

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

Costel Munteanu<sup>1</sup>, Péter Kovács<sup>2</sup>, and Marius Echim<sup>3</sup>

<sup>1</sup>Institute of Space Science

<sup>2</sup>Wigner Research Centre for Physics

<sup>3</sup>Belgian Institute for Space Aeronomy

January 20, 2023

## 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 from the Cluster 3 spacecraft during an inbound pass through the Earth's magnetosheath region. We show how specific features of the power spectral density (PSD) can be mapped to localised time-frequency regions in the spectrogram representation, and identify multiple intermittent events using the wavelet-based local intermittency measure (LIM). Multiscale probability density functions (PDFs) showed clear departures from Gaussianity, signifying the presence of intermittency. Structure functions (SFs) and rank-ordered multifractal analysis (ROMA) revealed the multifractal nature of the analysed 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 dataset, 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 October 21, 2022

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

4 **Costel Munteanu<sup>1</sup>, Petér Kovács<sup>2</sup>, and Marius Echim<sup>1,3</sup>**

5 <sup>1</sup>Institute of Space Science (ISS), Măgurele, Romania

6 <sup>2</sup>Wigner Research Centre for Physics, Budapest, Hungary

7 <sup>3</sup>The Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium

8 **Key Points:**

- 9 • We built an interactive software analysis tool designed to study space plasma tur-  
10 bulence and intermittency
- 11 • The tool can ingest data from a large variety of sources and includes a varied port-  
12 folio of analysis methods
- 13 • The performance of our analysis tool is demonstrated using in-situ magnetic field  
14 measurements from the Cluster 3 spacecraft

---

Corresponding author: Costel Munteanu, [costelm@spacescience.ro](mailto:costelm@spacescience.ro)

**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 from the Cluster 3 spacecraft during an inbound pass through the Earth's magnetosheath region. We show how specific features of the power spectral density (PSD) can be mapped to localised time-frequency regions in the spectrogram representation, and identify multiple intermittent events using the wavelet-based local intermittency measure (LIM). Multiscale probability density functions (PDFs) showed clear departures from Gaussianity, signifying the presence of intermittency. Structure functions (SFs) and rank-ordered multifractal analysis (ROMA) revealed the multifractal nature of the analysed 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 dataset, 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, e.g., the solar wind, are frequently in a turbulent state characterized by multiscale 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 (Matthaeus et al., 2015; Greco et al., 2018; Bruno, 2019). Significant efforts are devoted to understand turbulence and intermittency in space plasmas (Burlaga, 1991; Marsch & Tu, 1994; Bruno et al., 2001; Sorriso-Valvo et al., 2001; Hnat et al., 2002; Vörös et al., 2003; Yordanova et al., 2005; Echim et al., 2007; Chang et al., 2010; Wawrzaszek et al., 2015; Echim et al., 2021; Teodorescu et al., 2021; Wawrzaszek & Echim, 2021).

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 distribution 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 the Flatness (the normalized fourth-order moment of fluctuations) towards smaller scales, are generally considered signatures of intermittency (Marsch & Tu, 1994; Bruno et al., 2001). The anomalous scaling of structure functions (SFs) is the starting point for more elaborate models of turbulence based on fractals and multifractals (e.g. Sreenivasan, 1991; Benzi et al., 1993).

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: (1) weak plasma turbulence, which considers at the center of the turbulence paradigm different types of waves (e.g. Biskamp, 1993), and (2) 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

66 is a task that requires the simultaneous use of both spectral and statistical approaches  
67 (Chang, 2015).

68 During our studies, we often encountered a gap between software analysis tools and  
69 turbulence research. We found there is no easy, straightforward, and user-friendly way  
70 of performing at once most of the analyses mentioned above. Also, even when the soft-  
71 ware solutions are relatively easy-to-use for non-experienced programmers, they are of-  
72 ten limited in spread and scope of the methods. The Queen Mary Science Analysis Sys-  
73 tem (QSAS; <http://www.sp.ph.ic.ac.uk/csc-web/QSAS>), a software package provid-  
74 ing an environment for the selection, manipulation, and display of space physics data,  
75 is one such solution. Another example is the Space Physics Environment Data Analy-  
76 sis Software (SPEDAS; <http://spedas.org/wiki>), a generalized software development  
77 platform supporting multi-mission data ingestion, analysis and visualization. These tools  
78 function mostly as mission/instrument dedicated plug-ins. QSAS, for example, has a Clus-  
79 ter plug-in for computing the curl, gradient, divergence and barycentric average of a vec-  
80 tor quantity ([http://www.ninepeux.net/haaland/QSAS\\_plugins.html](http://www.ninepeux.net/haaland/QSAS_plugins.html)). SPEDAS in-  
81 cludes multiple plugins for loading and analyzing Magnetospheric Multiscale (MMS) mis-  
82 sion data ([http://spedas.org/wiki/index.php?title=Magnetospheric\\_Multiscale](http://spedas.org/wiki/index.php?title=Magnetospheric_Multiscale)).  
83 These plugins are usually created and supported by the respective mission/instrument  
84 teams, and are useful to analyze data from that specific mission/instrument. IRFU-Matlab  
85 is another useful collection of tools for analysing space physics data (<https://sites.google.com/site/irfumatlab>), strictly linked to the MATLAB environment (<https://www.mathworks.com/products/matlab.html>).  
87

88 General purpose commercial software, like IDL (<https://www.l3harrisgeospatial.com/Software-Technology/IDL>) or MATLAB, are broader in scope and are rather ex-  
89 pensive. These general purpose software do not offer straightforward ways of perform-  
90 ing all the necessary steps of a space data analysis process: from importing, cleaning and  
91 preprocessing a dataset, performing the analyses of interest, visualizing the results, and  
92 finally exporting high-quality figures. Thus, most researchers devote considerable time  
93 and effort to develop their own advanced algorithms and techniques necessary to study  
94 turbulence and intermittency. The problem is further complicated by the ever increas-  
95 ing amount of publicly available space physics data, which is provided in many differ-  
96 ent formats and file types. Thus, in addition to implementing their own analysis routines,  
97 researchers are also required to do their own data importing and preprocessing routines,  
98 and this is also a tedious and difficult task.  
99

100 A recent noteworthy contribution is the Open-source software analysis tool to in-  
101 vestigate space plasma turbulence and nonlinear DYNamics (ODYN; Teodorescu & Echim,  
102 2020). ODYN, a Python-based tool, includes a rich collection of analysis methods ded-  
103 icated to the investigation of space plasma turbulence and intermittency: PSD, PDFs  
104 and their moments, multifractals and discriminating statistics. Data analysis with ODYN  
105 can be performed either on selected events or iteratively (automatic) on larger sets of  
106 measurements through a configurable package of algorithms, which distinguishes it from  
107 the other tools mentioned above.

108 We designed and built an integrated software analysis tool -INA- dedicated specif-  
109 ically to investigate turbulence and intermittency in space plasmas. INA integrates a com-  
110 prehensive collection of analysis methods, from lower order spectral analysis methods  
111 to highly complex multiscale PDFs and multifractal analysis methods. One of the main  
112 distinctive feature of INA is the interactive and user-friendly environment in which one  
113 can easily import a dataset, customise key analysis parameters, apply multiple analy-  
114 sis methods, cross-validate the results, and export high-quality figures. These features  
115 make INA unique compared to similar software products dedicated to turbulence anal-  
116 ysis, like, for example, the ODYN tool.

117 INA was developed in the framework of the European Community’s Seventh Frame-  
118 work Programme project STORM (Solar system plasma Turbulence: Observations, in-  
119 teRmittency and Multifractals), and is publicly available from the project website at [https://](https://www.storm-fp7.eu)  
120 [www.storm-fp7.eu](https://www.storm-fp7.eu). The program was carefully tested and validated during STORM.  
121 The tool is primarily designed to ingest data from space missions like Cluster, Advanced  
122 Composition Explorer (ACE), Venus Express and Ulysses, but can also read data from  
123 other sources. INA can be used to perform a comprehensive analysis of time series, in-  
124 cluding methods for: descriptive statistics, spectral analysis, multiscale analysis of fluc-  
125 tuations, wavelet analysis and structure functions. INA also includes the Rank Ordered  
126 Multifractal Analysis (ROMA), an advanced multifractal analysis method, and this makes  
127 it one of only few publicly available software implementing this method. In fact, to our  
128 knowledge, INA and ODYN are the only freely available tools which include the ROMA  
129 analysis method.

130 INA was built using MATLAB, a proprietary programming and numerical com-  
131 puting platform used by engineers and scientists. In addition to classical script-based  
132 programming, MATLAB can also be used to create standalone applications with cus-  
133 tom, user-friendly graphical user interfaces (GUI) with standard components such as but-  
134 tons, check boxes, and drop-down lists. The standalone applications generated using MAT-  
135 LAB Compiler can be shared and used freely by any interested user, even if they don’t  
136 own a MATLAB licence. The only prerequisite needed is the free MATLAB Runtime,  
137 available at <https://www.mathworks.com/products/compiler/matlab-runtime.html>.

