Multi-signal Multifractal Detrended Fluctuation Analysis (MS-MFDFA) Software

Version 1.0 - User manual

Juan Carlos de la Torre Macías
Pablo Pavón Domínguez
Patricia Ruíz Villalobos
Bernabé Dorronsoro Díaz
Soledad Moreno Pulido

October 2023

Índice

1. Introduction

In the literature, there are few studies that address low-level directives. These directives consider the underlying architecture to create green programs. One of the main challenges is the strong knowledge required. Another challenge is the high level of uncertainty in consumption measurements. These measurements can be affected by existing software services, hardware, and ambient temperature, among other factors.

In our work, titled *Multi-signal Multifractal Detrended Fluctuation Analysis for Uncertain Systems. Application to the Energy Consumption of Software Programs in Microcontrollers*, we introduce a new approach. For the first time, we propose the characterization of software consumption using the Multifractal approach. This generates a signature of the software consumption profile. This signature is independent of the runtime, consumption levels, and uncertainty.

To achieve this, we construct a consumption signal. This signal, or time series, is derived from high-frequency sampling. It captures the instantaneous input current during the program's operation. This signal is then analyzed using MF-DFA.

There's a challenge with the uncertainty present in the energy consumption measurements. Analyzing a single time series doesn't provide significant conclusions. As a result, many experiments are required. To address this, we propose a novel methodology. This methodology can handle a large number of time series simultaneously during the multifractal analysis.

Additionally, we propose eight programs for analysis. These can serve as a reference for others. We have developed software that can automatically perform this analysis. This software has been made freely available to the scientific community.

The MS-MFDFA software achieves four main objectives:

- **Preprocessing**. Allows for trimming signals and addressing potential anomalous data (*outliers*).
- **Decomposition**. Enables the decomposition of signals into their trend, seasonality, and residue components, storing the residue of each signal for subsequent analysis.
- **Multianalysis**. Facilitates the MF-DFA analysis of the set of signals, producing the *Assembled* fluctuation function, which represents the set of fluctuation functions resulting from applying the MF-DFA algorithm to each signal.

- **Multianalysis Quality**. Allows for the generation of a summary table of the quality of the fit performed on the *Assembled* fluctuation function in relation to all fluctuation functions. The summary is provided in table form and is available in both *.xlsx* and *.tex* formats.

This software has been implemented in Matlab R2022b and is optimized for a screen resolution of 2560 x 1440 pixels. The self-installation version is available for use on Windows 10 or later.

2. Installation

There are two options to install and run the application:

- Option 1: *CMSeFracWare1D.mlappinstall*. This requires the user to have Matlab installed. Executing the file will install the application in the Matlab environment that the user has installed. The application will be located in the *APP* menu of Matlab.
- Option 2: *CMSeFracWare1D.exe*. This does not require the user to have Matlab installed. Running the file will launch an application installer for Windows systems and, if the user doesn't have it, it will automatically install the Matlab runtime engine. After installation, the user can run the software in the same way as any other Windows application.

3. Interface Description

Upon launching the application through any of the options mentioned in section 2, the initial screen shown in Figure 1 will appear, which corresponds to the **Home** tab. This tab provides direct access to each of the software's functionalities, which will be described further below. When clicking on a functionality button, the program will automatically open its corresponding tab. We can also access the desired functionality tab by directly clicking on the tab at the top of the screen. Generally, all application buttons are presented in blue. Therefore, on this screen, we have access to the six functionalities offered by the program:

- **Tab 1: *Preprocessing***. In this tab (Figure 2), we can trim a signal and even address anomalous data if they are found in our signal. The components of the tab are:

