column). The rugged lines in panels (a, b, e, and f), depict the fluctuation PDFs; smooth lines are also depicted: the blue one represents the Gaussian PDF corresponding to the smallest scale fluctuation PDF, and the red one represents the Gaussian PDF corresponding to the largest scale fluctuation PDF. In panels (g and h), the Flatness values for Gaussian distribution (equal to 3 for all scales) are marked with horizontal red lines.

a search for departures from Gaussianity, in other words, we search for those scales for which the fluctuation PDFs have different shapes than the corresponding Gaussian shapes; this signifies a possible presence of intermittency (see, e.g., Figure 5.10 in Echim et al., 2021). The multiscale PDFs in Figures 6a and 6b show significant departures from Gaussianity for small scales (≤ 2.86 s), suggesting the presence of intermittency. Larger scales

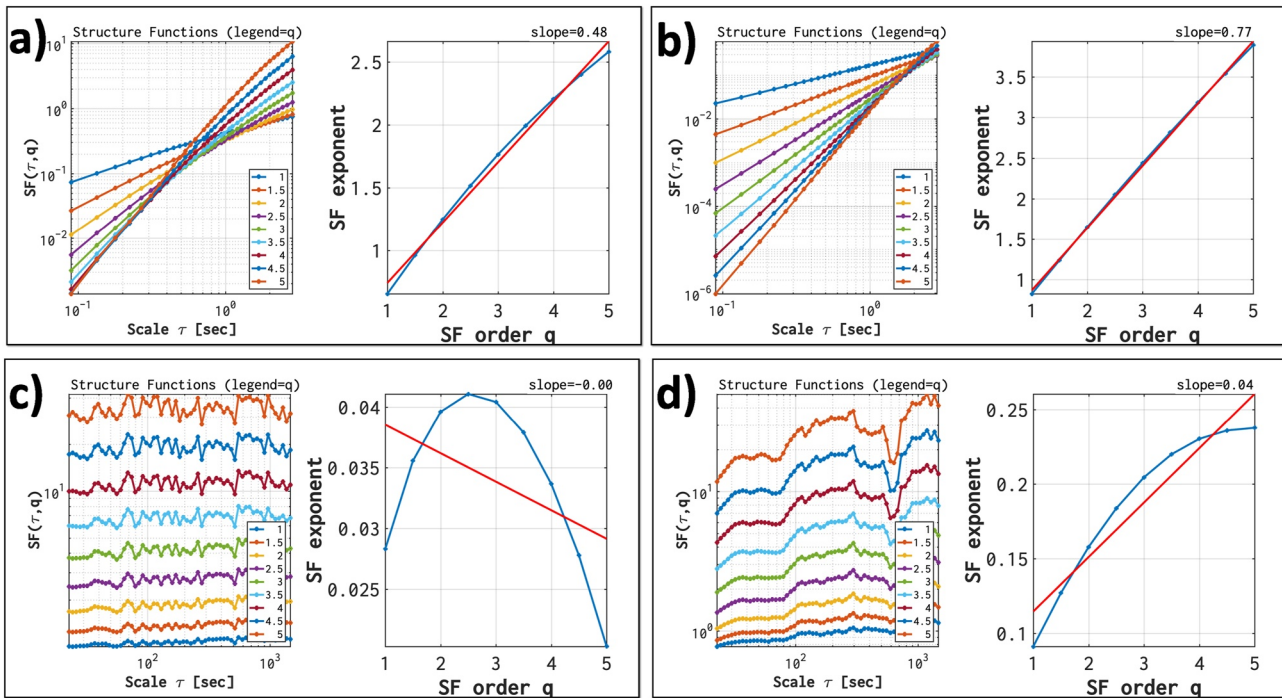


Figure 7. Structure Function analysis using integrated nonlinear analysis. Column 1 shows results for the bowshock interval: (a) small scales (τ from 0.08 to 2.86 s), and (c) large scales (τ from 22.84 to 1,928.80 s). Column 2 shows the corresponding results for the magnetopause interval. In each panel, the left-side depicts $SF_q(\tau)$, defined as in Equation 5, using the values for q shown in the legend; the right-side depicts the scaling exponents $\zeta_q(q)$ (in blue) and the corresponding linear fit (in red); the slope of the linear fit is shown in the top-right corner.