138 In a recent series of papers (Deak et al., 2018, 2021; Opincariu et al., 2019; Turicu  
139 et al., 2022; Munteanu et al., 2022), we used the INA software extensively during the de-  
140 sign, implementation and validation phases of hardware, Field-Programmable Gate Ar-  
141 ray (FPGA) implementations of various analysis methods. In Deak et al. (2018), the INA  
142 software implementation of PDFs computation, was the starting point for the FPGA im-  
143 plementation of the same method. Opincariu et al. (2019) added spectral analysis to our  
144 FPGA toolbox, again, starting from ideas and algorithms implemented in INA. Deak et  
145 al. (2021) used the INA mathematical kernel and algorithm for the Flatness parameter  
146 calculations, and generated an FPGA implementation of this method. In Turicu et al.  
147 (2022), a local stationarity measure is implemented in FPGA, again, based on an INA  
148 software implementation of the same measure. An FPGA implementation of a magnetic  
149 field directional discontinuity detector is presented in Munteanu et al. (2022). Same as  
150 before, the mathematical kernel of the discontinuity detector was first developed and val-  
151 idated using INA, and then ported onto FPGA.

152 The paper is organized as follows. Section 2 gives a brief mathematical description  
153 the analysis methods currently implemented in INA. In Section 3 we provide a detailed  
154 program overview. Section 4 presents an illustrative example: a comprehensive study  
155 of magnetic field magnitude measured by the Cluster 3 spacecraft during a magnetosheath  
156 inbound pass on 09 February 2007, employing the full set of INA analysis methods. Sec-  
157 tion 5 summarizes the paper.

## 158 2 Description of the Analysis Methods Implemented in INA

159 INA includes a comprehensive collection of analysis methods dedicated to study  
160 turbulence and intermittency: starting from descriptive statistics (histograms); advanc-  
161 ing to spectral analysis (both Fourier-based and wavelet-based methods); then moving  
162 on to advanced methods to study multiscale fluctuations (probability distribution func-  
163 tions and structure functions); and, finally, the rank-ordered multifractal analysis (ROMA),  
164 a recently developed method to study multifractals.

165

## 2.1 Spectral and Wavelet Analyses

166

167

168

169

The periodogram is the fastest way to estimating the power spectral density (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):

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

170

171

here  $i$  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  $S(f)$  vs.  $f$ .

172

173

174

175

176

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 cleaner PSD estimate.

177

178

179

180

Fourier analysis can also be 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 Welch and spectrogram methods are included in the INA Spectral Analysis module.

181

182

183

184

185

186

187

188

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 localising short-time, high-frequency events, compared to the Fourier-based spectrograms.

189

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

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

190

191

192

193

194

where  $a$  is the scale,  $b$  is the translation parameter,  $\psi$  is the wavelet (mother) function and  $W(a, b)$  are the wavelet coefficients. A scalogram, the wavelet equivalent of the 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.  $a \approx 1/f$ ).

195

196

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

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

197

198

199

200

201

202

where  $\langle |W(a, b)|^2 \rangle_b$  denotes the average over all (squared) wavelet coefficients corresponding to scale  $a$ . LIM quantifies the deviations between the local and the average power of the signal  $x(t)$  at each scale  $a$ . In other words, since  $|W(a, b)|^2$  is equivalent to the Fourier spectrum, a value of  $LIM = 1$  means that the signal  $x(t)$  does not show intermittency, since each portion of it has the same power as the average spectrum. Conversely, values of  $LIM > 1$  identify those portions of the the signal  $x(t)$  that have more

203 power than the average. Consequently, regions with  $LIM > 1$  can be used to locate  
 204 intermittent events in time and scale. A dedicated Wavelet Analysis module is included  
 205 in INA.

## 206 **2.2 Multiscale Study of Fluctuations**

207 An incremental measure for the variability of a physical variable  $x$  is constructed  
 208 from time differences computed at different scales  $\tau$ :

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

209 The probability density function (PDF) at each scale  $\tau$  is estimated from the nor-  
 210 malized histograms of  $\delta x(t, \tau)$ . Multiscale PDFs can be computed from the Probabil-  
 211 ity Density module of INA, which includes also the one parameter rescaling (OPR; Hnat  
 212 et al., 2002) method, which aims to find a single parameter able to rescale/collapse the  
 213 PDFs on a single (master) curve. The assumption behind OPR approach is that the mul-  
 214 tiscale probability densities can be described by a stable, symmetric universal shape. To  
 215 verify this assumption, one generates a log-log plot of the peak values of the PDFs ver-  
 216 sus  $\tau$ . The OPR scaling exponent is then determined from the slope of this plot.

217 Multiscale PDFs provide valuable data on the intermittency of a fluctuating phe-  
 218 nomenon. The conventional method to evaluate intermittency is to study the scaling be-  
 219 havior of the moments of order  $q$  of the time increments, known as the structure func-  
 220 tions (SFs; Frisch, 1995):

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

221 where  $P(\delta x, \tau)$  is the probability density function at scale  $\tau$ . A set of scaling exponents  
 222  $\zeta_q$  are then estimated from the (log-log) slopes of  $SF_q(\tau)$  vs.  $\tau$ . When  $\zeta_q$  varies linearly  
 223 with the moment order  $q$ , the fluctuating variable is considered monofractal or self-similar.  
 224 When the linear relation between  $\zeta_q$  and  $q$  is not satisfied, the fluctuating phenomenon  
 225 is considered multifractal. Structure function analysis has a dedicated module in INA.

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

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

228 A given time series is said to be intermittent if  $F(\tau)$  continually increases towards smaller  
 229 scales. Another time series can be said to more intermittent if  $F(\tau)$  grows faster. Gaus-  
 230 sian fluctuations constitute a special case of non-intermittent signals, for which  $F(\tau)$  is  
 231 equal to 3. Computations of the Flatness parameter can be performed from the Prob-  
 232 ability Density module.

## 233 **2.3 Rank-Ordered Multifractal Analysis (ROMA)**

234 Because the conventional structure function analysis reveals the statistics of the  
 235 full set of fluctuations (which is dominated by that of the small amplitudes), see Equa-  
 236 tion (5), the physical interpretation of the multifractal nature cannot be easily under-  
 237 stood by simply examining the curvatures of the deviations from the linearity between  
 238  $\zeta_q$  and  $q$ . This means that we have to search for a way of isolating out the subdominant  
 239 (large-amplitude) fluctuations and then perform the statistical analyses for each of the

240 isolated populations. Such grouping of fluctuations must depend on their amplitude, but  
 241 cannot depend merely on the raw values of the fluctuations because the amplitude ranges  
 242 will be different for different scales.

243 These ideas led Chang and Wu (2008) to propose a new multifractal analysis tech-  
 244 nique, called Rank Ordered Multifractal Analysis (ROMA), which is able to isolate the  
 245 subdominant fluctuations and investigate their multifractal nature (see also, Chang et  
 246 al., 2010; Chang, 2015; Echim et al., 2021).

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

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

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

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

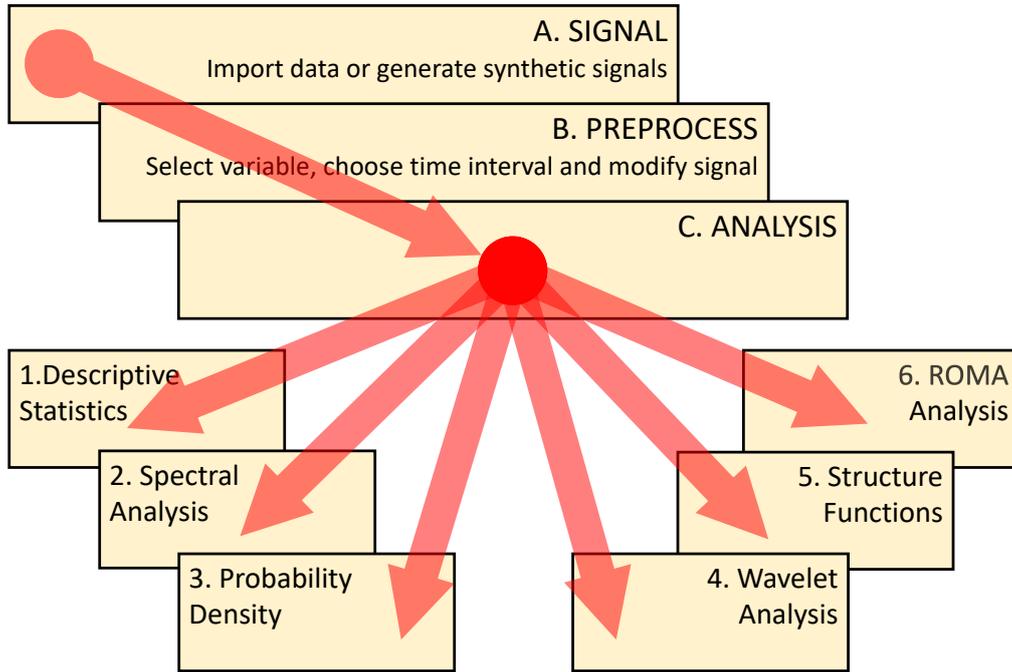
258 where  $a_1 = Y_1 \cdot (\tau/\tau_0)^s$  and  $a_2 = Y_2 \cdot (\tau/\tau_0)^s$ . As in the conventional structure func-  
 259 tion analysis, we may now search for the value of  $s$  such that the scaling property of the  
 260 range-limited structure function is:  $SF_q^{(\tau l)}(\delta x, \tau)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