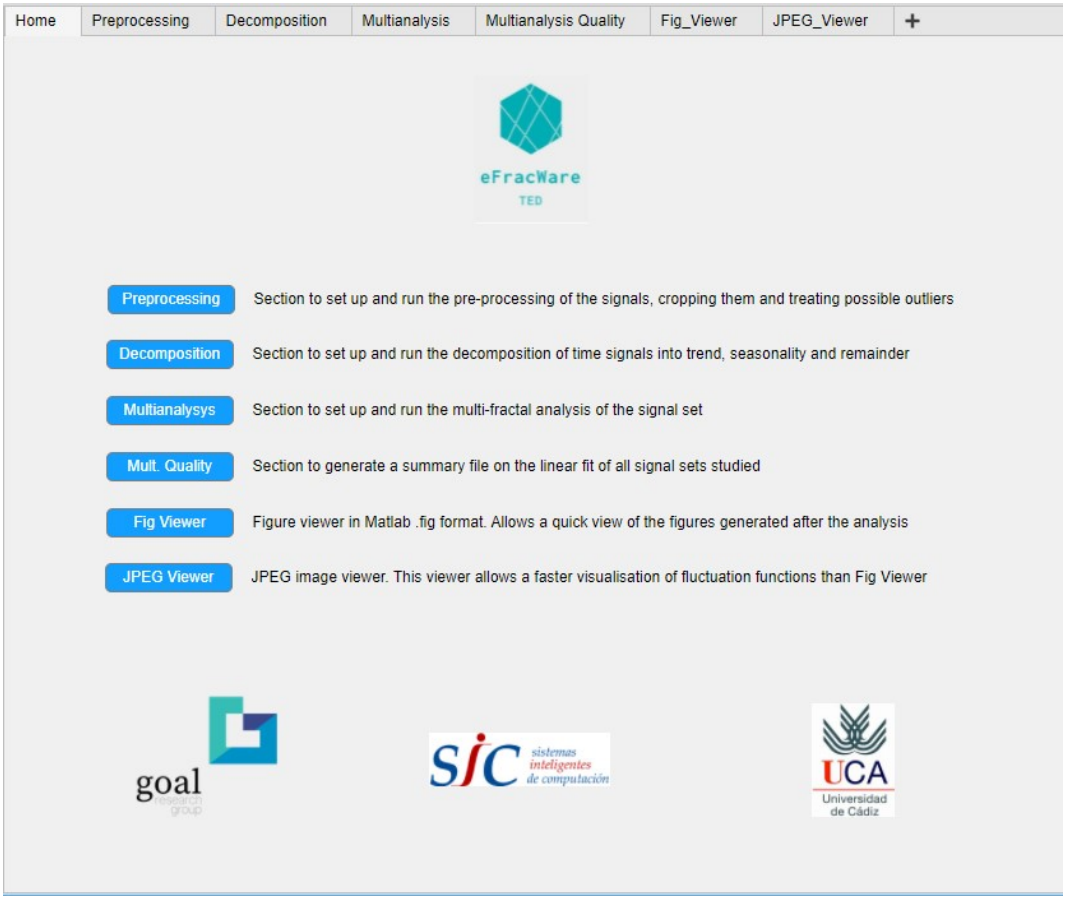

Figura 1: Home Tab.

- **Button *Source Folder***. Allows the user to select the folder containing the signals. Signal data should be in column format and in separate text files.
- **Button *Sig. Out. Folder***. Allows the user to select the folder where the processed signal results will be stored.
- **Button *Summary Folder***. Allows the user to select the folder where the *summary_preprocessed_signals.xlsx* file will be located, which will contain a summary table. This summary table displays data for both the original and processed signals. The provided data includes the smallest sample length, average value, standard deviation, maximum value, and minimum value from all signals. Lastly, it indicates the percentage of outlier values (calculated as selected in the program) in the signal with the highest number of these values.