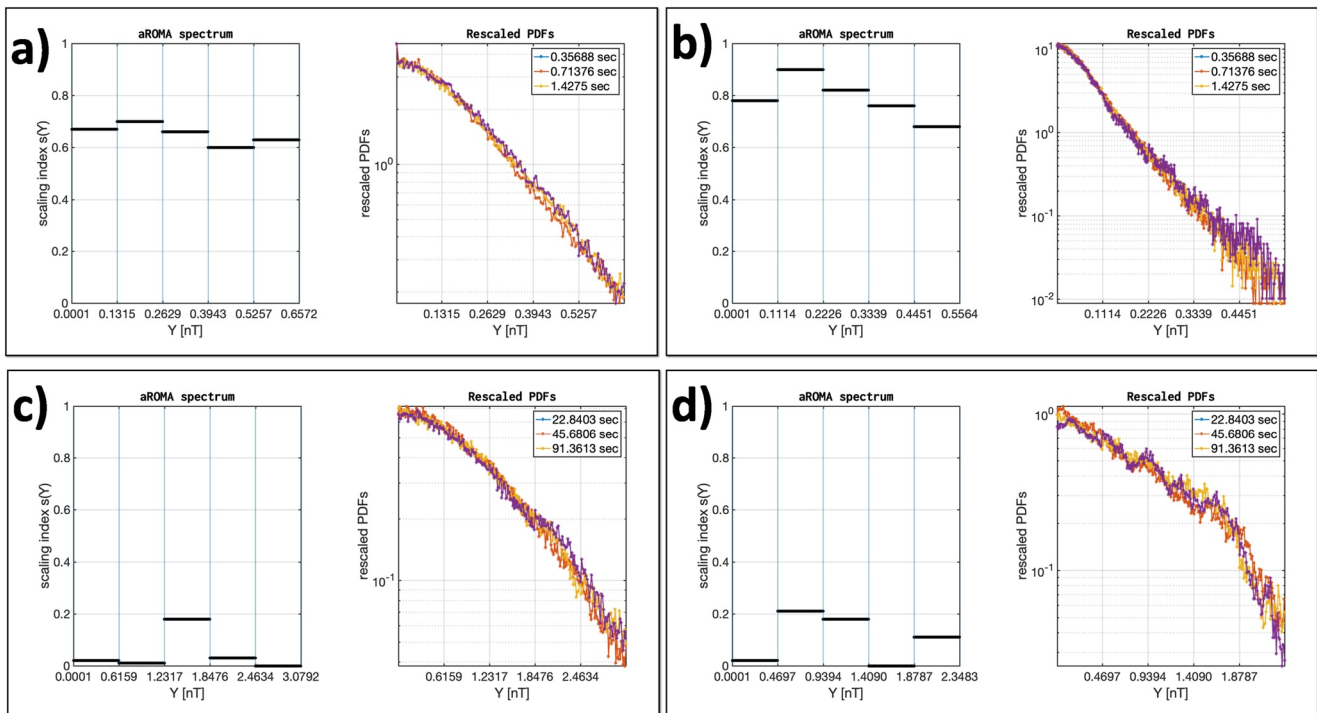


Figure 8. Rank Ordered Multifractal Analysis (ROMA) using integrated nonlinear analysis. Column 1 shows results for the bowshock interval: (a) small scales and (c) large scales. Column 2 shows the corresponding results for the magnetopause interval. In each panel, the left-side depicts the ROMA spectrum $s(Y)$, and the right-side includes an illustration of the rescaled probability density functions, where the legends show the scales τ used as input in Equation 8.

(≥ 22.84 s), on the other hand, closely follow a Gaussian shape. For the smallest scale depicted here (0.04 s), the large-amplitude fluctuations appear to be slightly further away from a Gaussian shape for interval 1 compared to interval 2; this implies that interval 1 has a slightly higher degree of intermittency compared to interval 2.