264 The physical interpretation of the ROMA spectrum is based on the fact that  $s$  is  
 265 equivalent to the local Hurst exponent for the subset of fluctuations in the given range  
 266  $\Delta Y$ . The Hurst exponent reveals the degree of persistency of a signal (e.g. Section 3.10  
 267 in Hergarten, 2002). Persistency means that a time series has a long-term tendency for  
 268 positive variations (with respect to the mean value) to be followed by other positive vari-  
 269 ations, and vice-versa. At the other end, anti-persistency means that a time series has  
 270 a long-term tendency for positive variations to be followed by negative variations, and  
 271 so on. A signal is said to be persistent if it has a Hurst exponent between 0.5 and 1, and  
 272 anti-persistent if its Hurst exponent is between 0 and 0.5. A Hurst exponent equal to  
 273 0.5 means that the signal fluctuations are completely uncorrelated, showing no long-term  
 274 tendency.

### 275 **3 Program Overview**

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

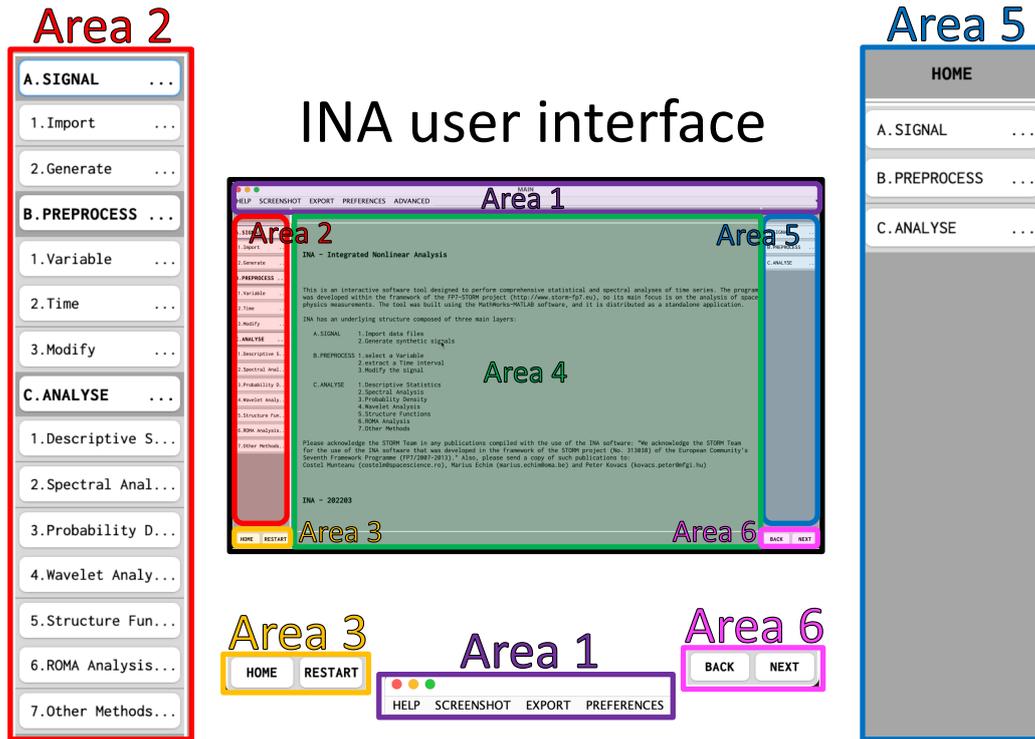
## Integrated Nonlinear Analysis - INA



**Figure 1.** Diagram representation of the Integrated Nonlinear Analysis toolbox (INA). The program is structured into three layers (A, B, and C), and six modules. The user is guided to follow a linear progression from data import to the analysis layer, which acts as a "hub" connecting all of the analysis modules.

283 Figure 2 shows a screenshot of the INA HOME screen highlighting six areas, each  
 284 providing a user interface for a specific scope. Area 1 allows access to HELP and PREF-  
 285 FERENCES. The two export options (SCREENSHOT and EXPORT) are also accessi-  
 286 ble from this panel. A set of buttons corresponding to the layers and modules mentioned  
 287 in the first paragraph, is always displayed on the left part of the main user interface (Area  
 288 2). Using these buttons one can freely navigate between layers/modules and apply vari-  
 289 ous analysis methods. The HOME and RESTART buttons are always displayed on the  
 290 lower left corner of the main interface (Area 3), and offer two options to open the HOME  
 291 screen: (a) without erasing any previous settings (imported signal, time selection, pre-  
 292 processing option, etc.), and (b) by erasing all settings and RESTART the program. Two  
 293 additional navigation buttons are always displayed on the lower right corner of the main  
 294 interface (Area 6), labeled NEXT and BACK, and allow a user to advance (or go back)  
 295 to the next (previous) layer/module, relative to the default linear progression from layer  
 296 (A) to layer (C) and from analysis module (1) to module (6).

297 All the results generated by INA are displayed in Area 4. Note that in Fig. 2, Area  
 298 4 shows the HOME screen. Most of the user interactions with INA take place in Area  
 299 5. This panel displays various user interface controls (buttons, editable text, dropdown  
 300 menus, etc.), depending on the specific layer/module selected by the user, as described  
 301 in the following sections.

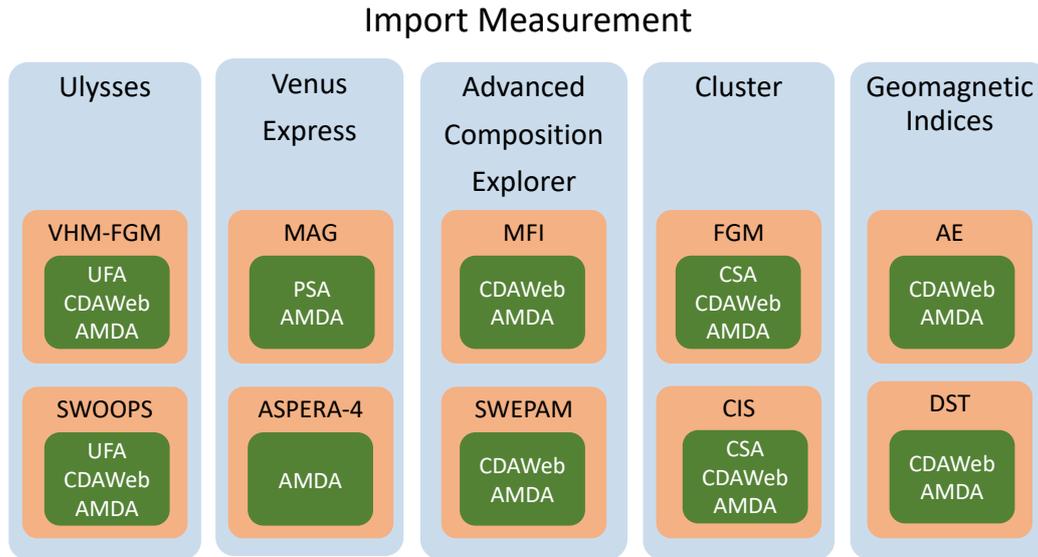


**Figure 2.** (central panel) Screenshot of the INA HOME screen highlighting six areas of the graphical user interface. Area 1 contains links to HELP, SCREENSHOT, EXPORT and PREFERENCES. Area 2 includes a static set of buttons accessible at any moment. Area 3 includes the HOME and RESTART buttons. Area 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)

302 **3.1 Importing Data**

303 The INA layer titled A. SIGNAL offers various options to import/ingest data for  
 304 further analysis. This layer is organized in two sections: Import and Generate. The Im-  
 305 port section includes several possible data sources: Measurement, File and Other. Import-  
 306 Measurement offers options to ingest in-situ satellite observations and ground-based ge-  
 307 omagnetic indices. Import-File allows one to import various types of data formats like  
 308 CDF (Common Data Format; <https://cdf.gsfc.nasa.gov>), TXT or MAT (binary MAT-  
 309 LAB files; <https://www.mathworks.com/help/matlab/workspace.html>). Import-Other  
 310 includes links to custom data formats requested by users. Note that the program is still  
 311 in development, and user requests for custom file formats can be added here.