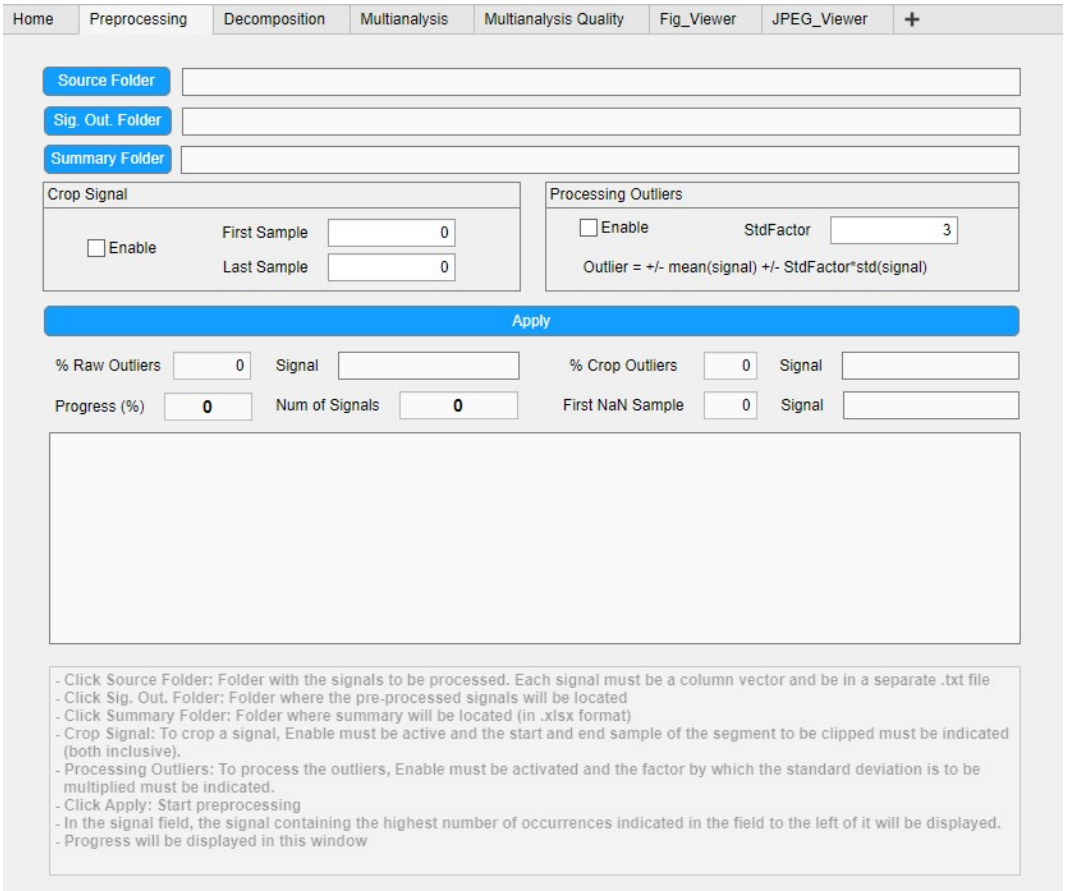

Figura 2: Preprocessing Tab.

- **Panel *Crop Signal***. Allows the user to choose whether to trim the signals by checking the box next to *Enable*. If desired, the user should input both the starting and ending sample numbers (both will be included in the trimmed signal).
- **Panel *Processing Outliers***. Allows the user to choose whether to address outlier data by checking the box next to *Enable*. If desired, the user should input the multiplicative factor over the standard deviation to consider a data point as an outlier.
- **Button *Apply***. Executes preprocessing based on the selected options.
- ***Information Section***. Located below the *Apply* button. The provided fields are:
    - **%Raw Outliers**. Indicates the highest percentage of detected outliers from all signals.

    - **Signal (next to % Raw Outliers)**. Displays the file name of the signal with the highest number of outliers.
    - **% Crop Outliers**. Indicates the highest percentage of detected outliers from all trimmed signals.
    - **Signal (next to % Crop Outliers)**. Displays the file name of the trimmed signal with the highest number of outliers.
    - **Progress (%)**. Indicates the preprocessing execution percentage.
    - **Num of Signals**. Displays the total number of processed signals.
    - **First NaN Sample**. Indicates the smallest sample with a *NaN* value from all signals.
    - **Signal (next to First NaN Sample)**. Displays the file name of the signal with the *NaN* value in its smallest sample from all original signals. This data is very useful for configuring signal trimming so that all have the same size for subsequent analysis.
    - **Message Box**. Displays the steps the program is taking until preprocessing completion.

- **Tab 2: *Decomposition***. In this tab (Figure 3), we can perform the decomposition of signals into their trend, seasonality, and residue components. The residue (resulting from subtracting the trend and seasonality from the signal) of each signal will be stored in the folder specified by the user. Computationally, this is the most demanding step. For this reason, if this step is executed, a *threadpool* will be initiated, allowing its parallel execution using all available processors in the system. The components of the tab are:

    - **Button *Source Folder***. Allows the user to select the folder containing the signals, whether they are original or preprocessed as deemed appropriate by the user. Signal data should be in column format and in separate text files.
    - **Button *Output Folder***. Allows the user to select the folder where the residues will be stored.
    - **Panel *Algorithm***. Allows the user to choose the algorithm to use (STL or SDA) for signal decomposition.
    - **Panel *STL Set up***. Allows the user to select the necessary parameters to apply the STL algorithm (if it was chosen in the previous panel). The options are:

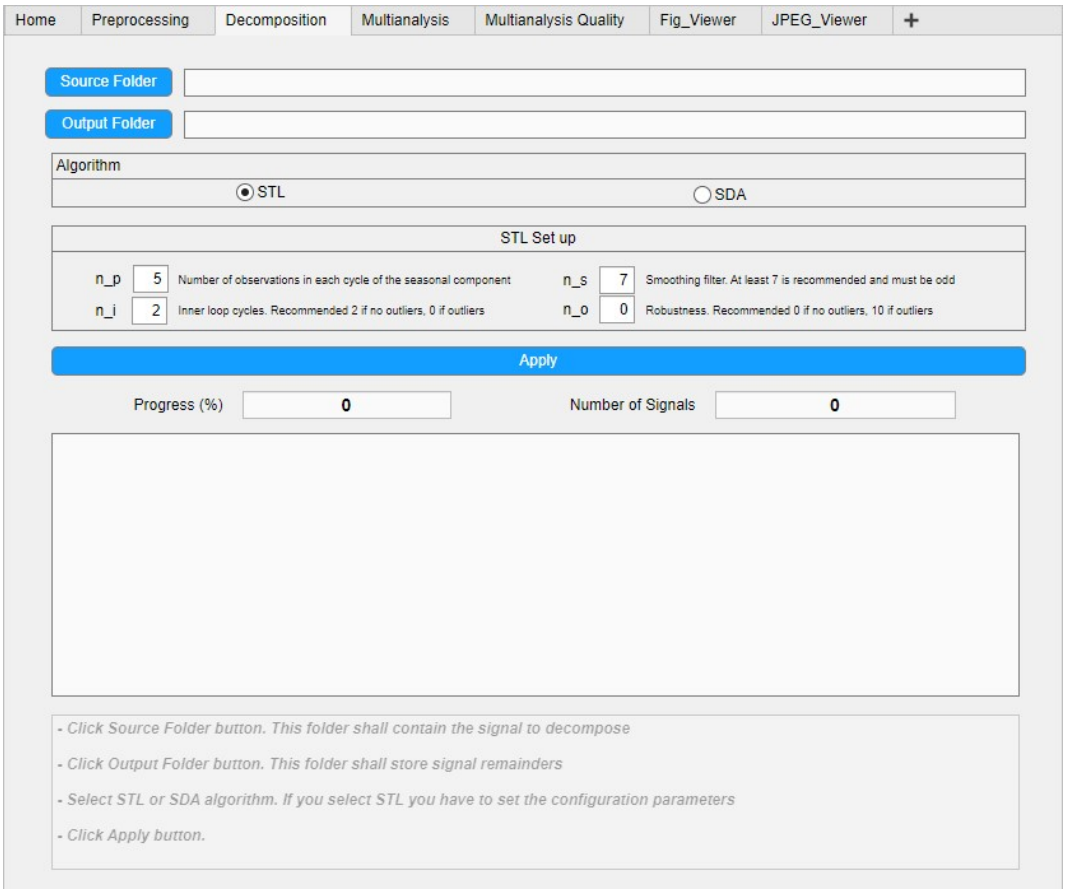

Figura 3: Decomposition Tab.

- $\mathbf{n}_p$. Number of observations in each cycle of the stationary component. The default value provided is 5.
- $\mathbf{n}_i$. Number of times the algorithm will be applied. The default value provided is 2, recommended when there are no outliers in the signal.
- $\mathbf{n}_s$. Value for the smoothing filter. The default value provided is 7.
- $\mathbf{n}_o$. Value representing the robustness of the algorithm. The default value provided is 0, recommended when there are no outliers in the signal.

- **Button *Apply***. Executes the decomposition of the signals based on the selected options.
- ***Information Section***. Located below the *Apply* button. The provided fields are:

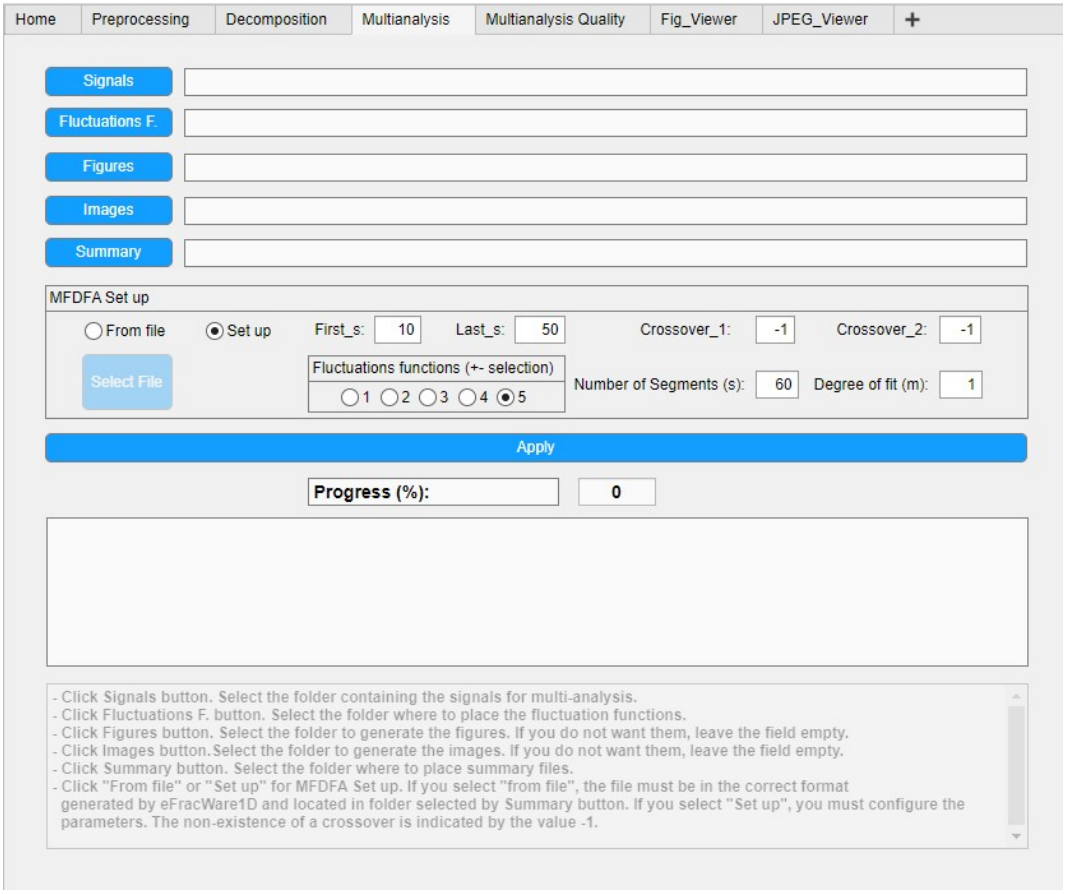

Figura 4: MF-DFA Algorithm Application Tab.

- **Progress %**. Indicates the execution percentage.
- **Number of Signals**. Displays the total number of processed signals.
- **Message Box**. Displays the steps the program is taking until the decomposition is complete.

- **Tab 3: *Multianalysis***. In this tab (Figure 4), we can apply the MF-DFA algorithm to the set of signals under study, obtain the fluctuation function of each signal, the representative fluctuation function of the signals, and the linear regression performed on the representative fluctuation function and its quality in relation to the fluctuation functions of each signal. The components of the tab are:

  - **Button *Signals***. Allows the user to select the folder containing the signals, whether they are original, preprocessed, or the residues

after decomposition, as deemed appropriate by the user. Signal data should be in column format and in separate text files.

- **Button *Fluctuations F.***. Allows the user to select the folder where the fluctuation function data will be stored. An *.xlsx* file will be generated for each signal with the same name, containing its fluctuation function data. Additionally, in the same folder, a file named *Assembly_fluctuation_functions.xlsx* will be generated, containing the representative fluctuation function values of all signals.
- **Button *Figures***. Allows the user to select the folder that will contain the fluctuation function graphs in Matlab's *.fig* format. If no folder is selected, figures will not be generated.
- **Button *Images***. Allows the user to select the folder that will contain the fluctuation function graphs in *.jpeg* format. This requires that the figures indicated in the previous button have been generated. If no folder is selected, images will not be generated.
- **Button *Summary***. Allows the user to select the folder that will contain the following files:
    - **LinearRegressions.xlsx**. A file containing the results of executing linear regression on the representative fluctuation functions and the quality of the fit in relation to all signal fluctuation functions. The file contains a summary of the fit quality for each segment and signal.
    - **Assembly_fluctuation_functions.xlsx**. A file containing the representative fluctuation functions of all signals.
    - **Assembly_fluctuation_functions_std.xlsx**. A file containing the standard deviation of the representative fluctuation functions in relation to the fluctuation functions of all signals.
- **Panel *MFDFA Set up***. Allows the user to choose whether to load configuration parameters from a file (option *From file*) or directly from the program (option *Set up*). If *Set up* is chosen, fields to input parameters will be enabled. The options are:
    - **First_s**. The first segment of the fluctuation functions from which linear regression will be performed (if m=1). The default value provided is 10, as it's common in literature.
    - **Last_s**. The last segment of the fluctuation functions up to which linear regression will be performed (if m=1). The default value provided is N/4 (with N being the signal size), as it's common in literature.