Figure 6 also depicts results from the OPR technique. OPR can be used to estimate the rescaling index, by computing the slope of PDF maxima versus scale (e.g., Hnat et al., 2002). For interval 1 (Figure 6c) the scaling index (denoted in the figure as “slope”) is 0.66; and for interval 2 (Figure 6d), the index is 0.79. The accuracy of the scaling index estimated using OPR, is verified by rescaling the PDFs. The results, depicted in Figures 6e and 6f, show that OPR gives better rescaling results for interval 2 compared to interval 1. Note that OPR is able to rescale only the small amplitude fluctuations, the large ones rest unscaled suggesting the system is multifractal (Chang & Wu, 2008). According to the interpretation given in Section 2.3, the scaling index is equivalent to the Hurst exponent of the small amplitude fluctuations. The OPR results suggest that both intervals are characterized by persistency, that is, the Hurst exponent is larger than 0.5, meaning that there is a long-term tendency for positive variations (with respect to the mean value) to be followed by other positive variations, or vice-versa for negative variations. Also, the index is larger for interval 2, meaning that interval 2 has a higher degree of persistence compared to interval 1.

The results extracted from the multiscale PDF analyses are confirmed by the Flatness depicted in Figures 6g and 6h. Flatness increases toward smaller scales, which is considered a signature of intermittency (Bruno & Carbone, 2013); also, the Flatness value for the smallest scale is larger for interval 1 compared to interval 2, implying that interval 1 has a higher degree of intermittency compared to interval 2. From the Flatness analysis we also observe that the intermittent structures, that is, Flatness values larger than Gaussian value of 3, pertain to the same range of scales in both intervals.

The results of the SF analysis are depicted in Figure 7. These results confirm the presence of two scale domains with different scaling properties, similar to the results obtained with Flatness. For the small scales region, the SF exponent shows a slightly non-linear variation with the moment order q for interval 1 (Figure 7a), signifying multifractality, and an approximately linear variation for interval 2 (Figure 7b), that is, quasi-monofractal process. The slope of the SF exponent versus q is 0.48 for interval 1 and 0.77 for interval 2. For the large scales region, the slope of the SF exponent versus q is 0 for interval 1 (Figures 7c) and 0.04 for interval 2 (Figure 7d). These results show a clear anti-persistent character for the large-scale fluctuations. This means that an increase will most likely be followed by a decrease and vice-versa. In other words, future values will tend to return to a long-term average value.

The multifractality of magnetosheath fluctuations at small scales revealed by the SF analysis is confirmed by an analysis with the ROMA method. Some of the results obtained using ROMA are depicted in Figure 8. The scale intervals chosen for the analysis are the same as those used in Figure 7. Figures 8a and 8b depict the small-scale ROMA multifractal spectra, that is, the multifractal index s versus the rescaled variable Y , for the two intervals. Figures 8c and 8d show the corresponding results for the large scales. When the scaling indices computed with ROMA are applied to the multiscale PDFs (see Equation 7), one obtains the rescaled PDFs depicted at the right-side of each ROMA-spectrum. Figure 8 shows some differences between the bowshock and the magnetopause interval. At small scales both fluctuations are persistent ($s > 0.5$). At these scales, kinetic effects are believed to be responsible for the persistence. Close to bowshock, a decrease at large rescaled amplitudes Y toward a Gaussian value of 0.5 is observed. At the magnetopause, a similar decrease is observed, but the scaling index remains large ($s > 0.65$) even for the largest rescaled amplitudes. Also, the overall-higher scaling index for magnetopause, correlates very well with the steeper slope of the PSD. At large scales the fluctuations are anti-persistent ($s < 0.5$) for both intervals, with the near-bowshock fluctuations appearing more monofractal, that is, showing a constant value for $s(Y)$, compared to the near-magnetopause fluctuations.

5. Summary and Conclusions

We built, tested and released a software analysis toolbox -INA-designed to study space plasma turbulence and intermittency. The toolbox was developed in MATLAB, but it is distributed as a standalone executable file which can be freely shared and used by any interested user. The toolbox can ingest various satellite and ground-based measurements, and includes a comprehensive collection of analysis methods. The software is designed to speed up and enhance the scientific output by offering a straightforward way from spacecraft data import to exporting analysis results, thus significantly reducing the time necessary for data preprocessing or software implementation

of analysis methods. It can also facilitate the collaboration between scientists with common research interests and can be a suitable environment for sharing and discussing scientific results.