312 Figure 3 depicts a diagram representation of the Import-Measurement section, which  
 313 includes five branches: Ulysses, Venus Express (VEX), Advanced Composition Explorer  
 314 (ACE), Cluster and Geomagnetic Indices; for each branch are defined three levels of con-  
 315 tent: satellite name, instrument name (acronym), and the source where the respective data  
 316 can be retrieved/downloaded. The targeted spacecraft are: ACE (<https://www.nasa.gov/ace>), Ulysses, VEX and Cluster (<https://sci.esa.int/web/home/-/51459-missions>);  
 317 the targeted instruments are VHM-FGM and SWOOPS, the magnetic field and plasma  
 318 instruments onboard Ulysses; MAG and ASPERA-4, the magnetic field and plasma in-  
 319 struments onboard VEX; MFI and SWEPAM, the magnetic field and plasma instruments  
 320



**Figure 3.** Diagram depicting the structure of the Import-Measurement section of INA. 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).

321 onboard ACE; FGM and CIS, the magnetic field and plasma instruments onboard Cluster.  
322 INA was optimized and extensively tested to ingest data from these spacecraft.

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

328 In Section 4 we present a case study to demonstrate the functionalities of INA. The  
329 case study uses a data file containing magnetic field measurements from the fluxgate mag-  
330 netometer (FGM) onboard the Cluster 3 spacecraft, downloaded from the Cluster Sci-  
331 ence Archive (CSA, <https://www.cosmos.esa.int/web/csa>), the official data repos-  
332 itory for the Cluster mission. In Fig. 3 the acronyms denote the following data sources:  
333 UFA stands for Ulysses Final Archive (<http://ufa.esac.esa.int/ufa>), and PSA stands  
334 for Planetary Science Archive (<https://archives.esac.esa.int/psa>). We also include  
335 two alternative data sources: CDAWeb (<https://cdaweb.gsfc.nasa.gov>) and  
336 AMDA (<http://amda.irap.omp.eu>).

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

345

### 3.2 Preliminary Preprocessing

346

347

348

349

350

351

352

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 data file imported. In order to pursue the analysis, the user has to select a variable and then define the time interval. The time selection section offers two options to select time: (a) interactively using the cursor, and (b) manually by setting the time limits.

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

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 within 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  $\sigma_y$  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 Fig. 3, the data gaps are flagged by the data providers, and linear interpolation across data gaps is used 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, e.g. 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.

369

### 3.3 Data Analysis Modules

370

371

372

373

374

375

376

377

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.

378

379

380

381

382

383

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 power spectral density (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.

384

385

386

387

388

389

390

Spectral Analysis (module 2) contains two methods titled PSD-Welch and Spectrogram. PSD-Welch estimates the power spectral density using the Welch method (see Section 2.1). Here, one can customize various analysis parameters like: the window type, the segment length and the overlap between adjacent segments. Spectrogram method estimates the PSD for a series of time windows, 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.

391

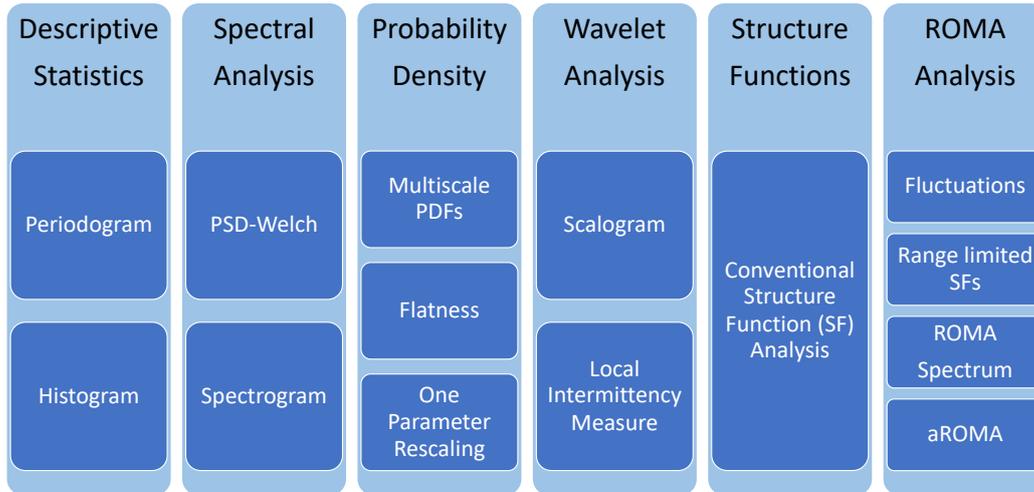
392

393

394

The PSD-Welch section offers an option called "Slope analysis". This functionality offers an interactive computation of PSD slopes from the log-log representation of PSD vs. 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),

### Analysis Modules



**Figure 4.** Diagram depicting the data analysis modules currently implemented in INA. From left to right, the INA analysis modules are: (1) Descriptive Statistics, (2) Spectral Analysis, (3) Probability Density, (4) Wavelet Analysis, (5) Structure Functions and (6) ROMA Analysis.

395 (b) *int*, where the user can choose to simultaneously fit one, two, or three different fre-  
 396 quency ranges of the same PSD and (c) *met*, where the user can choose how the frequency  
 397 ranges are selected: using either the mouse or by typing explicit limits for the frequency  
 398 interval. For reference, the PSD-Welch results depicted in Fig. 5, were obtained using  
 399 linear fits (*fit*="linfit") over two frequency ranges (*int*="two").

400 Analysis module 3, titled Probability Density, includes three statistical approaches:  
 401 PDF, Flatness, and OPR. The PDF method computes the multiscale probability den-  
 402 sity functions, Flatness is a quantitative measure of intermittency, and OPR is the one  
 403 parameter rescaling method (see Section 2.2). The time differences  $\delta x(t, \tau)$  computed  
 404 as described in Equation 4, require the choosing of a range of time scales  $\tau$ . The small-  
 405 est possible time scale is given by the time resolution of the signal, while the largest scale  
 406 is limited by its length. The range of values for  $\tau$  can be customized by the user. Esti-  
 407 mating the probability density functions involves the computation of normalized histograms  
 408 of  $\delta x(t, \tau)$ . The number of bins required by the histogram method can be adjusted by  
 409 the user (using fewer bins can lead to an incorrect estimation of the shape of the PDFs,  
 410 while using a very large number of bins increases the variability of the results).

411 Wavelet Analysis (module 4) can be used to compute scalograms and the associ-  
 412 ated Local Intermittency Measures (LIM). The scalogram is a 3D representation of the  
 413 logarithm of the (squared) wavelet coefficients, and LIM can be regarded as a normal-  
 414 ized scalogram (see more in Section 2.1). Wavelet analysis involves the choosing of a wavelet  
 415 mother function. For example, Torrence and Compo (1998) argue that the wavelet func-  
 416 tion should reflect the type of features present in a time series: for time series with sharp  
 417 jumps, one should choose a function such as the Haar, while for smoothly varying time  
 418 series one should choose a smoother wavelet function. In INA, the user can select from  
 419 various wavelet functions: *db* are the Daubechies' extremal phase wavelets, with *db1* be-  
 420 ing the Haar wavelet; *sym* are the Symlets family of wavelets and *coif* stands for Coiflets  
 421 ([https://www.mathworks.com/help/wavelet/gs/introduction-to-the-wavelet-families](https://www.mathworks.com/help/wavelet/gs/introduction-to-the-wavelet-families.html)  
 422 [.html](https://www.mathworks.com/help/wavelet/gs/introduction-to-the-wavelet-families.html)).

Structure Functions (analysis module 5) can be used to perform a conventional structure function analysis of the signal. The user can customize the range of time scales involved in structure functions calculation, as described in Equation 5. The moment order  $q$  is another customizable parameter. In practice, structure functions 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 2013 (see also, Teodorescu et al., 2021) recommends the use of  $q_{max} = \log(N) - 1$ , where  $N$  is the number of samples in the dataset.

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, i.e. 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.