- **Crossover_1**. The segment where a crossover point occurs. If there's no crossover point, the value -1 should be indicated (provided by default). If indicated, its value must be $First_s < Crossover_1 < Last_s$, and two linear regressions will be performed: between *First_s* and *Crossover_1* and between *Crossover_1* and *Last_s*.
- **Crossover_2**. The segment where a second crossover point occurs. If there's no crossover point, the value -1 should be indicated (provided by default). If indicated, its value must be $Crossover_1 < Crossover_2 < Last_s$, and three linear regressions will be performed: between *First_s* and *Crossover_1*, between *Crossover_1* and *Crossover_2*, and between *Crossover_2* and *Last_s*.
- **Panel *Fluctuations functions (+- selection)***. Allows the user to select the end of the fluctuation functions to be performed. For example, if 2 is selected, fluctuation functions corresponding to q=-2, q=-1, q=0, q=1, and q=2 will be performed.
- **Number of Segments (s)**. Allows the user to specify the number of segments to study on the signal.
- **Degree of fit (m)**. Allows the user to specify the polynomial degree for regression, usually m=1.

- **Button *Apply***. Executes the multianalysis of the signals based on the selected options.
- **Progress %**. Indicates the execution percentage of each of the analyses being performed.
- **Message Box**. Displays the steps the program is taking until the analyses are complete.

- **Tab 4: *Multianalysis Quality***. In this tab (Figure 5), we can obtain a summary table displaying the quality of the obtained representative fluctuation functions, both in *.xlsx* and *.tex* formats. Additionally, we will obtain the dimension functions of each segment on which linear regressions were performed in *.xlsx* format. The components of the tab are:

  - **Button *Input File***. Allows the user to select the *LinearRegressions.xlsx* file generated in Tab 3 *Multianalysis*.
  - **Button *Output Folder***. The folder in which the files *fluctuations_functions_summary.xlsx*, *fluctuations_functions_summary.tex*, and