One of the main distinctive features of INA is the integrated, interactive, and user-friendly environment in which one can import a data set, customize key analysis parameters, apply multiple methods on the same signal and then produce high-quality graphical results. Other distinctive features include: the versatility to import data from multiple sources, the rich collection of analysis methods, and the detailed configuration options that allow a fine customization of analysis parameters and graphical presentation of results.

Data from specific spacecraft missions such as Venus Express, Ulysses, ACE or Cluster, can be processed through INA while other types of data sets can be easily added. For example, using the Import-File-CDF functionality, most CDF files downloaded from CDAWeb can be imported in INA; this includes the very popular OMNI data sets, and also data from spacecraft like Solar Orbiter or Parker Solar Probe. Considerable effort has been invested in data preprocessing such that INA can guarantee output at the highest scientific quality. In this regard, several approaches have been implemented to manage data gaps. The analysis methods in INA span a wide range, from descriptive statistics, spectral analysis (both Fourier and wavelet-based methods), statistical analysis of fluctuations (multiscale PDFs and their moments), to multifractal analysis based on the newly developed ROMA method. INA offers an easy access to these advanced analysis and visualization methods, and can be used by experienced scientists and also as a training tool for students getting accustomed to both methodology and technical aspects of a comprehensive data analysis approach for the study of turbulence and intermittency in space plasmas.

An illustrative example is provided, demonstrating the performance of INA when applied to a specific data set. We analyzed magnetic field intensity measured by the Cluster 3 spacecraft during an inbound pass through the Earth's magnetosheath on 09 February 2007. Two time intervals were selected and compared to each other, one near the bowshock and the other one near the magnetopause. For the near-bowshock interval, we showed how specific features of the power spectrum can be mapped to localized time-frequency regions in the spectrogram representation, demonstrating that the interval is nonstationary. For the near-magnetopause interval, a series of intermittent events, that is, regions where local power is larger than background power, identified in the wavelet-based LIM representation, signified that this interval is also nonstationary.

Multiscale PDFs and higher order moments of fluctuations were used to study intermittency. Small scale PDFs showed a clear departure from Gaussian shapes, signifying the presence of intermittency in both intervals, with the near-bowshock interval appearing slightly more intermittent than the near-magnetopause one. This result was confirmed by the analysis of the Flatness parameter. The PDFs of large scale fluctuation were shown to be Gaussian. SF analysis also revealed the presence of two scale domains with different properties. For the small scale domain of the near-bowshock interval, a clear nonlinear variation of SF exponent with moment order q is found, signifying multifractality, while the same small scale region is approximately linear for the near-magnetopause interval, implying quasi-monofractality. The multifractality of magnetosheath fluctuations is also confirmed by the ROMA results. Using ROMA, small scale fluctuations were shown to be persistent for both intervals ($s > 0.5$), with an overall higher degree of persistence for the near-magnetopause interval, compared to the near-bowshock interval.

We showed that INA puts a lot of emphasis on cross-validating analysis results. For example, multiple spectral analysis methods are implemented in INA, and are capable of reaching comparable results, for example, via Fourier spectrograms and wavelet scalograms. Multiscale PDFs can be used to estimate the scaling exponent, but so do SF analysis, and ROMA. Many of the methods implemented in INA are also complementary to each other. For example, multifractal analysis can be used to study monofractal signals, but can also reveal multifractality, qualitatively using SFs and also quantitatively, by providing the range limited scaling exponent (ROMA).

To our knowledge, INA is the only software tool providing such an integrated, interactive and user-friendly environment, including a comprehensive set of both low- and high-order analysis methods dedicated to study turbulence and intermittency in space plasmas, and distributed for free to any interested user.

Data Availability Statement

The INA standalone application and the MATLAB source code for INA are publicly available for download from the STORM-FP7 project website: <http://www.storm-fp7.eu/index.php/data-analysis-tools>. In Section 4 we used magnetic field measurements from the FGM instrument (Balogh et al., 2001) on-board Cluster 3 spacecraft. We

used high time resolution (22 Hz) data measured on 9 February 2007, available from: <https://www.cosmos.esa.int/web/csa/access>.