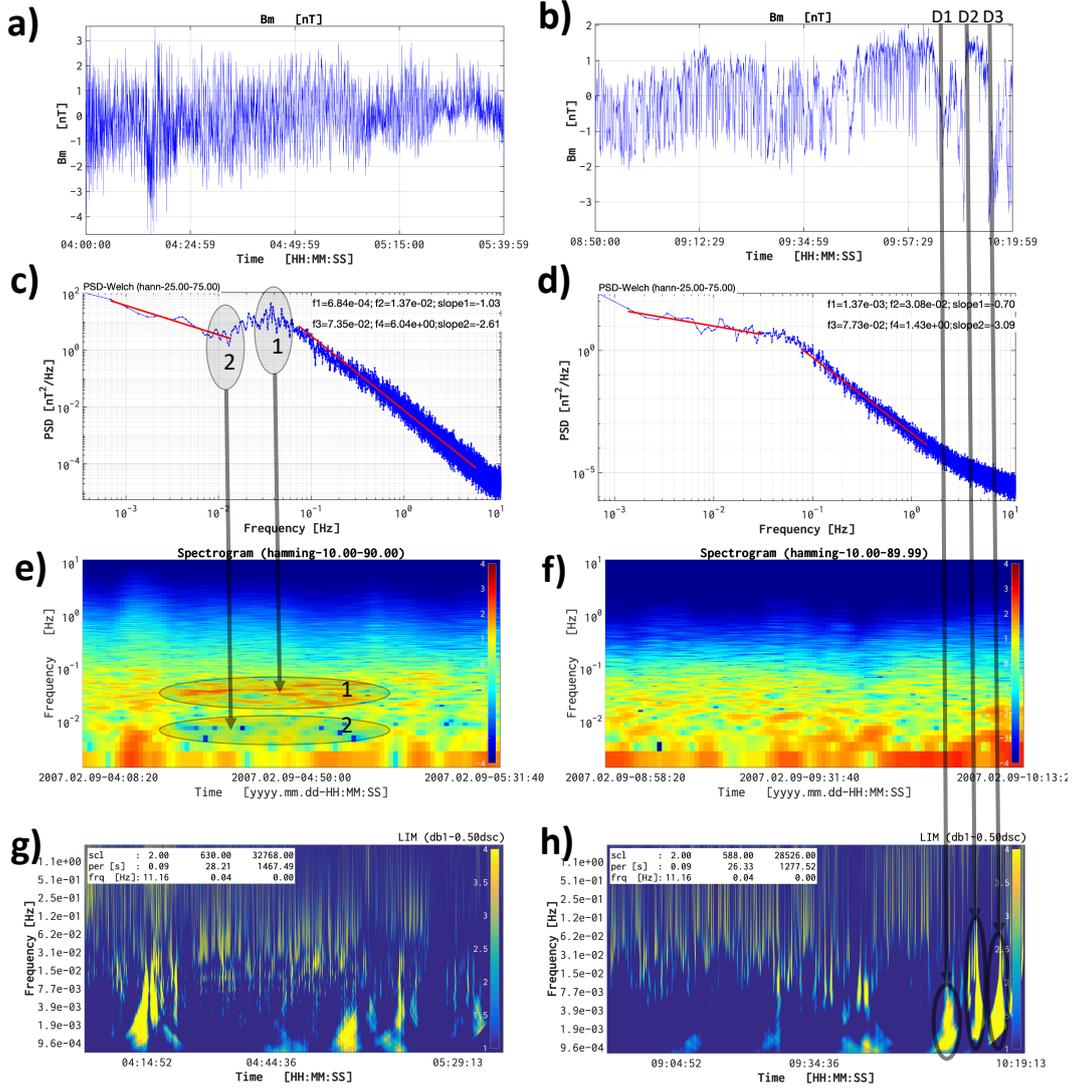
### 3.4 Exporting Results from INA

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 panel 1 (see Fig. 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 screenshot functionality of the Mac OS operating system.

The second, and recommended option to generate publication-quality figures is the EXPORT functionality, also accessible from area 1 illustrated in Fig. 2. Note that each individual plot in Figures 5, 6, 7 and 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 datasets to be exported. The content of this window depends on the layer/module from which it was called from.

## 4 An Illustrative Example

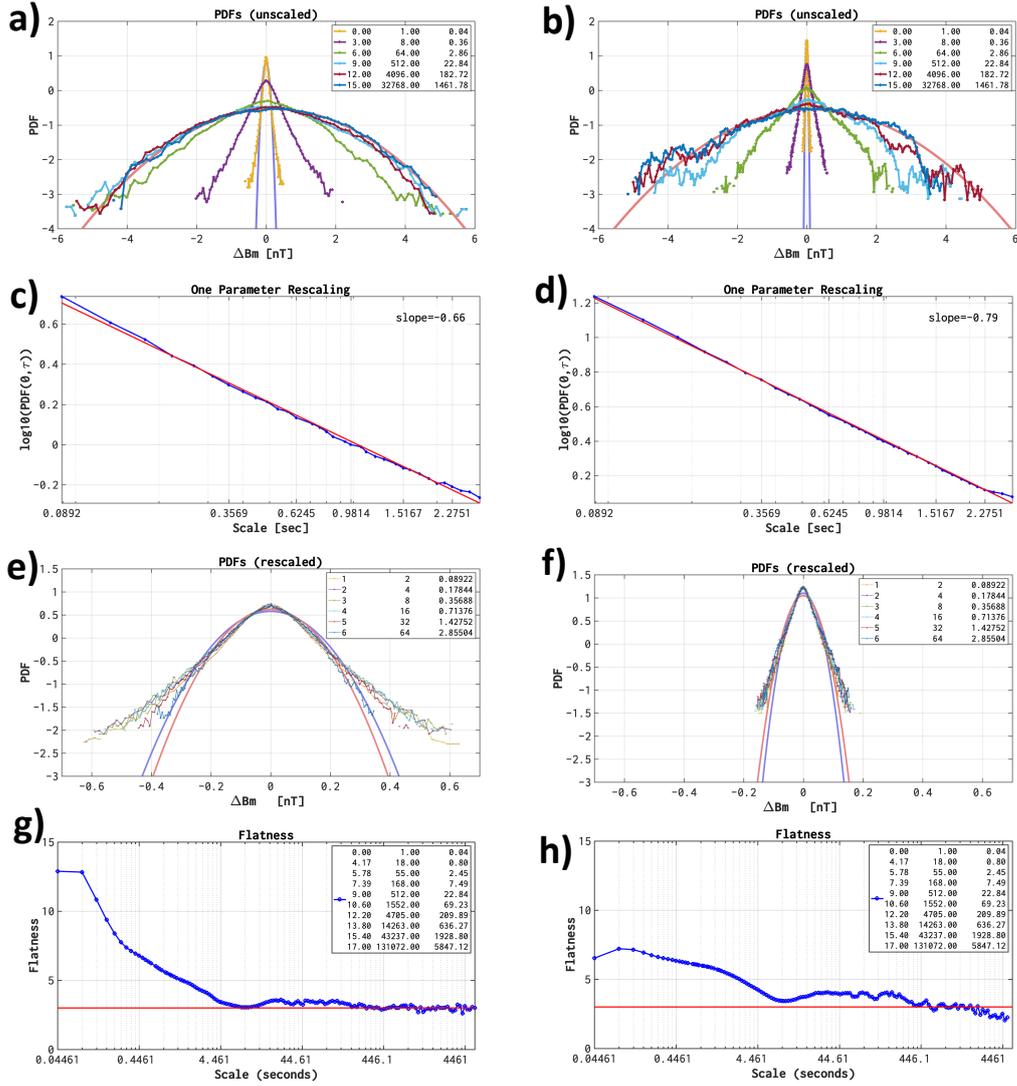
In this section we provide an example which demonstrates most of INA capabilities. We analyze the magnetic field intensity measured by Cluster 3 spacecraft during a magnetosheath inbound pass on 9 February 2007. We use high (22 Hz) time resolution magnetic field data from the fluxgate magnetometer (FGM; 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, 6, 7 and 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 structure function 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.



**Figure 5.** Spectral analysis using INA of the magnetic field intensity measured by Cluster 3 in the magnetosheath on 9 February 2007. Column 1: a) time series of measurements for the bowshock interval, c) PSD-Welch, e) Fourier Spectrogram, and g) wavelet LIM. Column 2 show the corresponding results for the magnetopause interval.