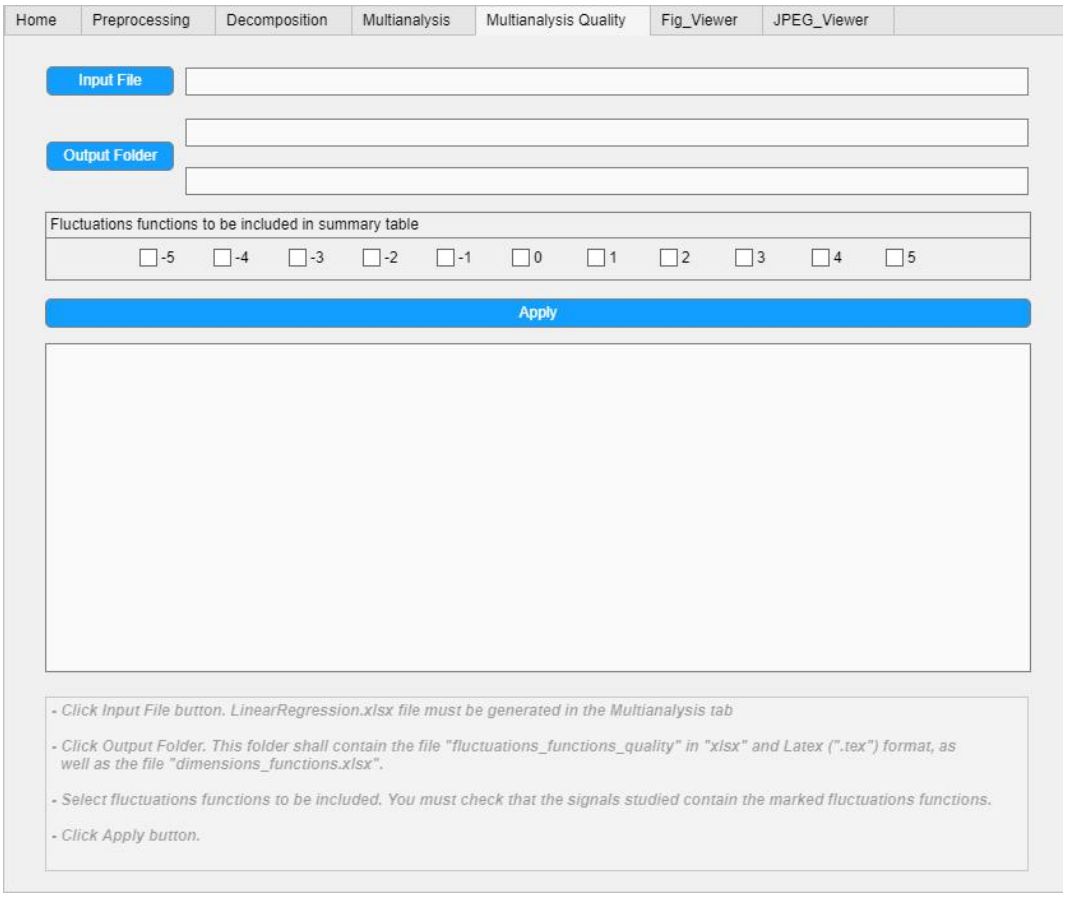

Figura 5: Tab to obtain a summary of the study's quality and the dimension functions.

dimension_function.xlsx will be generated. Each of the first two files contains the R^2 value and *RMSE* of the regressions performed for the fluctuation functions chosen as representative, the average of the functions from all signals, and their standard deviation. The last file contains the slopes of the regressions for the selected fluctuation functions.

- **Tab 5: *Fig_Viewer***. In this tab (Figure 6), we can view the figures in Matlab format obtained from the *Multianalysis* tab. This viewer allows the user to quickly and conveniently view the fluctuation functions obtained from each signal. The components of the tab are:
  - **Button *Select Figure Folder***. Allows the user to select the folder containing the figures.
  - **Display Panel**. This will display the figures.

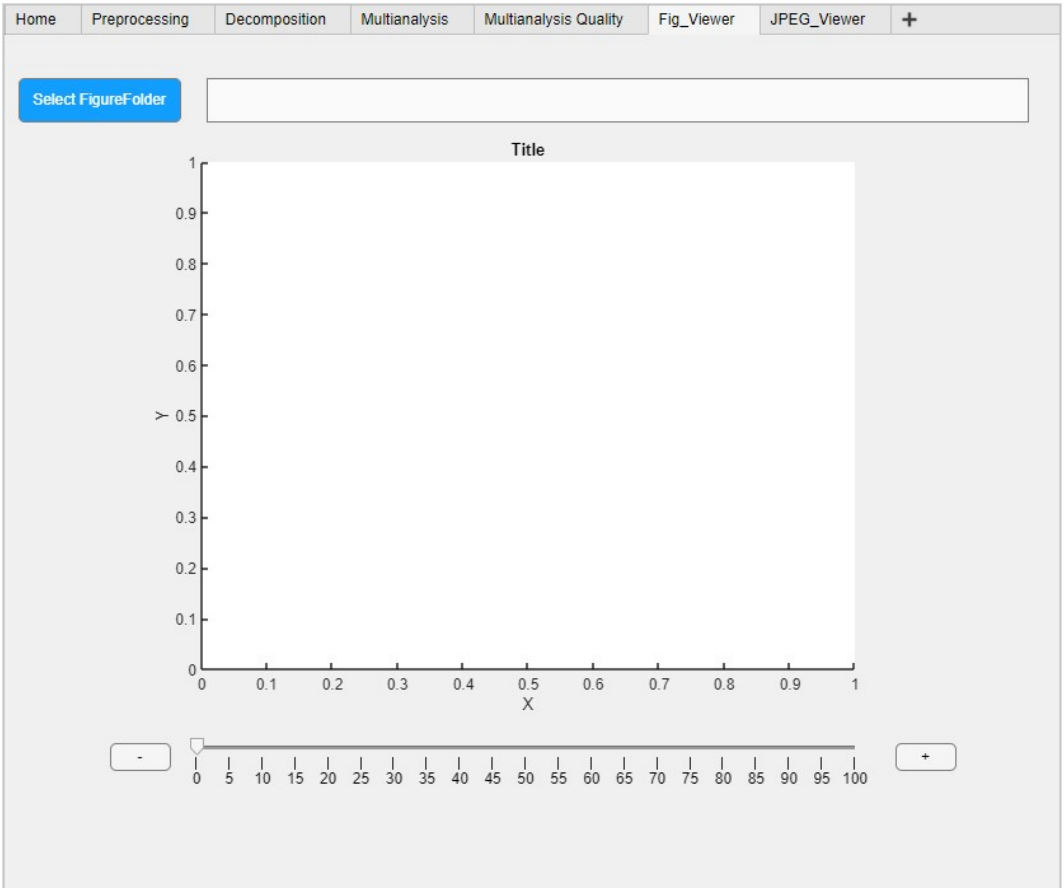

Figura 6: Tab to view figures in Matlab's *.fig* format.