Acknowledgments

This work was supported by the European Community's Seventh Framework Programme (FP7) under Grant 313038/STORM; Romanian Ministry of Research and Innovation via PCCDI Grant 18PCCDI/2018; and National Core Program (LAPLAS 2019). The work of M. E. was supported by ESA PRODEX CLUSTER, BRAIN-be 2.0 Grant B2/223/P1/PLATINUM, and Belgian Solar Terrestrial Center of Excellence (STCE). The work of C. M. was supported by ESA PRODEX MISION, and National projects PN-III-P1-1.1-TE-2019-1288 and PN-III-P1-1.1-TE-2021-0102. P. K. acknowledges the support of the Hungarian National Research, Development and Innovation Fund (OTKA, Project FK128548).

References

- Alexandrova, O., Chen, C., Sorriso-Valvo, L., Horbury, T., & Bale, S. (2013). Solar wind turbulence and the role of ion instabilities. *Space Science Reviews*, 178(2–4), 101–139. <https://doi.org/10.1007/s11214-013-0004-8>
- Balogh, A., Carr, C. M., Acuña, M. H., Dunlop, M. W., Beek, T. J., Brown, P., et al. (2001). The cluster magnetic field investigation: Overview of in-flight performance and initial results. *Annales Geophysicae*, 19(10/12), 1207–1217. <https://doi.org/10.5194/angeo-19-1207-2001>
- Benzi, R., Ciliberto, S., Tripiccone, R., Baudet, C., Massaioli, F., & Succi, S. (1993). Extended self-similarity in turbulent flows. *Physical Review E*, 48(1), R29–R32. <https://doi.org/10.1103/PhysRevE.48.R29>
- Biskamp, D. (1993). *Nonlinear magnetohydrodynamics*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511599965>
- Bloomfield, P. (2000). *Fourier analysis of time series: An introduction*. Wiley Series in Probability and Statistics.
- Bruno, R. (2019). Intermittency in solar wind turbulence from fluid to kinetic scales. *Earth and Space Science*, 6(5), 656–672. <https://doi.org/10.1029/2018EA000535>
- Bruno, R., & Carbone, V. (2013). The solar wind as a turbulence laboratory. *Living Reviews in Solar Physics*, 10, 1–208. <https://doi.org/10.1294/lrsp-2013-2>
- Bruno, R., Carbone, V., Veltri, P., Pietropaolo, E., & Bavassano, B. (2001). Identifying intermittency events in the solar wind. *Planetary and Space Science*, 49(12), 1201–1210. [https://doi.org/10.1016/s0032-0633\(01\)00061-7](https://doi.org/10.1016/s0032-0633(01)00061-7)
- Burlaga, L. F. (1991). Intermittent turbulence in the solar wind. *Journal of Geophysical Research*, 96(A4), 5847–5851. <https://doi.org/10.1029/91JA00087>
- Chang, T. (2015). *An introduction to space plasma complexity*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511980251>
- Chang, T., & Wu, C. C. (2008). Rank-ordered multifractal spectrum for intermittent fluctuations. *Physical Review E*, 77(4), 045401. <https://doi.org/10.1103/PhysRevE.77.045401>
- Chang, T., Wu, C. C., Podesta, J., Echim, M., Lamy, H., & Tam, S. W. Y. (2010). ROMA (Rank-Ordered Multifractal Analyses) of intermittency in space plasmas - A brief tutorial review. *Nonlinear Processes in Geophysics*, 17(5), 545–551. <https://doi.org/10.5194/npg-17-545-2010>
- Consolini, G., & De Michelis, P. (2011). Rank ordering multifractal analysis of the auroral electrojet index. *Nonlinear Processes in Geophysics*, 18(3), 277–285. <https://doi.org/10.5194/npg-18-277-2011>
- Daubechies, I. (1992). *Ten lectures on wavelets*. CBMS-NSF Regional Conference Series in Applied Mathematics. <https://doi.org/10.1137/1.9781611970104>
- Deak, N., Creț, O., Echim, M., Teodorescu, E., Negrea, C., Văcariu, L., et al. (2018). Edge computing for space applications: Field programmable gate array-based implementation of multiscale probability distribution functions. *Review of Scientific Instruments*, 89(12), 125005. <https://doi.org/10.1063/1.5044425>
- Deak, N., Creț, O., Munteanu, C., Teodorescu, E., & Echim, M. M. (2021). FPGA design for on-board measurement of intermittency from in-situ satellite data. *Earth and Space Science*, 8(12), e2021EA001678. <https://doi.org/10.1029/2021EA001678>