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

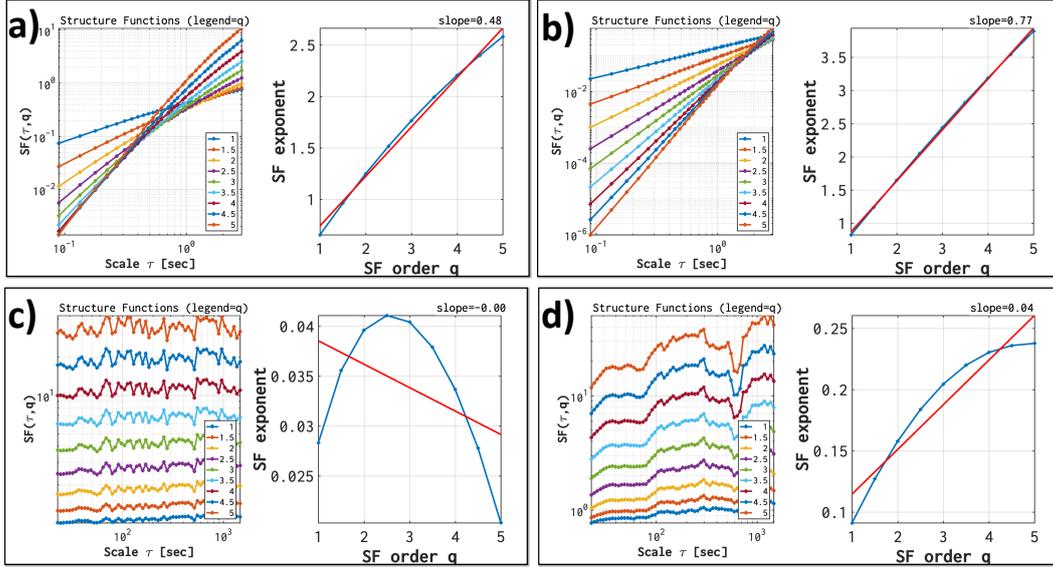
475 The results of the PSD-Welch method are depicted in Figures 5c and 5d, and in-  
 476 dicate the presence of two spectral regions characterized by different power laws, for both  
 477 intervals. At low frequencies, the near-bowshock spectrum is characterized by a spec-  
 478 tral index close to  $f^{-1}$ ; a slightly shallower ( $f^{-0.70}$ ) spectrum is found for the near-magnetopause  
 479 interval. At higher frequencies, the spectrum scales approximately as  $f^{-8/3}$  near the bow-  
 480 shock; a slightly steeper scaling ( $f^{-3.09}$ ) is observed at the magnetopause. Similar find-  
 481 ings are reported by Huang et al. (2017) (see also Teodorescu & Echim, 2020). Damp-



**Figure 6.** Multiscale statistical analysis using INA. Column 1 shows: a) unscaled PDFs, c) one parameter rescaling, e) Rescaled PDFs and f) Flatness, for the bowshock interval. Column 2 shows the same results, but for the magnetopause interval. The legends show the scales in powers of two (first column), number of points (second column) and seconds (third column).

482 ing of kinetic Alfvén waves through Landau resonances is invoked as a possible cause of  
 483 the steepening of the high-frequency spectrum (Howes et al., 2008).

484 Figures 5e and 5f depict the Fourier spectrograms for the two intervals. These re-  
 485 sults suggest that both signals are nonstationary, i.e. their spectra changes in time. Non-  
 486 stationarity should cast doubt, or even invalidate the results obtained using the PSD-  
 487 Welch analysis. Two features dominate the PSD of the bowshock magnetic field depicted  
 488 in Fig. 6c: (a) a frequency band of increased power between  $10^{-2}$  and  $10^{-1}$  Hz; and (b)  
 489 a frequency band of decreased power around  $10^{-2}$  Hz. The Fourier spectrogram depicted  
 490 in Fig. 6e clearly shows that these two features are in fact localized in time. The Spec-  
 491 trogram shown in Figure 5 suggests that the time-frequency bands highlighted in panel



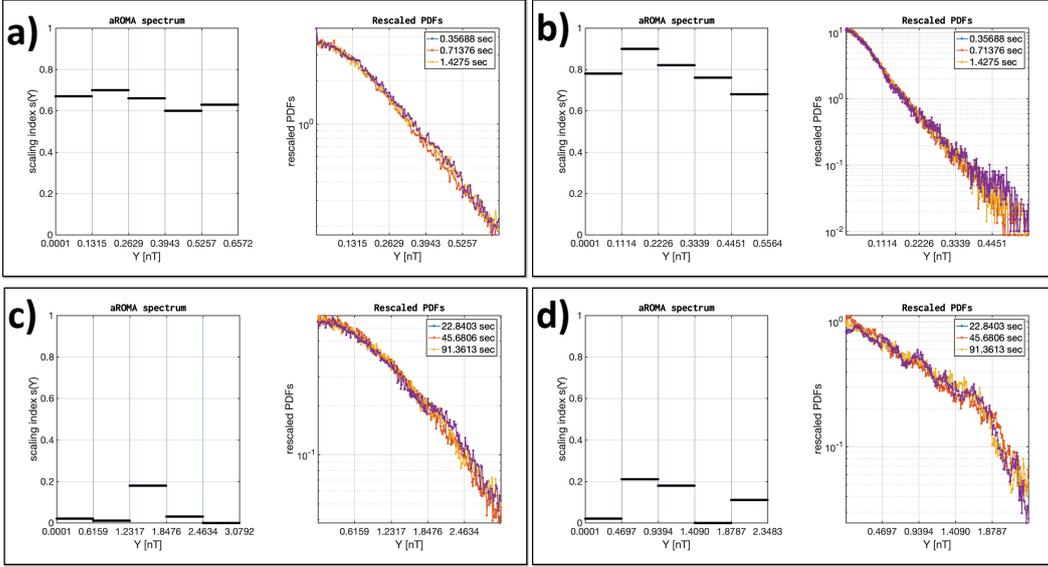
**Figure 7.** Structure Function analysis using INA. Column 1 shows: a) SFs for small scales (time range 1: from 0.08 to 2.86 s) and c) for large scale (time range 2: from 22.84 to 1928.80 s) for the bowshock interval. Column 2, same results, but for the magnetopause interval.

492 e are responsible for the corresponding features observed in the PSD presented in panel  
 493 c.

494 A wavelet analysis was also applied on these two intervals. Figures 5g and 5h show  
 495 the local intermittency measures (LIM), i.e., normalized wavelet scalograms, for the two  
 496 intervals. The results indicate that both time series contain strong intermittent events,  
 497 identified by localised bright yellow regions. Note that the LIM colour scale varies from  
 498 1 to 4, meaning that the brightest features correspond to regions where the local power  
 499 spectral density is (at least) four times larger than the background power. Multiple in-  
 500 termittent events can be found in both intervals, see, e.g., the three strong discontinu-  
 501 ities observed in interval 2 between 09:57 and 10:19 UT, labeled in Fig. 5b as (D1), (D2)  
 502 and (D3). One notes that these rapid signal variations correspond to three localised bright  
 503 yellow regions highlighted in the LIM representation of Fig. 6h.

504 Figure 6 depicts the results of the statistical analysis of the two intervals, adding  
 505 supplementary information to the results provided by the spectral analysis above. The  
 506 multiscale PDFs (Fig. 6a and 6b) show significant departures from Gaussianity for small  
 507 scales ( $\leq 2.86$  s), suggesting the presence of intermittency. Larger scales ( $\geq 22.84$   
 508 s), on the other hand, closely follow a Gaussian shape. For the smallest scale depicted  
 509 here (0.04 s), the large-amplitude fluctuations appear to be slightly further away from  
 510 a Gaussian shape for interval 1 compared to interval 2; this implies that interval 1 has  
 511 a slightly higher degree of intermittency compared to interval 2.

512 Figure 6 also depicts results from the one parameter rescaling (OPR) technique.  
 513 OPR can be used to estimate the rescaling index, by computing the slope of PDF max-  
 514 ima vs. scale (e.g., Hnat et al., 2002). For interval 1 (Fig. 6c) the scaling index is 0.66;  
 515 and for interval 2 (Fig. 6d), the index is 0.79. The accuracy of the scaling index estimated  
 516 using OPR, is verified by rescaling the PDFs. The results, depicted in Figs. 6e and 6f,  
 517 show that OPR gives better rescaling results for interval 2 compared to interval 1. Note  
 518 that OPR is able to rescale only the small amplitude fluctuations, the large ones rest un-  
 519 scaled suggesting the system is multifractal (Chang & Wu, 2008). According to the in-



**Figure 8.** ROMA analysis using INA. Column 1 depicts the aROMA spectrum for a) small scales and c) large scales, for the bowshock interval. Column 2, same results, but for the magnetopause interval. Each panel also includes an illustration of the rescaled PDFs.

520 interpretation given in Section 2.3, the scaling index is equivalent to the Hurst exponent  
 521 of the small amplitude fluctuations. The OPR results suggest that both intervals are char-  
 522 acterized by persistency, i.e., the Hurst exponent is larger than 0.5, meaning that there  
 523 is a long-term tendency for positive variations (with respect to the mean value) to be  
 524 followed by other positive variations, or vice-versa for negative variations. Also, the in-  
 525 dex is larger for interval 2, meaning that interval 2 has a higher degree of persistency  
 526 compared to interval 1.