- **Navigation Panel**. Allows the user to move forward or backward from the displayed figure either by pressing the + or − buttons, or by directly clicking on the bar between those buttons.

- **Tab 6: *JPEG_Viewer***. In this tab (Figure 7), we can view the images in jpeg format obtained from the *Multianalysis* tab. This viewer allows the user to see the fluctuation functions obtained from each signal faster than when using the *Fig_Viewer*, as it requires fewer computational resources. The components of the tab are:

  - **Button *Select Image Folder***. Allows the user to select the folder containing the images.
  - **Display Panel**. This will display the images.
  - **Navigation Panel**. Allows the user to move forward or backward from the displayed image either by pressing the + or − buttons, or by directly clicking on the bar between those buttons.

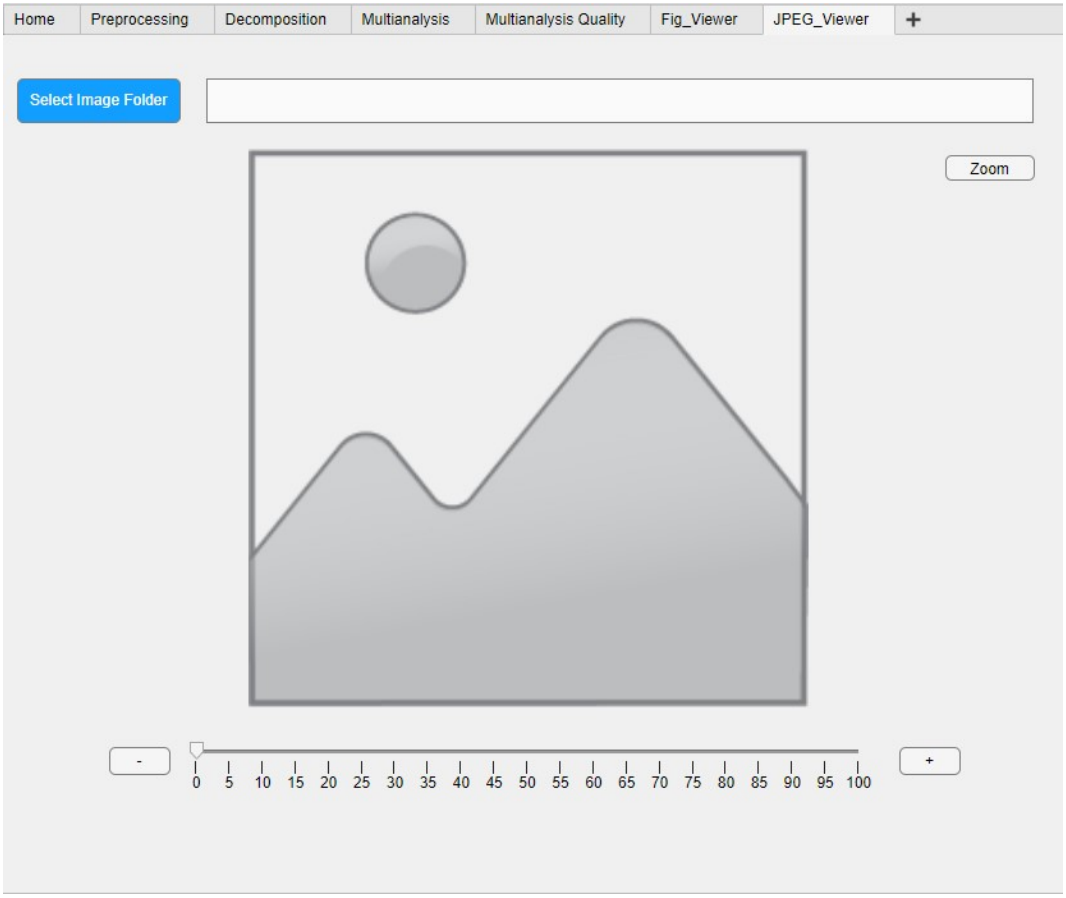

Figura 7: Tab to view images.

- **Button *Zoom***. Opens a window that allows the user to enlarge the image.