- Dudok de Wit, T., Alexandrova, O., Furno, I., Sorriso-Valvo, L., & Zimbardo, G. (2013). Methods for characterising microphysical processes in plasmas. *Space Science Reviews*, 178(2), 665–693. <https://doi.org/10.1007/s11214-013-9974-9>
- Echim, M., Chang, T., Kovacs, P., Wawrzaszek, A., Yordanova, E., Narita, Y., et al. (2021). Turbulence and complexity of magnetospheric plasmas. In *Magnetospheres in the solar system* (pp. 67–91). American Geophysical Union (AGU). <https://doi.org/10.1002/9781119815624.ch5>
- Echim, M., Lamy, H., & Chang, T. (2007). Multi-point observations of intermittency in the cusp regions. *Nonlinear Processes in Geophysics*, 14(4), 525–534. <https://doi.org/10.5194/npg-14-525-2007>
- Farge, M. (1992). Wavelet transforms and their applications to turbulence. *Annual Review of Fluid Mechanics*, 24(1), 395–457. <https://doi.org/10.1146/annurev.fl.24.010192.002143>
- Goldreich, P., & Sridhar, S. (1995). Toward a theory of interstellar turbulence. ii. strong Alfvénic turbulence. *The Astrophysical Journal*, 438, 763. <https://doi.org/10.1086/175121>
- Greco, A., Matthaeus, W. H., Perri, S., Osman, K. T., Servidio, S., Wan, M., & Dmitruk, P. (2018). Partial variance of increments method in solar wind observations and plasma simulations. *Space Science Reviews*, 214(1), 1. <https://doi.org/10.1007/s11214-017-0435-8>
- Hergarten, S. (2002). *Self-organized criticality in earth systems*. Springer. <https://doi.org/10.1007/978-3-662-04390-5>
- Hnat, B., Chapman, S. C., Rowlands, G., Watkins, N. W., & Farrell, W. M. (2002). Finite size scaling in the solar wind magnetic field energy density as seen by wind. *Geophysical Research Letters*, 29(10), 86–1–86. <https://doi.org/10.1029/2001GL014587>
- Howes, G. G., Cowley, S. C., Dorland, W., Hammett, G. W., Quataert, E., & Schekochihin, A. A. (2008). A model of turbulence in magnetized plasmas: Implications for the dissipation range in the solar wind. *Journal of Geophysical Research*, 103, A05103. <https://doi.org/10.1029/2007JA012665>
- Huang, S. Y., Hadid, L. Z., Sahraoui, F., Yuan, Z. G., & Deng, X. H. (2017). On the existence of the Kolmogorov inertial range in the terrestrial magnetosheath turbulence. *The Astrophysical Journal Letters*, 836(1), L10. <https://doi.org/10.3847/2041-8213/836/1/L10>
- Lion, S., Alexandrova, O., & Zaslavsky, A. (2016). Coherent events and spectral shape at ion kinetic scales in the fast solar wind turbulence. *The Astrophysical Journal*, 824(1), 47. <https://doi.org/10.3847/0004-637X/824/1/47>
- Marsch, E., & Tu, C. Y. (1994). Non-Gaussian probability distributions of solar wind fluctuations. *Annales Geophysicae*, 12(H15), 1127–1138. <https://doi.org/10.1007/s00585-994-1127-8>
- Matthaeus, W. H., Wan, M., Servidio, S., Greco, A., Osman, K., Oughton, S., & Dmitruk, P. (2015). Intermittency, nonlinear dynamics and dissipation in the solar wind and astrophysical plasmas. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 373(2041), 20140154. <https://doi.org/10.1098/rsta.2014.0154>
- Munteanu, C., Turicu, D. C., Creț, O., & Echim, M. (2022). Detecting discontinuities from in situ space measurements: Method and FPGA implementation. *Earth and Space Science*, 9(10), e2022EA002537. <https://doi.org/10.1029/2022EA002537>
- Opincariu, L., Deak, N., Creț, O., Echim, M., Munteanu, C., & Văcariu, L. (2019). Edge computing in space: Field programmable gate array-based solutions for spectral and probabilistic analysis of time series. *Review of Scientific Instruments*, 90(11), 114501. <https://doi.org/10.1063/1.5119231>
- Roberts, O. W., Alexandrova, O., Kajdic, P., Turc, L., Perrone, D., Escoubet, C. P., & Walsh, A. (2017). Variability of the magnetic field power spectrum in the solar wind at electron scales. *The Astrophysical Journal*, 850(2), 120. <https://doi.org/10.3847/1538-4357/aa93e5>
- Rodríguez-Iturbe, I., & Rinaldo, A. (2001). *Fractal river basins: Chance and Self-Organization*. Cambridge University Press. <https://cambridge.org/9780521004053>