527 The results extracted from the multiscale PDF analyses are confirmed by the Flat-  
 528 ness depicted in Figs. 6g and 6h. Flatness increases towards smaller scales, which is con-  
 529 sidered a signature of intermittency (Bruno & Carbone, 2013); also, the Flatness value  
 530 for the smallest scale is larger for interval 1 compared to interval 2, implying that inter-  
 531 val 1 has a higher degree of intermittency compared to interval 2. From the Flatness anal-  
 532 ysis we also observe that the intermittent structures, i.e. Flatness values larger than Gaus-  
 533 sian value of 3, pertain to the same range of scales in both intervals.

534 The results of the structure function (SF) analysis are depicted in Figure 7. These  
 535 results confirm the presence of two scale domains with different scaling properties, simi-  
 536 lar to the results obtained with Flatness. For the small scales region, the SF exponent  
 537 shows a slightly non-linear variation with the moment order  $q$  for interval 1 (Fig. 7a),  
 538 signifying multifractality, and an approximately linear variation for interval 2 (Fig. 7b),  
 539 i.e. quasi-monofractal process. The slope of the SF exponent vs  $q$  is 0.48 for interval 1  
 540 and 0.77 for interval 2. For the large scales region, the slope of the SF exponent vs  $q$  is  
 541 0 for interval 1 (Fig. 7c) and 0.04 for interval 2 (Fig. 7d). These results show a clear anti-  
 542 persistent character of large scale fluctuations. This means that an increase will most  
 543 likely be followed by a decrease or vice-versa, i.e., values will tend to revert to a mean.  
 544 This means that future values have a tendency to return to a long-term mean. A Hurst  
 545 exponent closer to 0, signifies a stronger tendency for the time series to revert to its long-  
 546 term mean value.

547 The multifractality of magnetosheath fluctuations at small scales revealed by the  
 548 structure function analysis is confirmed by an analysis with the ROMA method. Some  
 549 of the results obtained using ROMA are depicted in Figure 8. The scale intervals cho-  
 550 sen for the analysis are the same as those used in Fig. 7. Figs. 8a and 8b depict the small-  
 551 scale ROMA multifractal spectra, that is, the multifractal index  $s$  versus the rescaled vari-  
 552 able  $Y$ , for the two intervals. Figs. 8c and 8d show the corresponding results for the large  
 553 scales. When the scaling indices computed with ROMA are applied to the multiscale PDFs  
 554 (see Equation 7), one obtains the rescaled PDFs depicted at the right-side of each ROMA-  
 555 spectrum. Fig. 8 shows some differences between the bowshock and the magnetopause  
 556 interval. At small scales both fluctuations are persistent ( $s > 0.5$ ). At these scales, ki-  
 557 netic effects are believed to be responsible for the persistency. Close to bowshock, a de-  
 558 crease at large rescaled amplitudes  $Y$  towards a Gaussian value of 0.5 value is observed.  
 559 At the magnetopause, a similar decrease is observed, but the scaling index remains large  
 560 ( $s > 0.65$ ) even for the largest rescaled amplitudes. Also, the overall-higher scaling in-  
 561 dex for magnetopause, correlates very well with the steeper slope of the PSD. At large  
 562 scales the fluctuations are anti-persistent ( $s < 0.5$ ) for both intervals, with the near-  
 563 bowshock fluctuations appearing more monofractal, i.e., showing a constant value for  $s(Y)$ ,  
 564 compared to the near-magnetopause fluctuations.

## 565 5 Summary and Conclusions

566 We built, tested and released a software analysis toolbox - INA- designed to study  
 567 space plasma turbulence and intermittency. The toolbox was developed in MATLAB,  
 568 but it is distributed as a standalone executable file which can be freely shared and used  
 569 by any interested user. The toolbox can ingest various satellite and ground-based mea-  
 570 surements, and includes a comprehensive collection of analysis methods. The software  
 571 is designed to speed up and enhance the scientific output by offering a straightforward  
 572 way from spacecraft data import to exporting analysis results, thus significantly reduc-  
 573 ing the time necessary for data preprocessing or software implementation of analysis meth-  
 574 ods. It can also facilitate the collaboration between scientists with common research in-  
 575 terests and can be a suitable environment for sharing and discussing scientific results.

576 One of the main distinctive features of INA is the integrated, interactive, and user-  
 577 friendly environment in which one can import a dataset, customize key analysis param-  
 578 eters, apply multiple methods on the same signal and then produce high-quality graphi-  
 579 cal results. Other distinctive features include: the versatility to import data from mul-  
 580 tiple sources, the rich collection of analysis methods, and the detailed configuration op-  
 581 tions that allow a fine customisation of analysis parameters and graphical presentation  
 582 of results.

583 Data from specific spacecraft missions such as Venus Express, Ulysses, ACE or Clus-  
 584 ter, can be processed through INA while other types of datasets can be easily added. Con-  
 585 siderable effort has been invested in data preprocessing such that INA can guarantee out-  
 586 put at the highest scientific quality. In this regard, several approaches have been imple-  
 587 mented to manage data gaps. The analysis methods in INA span a wide range, from de-  
 588 scriptive statistics, spectral analysis (both Fourier and wavelet-based methods), statis-  
 589 tical analysis of fluctuations (multiscale PDFs and their moments), to multifractal anal-  
 590 ysis based on the newly developed ROMA method. INA offers an easy access to these  
 591 advanced analysis and visualization methods, and can be used by experienced scientists  
 592 and also as a training tool for students getting accustomed to both methodology and tech-  
 593 nical aspects of a comprehensive data analysis approach for the study of turbulence and  
 594 intermittency in space plasmas.

595 An illustrative example is provided, demonstrating the performance of INA when  
 596 applied to a specific dataset. We analysed magnetic field intensity measured by the Clus-  
 597 ter 3 spacecraft during an inbound pass through the Earth's magnetosheath on 9 Febru-

ary 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 localised time-frequency regions in the spectrogram representation, demonstrating that the interval is nonstationary. For the near-magnetopause interval, a series of intermittent events, i.e. 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 fluctuations were shown to be Gaussian. Structure Function 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 persistency 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, e.g. via Fourier spectrograms and wavelet scalograms. Multiscale PDFs can be used to estimate the scaling exponent, but so do Structure Function 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 and interactive 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.

## Acknowledgments

This work was supported by the European Community's Seventh Framework Programme (FP7) under grant agreement 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 and Belgian Solar Terrestrial Center of Excellence (STCE). The work of C. M. was supported by ESA PRODEX MISSION, and National projects PN-III-P1-1.1-TE-2019-1288 and PN-III-P1-1.1-TE-2021-0102.

## Availability Statement

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>. The INA software is publicly available for download from the STORM-FP7 project website: <http://www.storm-fp7.eu/index.php/data-analysis-tools>.

## References

- Alexandrova, O., Chen, C., Sorriso-Valvo, L., Horbury, T., & Bale, S. (2013). Solar wind turbulence and the role of ion instabilities. *Space Sci. Rev.*
- Balogh, A., Carr, C. M., Acuña, M. H., Dunlop, M. W., Beek, T. J., Brown, P., ... Schwingenschuh, K. (2001). The cluster magnetic field investigation: overview of in-flight performance and initial results. *Annales Geophysicae*, *19*, 1207-1217. doi: 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. *Phys. Rev. E*, *48*, R29-R32. doi: 10.1103/PhysRevE.48.R29
- Biskamp, D. (1993). *Nonlinear magnetohydrodynamics*. Cambridge University Press. Retrieved from <https://doi.org/10.1017/CB09780511599965>
- Bloomfield, P. (2000). *Fourier analysis of time series: An introduction*. Wiley Series in Probability and Statistics. (ISBN: 978-0-471-88948-9)
- Bruno, R. (2019). Intermittency in solar wind turbulence from fluid to kinetic scales. *Earth and Space Science*, *6*, 656-672. doi: <https://doi.org/10.1029/2018EA000535>
- Bruno, R., & Carbone, V. (2013). The solar wind as a turbulence laboratory. *Living Reviews in Solar Physics*, *10*, 2(2). doi: 10.12942/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*.
- Burlaga, L. F. (1991). Intermittent turbulence in the solar wind. *JGR Space Physics*.
- Chang, T. (2015). *An introduction to space plasma complexity*. Cambridge University Press. doi: <https://doi.org/10.1017/CBO9780511980251>
- Chang, T., & Wu, C. C. (2008). Rank-ordered multifractal spectrum for intermittent fluctuations. *Physical Review E*, *77*(4). doi: 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*, 545-551. doi: 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. doi: 10.5194/npg-18-277-2011
- Daubechies, I. (1992). Ten lectures on wavelets. *CBMS-NSF Regional Conference Series in Applied Mathematics*. (ISBN: 978-0-89871-274-2) doi: <https://doi.org/10.1137/1.9781611970104>
- Deak, N., Creț, O., Echim, M., Teodorescu, E., Negrea, C., Văcariu, L., ... Hângan, A. (2018). Edge computing for space applications: Field programmable gate array-based implementation of multiscale probability distribution functions. *Review of Scientific Instruments*, *89*(12), 125005. doi: 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. doi: <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. doi: 10.1007/s11214-013-9974-9
- Echim, M., Chang, T., Kovacs, P., Wawrzaszek, A., Yordanova, E., Narita, Y., ... Consolini, G. (2021). Turbulence and complexity of magnetospheric plasmas. In *Magnetospheres in the solar system* (p. 67-91). American Geophysical Union (AGU). doi: <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. doi: 10.5194/npg-14-525-2007

