PyROEX User Manual

V1.0

Software： **Python-Based ROEX Data Analysis Software** (PyROEX)

Developer: Mingyi Zhang
Address: Zhongguancun Nan'ertiao 1, Haidian District, Beijing 100190, China
E-mail: zhangmingyi24@mails.ucas.ac.cn

Mar. 2025

Table of Contents

1 Overview

The application of GNSS radio occultation (RO) technology in the meteorological field can be traced back to the GPS/MET experiment conducted by the United States in 1995. This project demonstrated the feasibility of using GNSS occultation technology to probe the Earth's atmosphere and ionosphere, and directly contributed to the subsequent rapid development of occultation missions. These include Germany's CHAMP satellite Argentina's SAC-C satellite, the joint U.S.-Germany GRACE mission the U.S.-Taiwan COSMIC satellite series, and China's FengYun satellite series. To date, GNSS radio occultation technology has become relatively mature, offering advantages such as global coverage, high precision, high vertical resolution, long-term stability, and all-weather observation. It has provided significant advancements in climate research, operational weather forecasting, and space weather studies.

The current international standard for GNSS receiver data exchange is the RINEX, which was proposed by Werner Gurter during the first large European GPS campaign EUREF 89 in 1989. The primary purpose of the RINEX was to facilitate the comprehensive processing of GNSS positioning network data, making it mainly suitable for GNSS positioning and orbit determination users. However, it is not entirely suitable for radio occultation observations. For instance, conventional geodetic GNSS receivers typically use closed-loop tracking, while GNSS radio occultation receivers generally employ open-loop tracking at the lower tracking stages. Some parameters related to open-loop tracking also need to be recorded in the observation file, but the current RINEX file does not include definitions for these parameters. In the industry, there are significant differences in the definition of GNSS radio occultation data. For example, the radio occultation data from COSMIC satellites and MetOp satellites use custom binary formats developed by researchers. Due to the lack of a unified standard, researchers must develop specialized data reading interfaces for each radio occultation mission when processing the data. This practice is not conducive to data exchange, unified management, and joint processing. The importance and urgency of establishing a standard for radio occultation data exchange were highlighted in the summary of the 8th International Radio Occultation Working Group meeting[1]. In response to these issues, the National Space Science Center of China, based on current GNSS standard data file formats such as RINEX, ANTEX, and IONEX, proposed the GNSS Radio Occultation Data Independent Exchange Format (ROEX). A brief introduction to the ROEX format will be provided below, and more detailed information can be found at (website). The ROEX has already been applied in actual scientific research, and its feasibility has been demonstrated.

The National Space Science Center of China has developed a software named "PyROEX" for ROEX files. This software is a graphical user interface tool written in the Python programming language, offering users options for data monitoring, viewing data combinations and integrity, as well as file editing.

[1] https://irowg.org/workshops/irowg-8/

2 Software Requirements

2.1 Environment requirement

The executable program in the compressed package is built using Conda 24.11.3 and Python 3.11.7, along with the Spyder 6.0.0 integrated development environment. The computer system requirements are as follows:

- Operating System: Windows, Linux, Mac
- System Type: 64-bit
- Memory: At least 512MB
- Hard Disk Space: At least 500MB

2.2 Software license

PyROEX is open-source software licensed under the GNU General Public License (Version 3) (https://www.gnu.org/licenses/gpl-3.0.html). Users can redistribute or modify it under the terms of the GNU General Public License as published by the Free Software Foundation.

2.3 Software installation

After downloading the PyROEX compressed file, please follow these steps to compile the software:

- **Step 1**: Extract the PyROEX.zip file. Users of different operating systems can use different commands for the extraction process:

  For Windows users, open the cmd console and navigate to the folder where PyROEX.zip is located using the cd command. Then, enter the following command in the console:

  ```
  7z x PyROEX.zip
  ```

  enter the following command in the terminal to extract the PyROEX.zip file:

  ```
  unzip PyROEX.zip -d target_folder
  ```

  The user will receive three folders: doc, data, and src. The doc folder contains the user manuals in both Chinese and English, as well as the ROEX documentation in both Chinese and English. The data folder contains atmospheric and ionospheric occultation observation files used for analysis. The src folder contains Python source code.

- **Step 2**: Use the `cd` command to navigate to the src folder, then execute the installation command in the terminal to batch install the necessary libraries by using Python's pip command. Enter the following:

  ```
  pip install -r requirements.txt
  ```

- **Step 3**: To run the PyROEX program, enter the following command in the terminal:

```
Python main.py
```

3 Software Tool Instructions

The PyROEX software mainly includes three functional modules: Data Monitoring, Data Combination, Data Integrity and File Cutting. The overall structure is shown in Figure 1.