- Schmitt, F. G., & Huang, Y. (2016). *Stochastic analysis of scaling time series: From turbulence theory to applications*. Cambridge University Press. <https://doi.org/10.1017/CBO9781107705548>
- Sorriso-Valvo, L., Carbone, V., Giuliani, P., Veltri, P., Bruno, R., Antoni, V., & Martinez, E. (2001). Intermittency in plasma turbulence. *Planetary and Space Science*, 49(12), 1193–1200. [https://doi.org/10.1016/S0032-0633\(01\)00060-5](https://doi.org/10.1016/S0032-0633(01)00060-5)
- Sreenivasan, K. R. (1991). Fractals and multifractals in fluid turbulence. *Annual Review of Fluid Mechanics*, 23(1), 539–604. <https://doi.org/10.1146/annurev.fl.23.010191.002543>
- Teodorescu, E., & Echim, M. M. (2020). Open-source software analysis tool to investigate space plasma turbulence and nonlinear DYNamics (ODYN). *Earth and Space Science*, 7(4), e2019EA001004. <https://doi.org/10.1029/2019EA001004>
- Teodorescu, E., Echim, M. M., & Voitu, G. (2021). A perspective on the scaling of magnetosheath turbulence and effects of bow shock properties. *The Astrophysical Journal*, 910(1), 66. <https://doi.org/10.3847/1538-4357/abe12d>
- Torrence, C., & Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, 79(1), 61–78. [https://doi.org/10.1175/1520-0477\(1998\)079<0061:APGTWA>2.0.CO;2](https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2)
- Turicu, D. C., Creț, O., Echim, M., & Munteanu, C. (2022). An FPGA-based solution for computing a local stationarity measure from satellite data. *IEEE Access*, 10, 9668–9676. <https://doi.org/10.1109/ACCESS.2022.3143239>
- Vörös, Z., Baumjohann, W., Nakamura, R., Runov, A., Zhang, T. L., Volwerk, M., et al. (2003). Multi-scale magnetic field intermittency in the plasma sheet. *Annales Geophysicae*, 21(9), 1955–1964. <https://doi.org/10.5194/angeo-21-1955-2003>
- Wawrzaszek, A., & Echim, M. (2021). On the variation of intermittency of fast and slow solar wind with radial Distance, heliospheric Latitude, and Solar Cycle. *Frontiers in Astronomy and Space Sciences*, 7, 617113. <https://doi.org/10.3389/fspas.2020.617113>
- Wawrzaszek, A., Echim, M., Macek, W. M., & Bruno, R. (2015). Evolution of intermittency in the slow and fast solar wind beyond the ecliptic plane. *The Astrophysical Journal Letters*, 814(2), L19. <https://doi.org/10.1088/2041-8205/814/2/L19>
- Welch, P. D. (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, 15(2), 70–73. <https://doi.org/10.1109/tau.1967.1161901>
- Yordanova, E., Bergman, J., Consolini, G., Kretzschmar, M., Materassi, M., Popielawska, B., et al. (2005). Anisotropic scaling features and complexity in magnetospheric-cusp: A case study. *Nonlinear Processes in Geophysics*, 12(6), 817–825. <https://doi.org/10.5194/npg-12-817-2005>