- 700 Farge, M. (1992). Wavelet transforms and their applications to turbulence. *Annual*  
 701 *Review of Fluid Mechanics*, *24*, 395-457. doi: 10.1146/annurev.fl.24.010192  
 702 .002143
- 703 Frisch, U. (1995). *Turbulence: The Legacy of A.N. Kolmogorov*. Cambridge Univer-  
 704 sity Press. doi: <https://doi.org/10.1017/CBO9781139170666>
- 705 Goldreich, P., & Sridhar, S. (1995). Toward a theory of interstellar turbulence. ii.  
 706 strong alfvénic turbulence. *Astrophysical Journal*, *438*. doi: 10.1086/175121
- 707 Greco, A., Matthaeus, W. H., Perri, S., Osman, K. T., Servidio, S., Wan, M., &  
 708 Dmitruk, P. (2018). Partial variance of increments method in solar wind  
 709 observations and plasma simulations. *Space Science Reviews*, *214*(1). doi:  
 710 <https://doi.org/10.1007/s11214-017-0435-8>
- 711 Hergarten, S. (2002). *Self-organized criticality in earth systems*. Springer, Berlin,  
 712 Heidelberg. doi: <https://doi.org/10.1007/978-3-662-04390-5>
- 713 Hnat, B., Chapman, S. C., Rowlands, G., Watkins, N. W., & Farrell, W. M.  
 714 (2002). Finite size scaling in the solar wind magnetic field energy density  
 715 as seen by wind. *Geophysical Research Letters*, *29*(10), 86-1-86-4. doi:  
 716 10.1029/2001GL014587
- 717 Howes, G. G., Cowley, S. C., Dorland, W., Hammett, G. W., Quataert, E., &  
 718 Schekochihin, A. A. (2008). A model of turbulence in magnetized plasmas:  
 719 Implications for the dissipation range in the solar wind. *J. Geophys. Res.*,  
 720 *113*(A05103). doi: <https://doi.org/10.1029/2007JA012665>
- 721 Huang, S. Y., Hadid, L. Z., Sahraoui, F., Yuan, Z. G., & Deng, X. H. (2017).  
 722 On the existence of the kolmogorov inertial range in the terrestrial mag-  
 723 netosheath turbulence. *The Astrophysical Journal Letters*, *836*(1). doi:  
 724 <https://doi.org/10.3847/2041-8213/836/1/L10>
- 725 Lion, S., Alexandrova, O., & Zaslavsky, A. (2016). Coherent events and spectral  
 726 shape at ion kinetic scales in the fast solar wind turbulence. *The Astrophysical*  
 727 *Journal*, *824*(1). doi: <https://doi.org/10.3847/0004-637X/824/1/47>
- 728 Marsch, E., & Tu, C. Y. (1994). Non-gaussian probability distributions of solar  
 729 wind fluctuations. *Ann. Geophys.*, *12*(H15), 1127-1138. doi: <https://doi.org/10.1007/s00585-994-1127-8>
- 730
- 731 Matthaeus, W. H., Wan, M., Servidio, S., Greco, A., Osman, K., Oughton, S., &  
 732 Dmitruk, P. (2015). Intermittency, nonlinear dynamics and dissipation in the  
 733 solar wind and astrophysical plasmas. *Philosophical Transactions of the Royal*  
 734 *Society A: Mathematical, Physical and Engineering Sciences*.
- 735 Munteanu, C., Turicu, D. C., Creț, O., & Echim, M. (2022). Detecting disconti-  
 736 nuousities from in situ space measurements: Method and fpga implementation.  
 737 *Earth and Space Science*, *9*(10). doi: <https://doi.org/10.1029/2022EA002537>
- 738 Opincariu, L., Deak, N., Creț, O., Echim, M., Munteanu, C., & Văcariu, L. (2019).  
 739 Edge computing in space: Field programmable gate array-based solutions for  
 740 spectral and probabilistic analysis of time series. *Review of Scientific Instru-*  
 741 *ments*, *90*(11), 114501. doi: 10.1063/1.5119231
- 742 Roberts, O. W., Alexandrova, O., Kajdic, P., Turc, L., Perrone, D., Escoubet, C. P.,  
 743 & Walsh, A. (2017). Variability of the magnetic field power spectrum in  
 744 the solar wind at electron scales. *The Astrophysical Journal*, *850*(2). doi:  
 745 <https://doi.org/10.3847/1538-4357/aa93e5>
- 746 Sorriso-Valvo, L., Carbone, V., Giuliani, P., Veltri, P., Bruno, R., Antoni, V.,  
 747 & Martinez, E. (2001). Intermittency in plasma turbulence. *Plane-*  
 748 *tary and Space Science*, *49*(12), 1193-1200. doi: [https://doi.org/10.1016/S0032-0633\(01\)00060-5](https://doi.org/10.1016/S0032-0633(01)00060-5)
- 749
- 750 Sreenivasan, K. R. (1991). Fractals and multifractals in fluid turbulence. *Annual*  
 751 *Review of Fluid Mechanics*, *23*(1), 539-604. doi: 10.1146/annurev.fl.23.010191  
 752 .002543
- 753 Teodorescu, E., & Echim, M. M. (2020). Open-source software analysis tool to in-  
 754 vestigate space plasma turbulence and nonlinear dynamics (odyn). *Earth and*

- 755 *Space Science*. doi: <https://doi.org/10.1029/2019EA001004>
- 756 Teodorescu, E., Echim, M. M., & Voitcu, G. (2021). A perspective on the scaling of  
757 magnetosheath turbulence and effects of bow shock properties. *The Astrophys-*  
758 *ical Journal*, *910*(1). doi: <https://doi.org/10.3847/1538-4357/abe12d>
- 759 Torrence, C., & Compo, G. P. (1998). A practical guide to wavelet analy-  
760 sis. *Bulletin of the American Meteorological Society*, *79*, 61-78. doi:  
761 [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)
- 762 Turicu, D. C., Creț, O., Echim, M., & Munteanu, C. (2022). An fpga-based solu-  
763 tion for computing a local stationarity measure from satellite data. *IEEE Ac-*  
764 *cess*, *10*, 9668-9676. doi: [10.1109/ACCESS.2022.3143239](https://doi.org/10.1109/ACCESS.2022.3143239)
- 765 Vörös, Z., Baumjohann, W., Nakamura, R., Runov, A., Zhang, T. L., Volw-  
766 erk, M., ... Rème, H. (2003). Multi-scale magnetic field intermit-  
767 tence in the plasma sheet. *Annales Geophysicae*, *21*(9), 1955–1964. doi:  
768 [10.5194/angeo-21-1955-2003](https://doi.org/10.5194/angeo-21-1955-2003)
- 769 Wawrzaszek, A., & Echim, M. (2021). On the variation of intermittency of  
770 fast and slow solar wind with radial distance, heliospheric latitude, and so-  
771 lar cycle. *Frontiers in Astronomy and Space Sciences*, *7*. doi: [10.3389/](https://doi.org/10.3389/fspas.2020.617113)  
772 [fspas.2020.617113](https://doi.org/10.3389/fspas.2020.617113)
- 773 Wawrzaszek, A., Echim, M., Macek, W. M., & Bruno, R. (2015). Evolution of inter-  
774 mittency in the slow and fast solar wind beyond the ecliptic plane. *The Astro-*  
775 *physical Journal Letters*, *814*(2), L19.
- 776 Welch, P. D. (1967). The use of fast fourier transform for the estimation of power  
777 spectra: A method based on time averaging over short, modified periodograms.  
778 *IEEE Trans. Audio Electroacoust.*, *Volume AU-15*, p. 70-73, 15, 70-73.
- 779 Yordanova, E., Bergman, J., Consolini, G., Kretzschmar, M., Materassi, M.,  
780 Popielawska, B., ... Wernik, A. W. (2005). Anisotropic scaling features  
781 and complexity in magnetospheric-cusp: a case study. *Nonlinear Processes in*  
782 *Geophysics*, *12*(6), 817–825. doi: [10.5194/npg-12-817-2005](https://doi.org/10.5194/npg-12-817-2005)