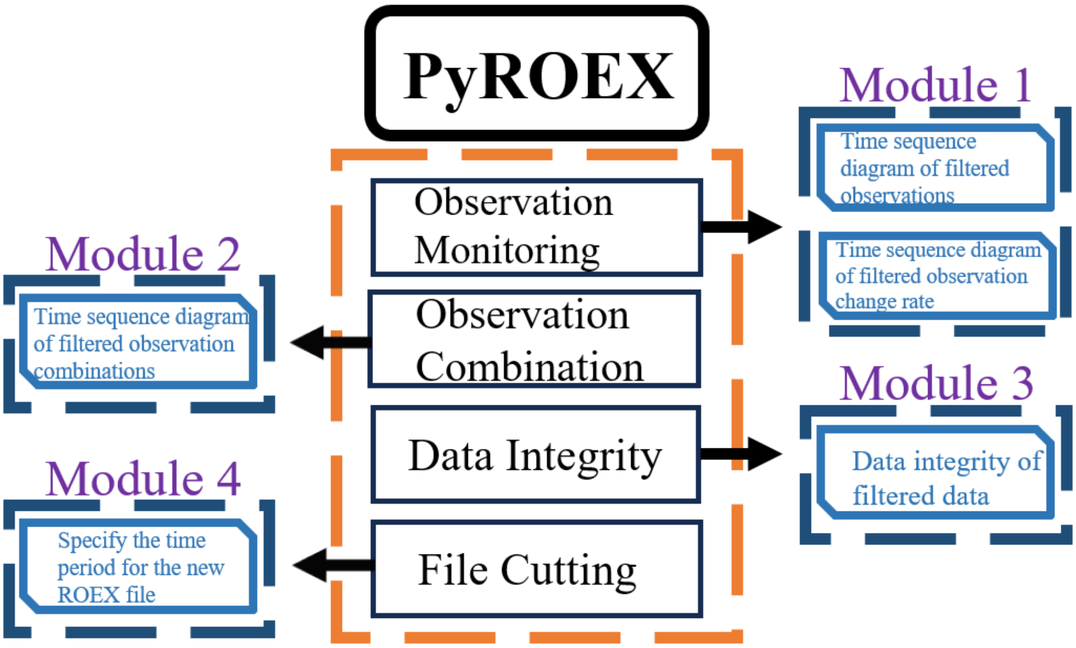

Fig.1. Software modules and functions

For the convenience of explanation, the following abbreviations are introduced:

- GNSS Radio Occultation Observation Independent Exchange Format: ROEX
- GNSS atmospheric occultation observation independent exchange format: ROEX-A
- GNSS ionospheric occultation observation independent exchange format: IROEX-I
- Occultation satellite: OCC
- Reference satellite: REF

3.1 File Cutting

If users need to extract data within a period of good quality, they can use the file cutting module to obtain new ROEX files that conform to the time set by the users. The interface of the file editing module is shown in Figure 1. For the ROEX-A file, users need to input five quantities, including the start time of closed-loop observation, the end time of closed-loop observation, the start time of open-loop observation, the end time of open-loop observation, and the output file name. For ROEX-I files, users only need to enter the start time of the observation, the end time of the observation and the output file name. After entering the information, click "Execute". The input time format is "YY-MM-DD HH:MM:SS". It should be noted that the seconds of the input time need to have a decimal point. For example, entering

"2024-2-21 18:42:56.00" is feasible, but entering "2024-2-21 18:42:56" will result in an error. The new ROEX file will be named after the input output-file-name plus ".rox ", and the generated file will be located in the data folder of the parent folder where the execution file is situated. The new ROEX file strictly adheres to the ROEX format.

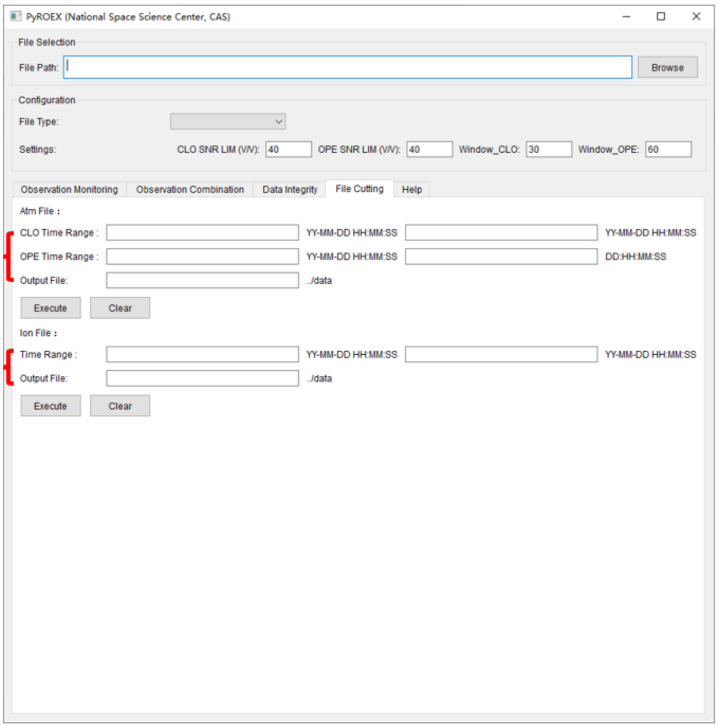

Fig.2. File cutting interface

3.2 Observation Monitoring

Through the observation monitoring module, users can obtain visualized images of ROEX file data. After running the program, users should click "Input" to import the ROEX file for analysis, then select the file type. The specific steps are shown in Figure 3.

Fig.3. Parameter configuration operation illustration

The user selects the file type in the configuration bar, sets the signal-to-noise ratio threshold and the sliding window length. The filtering principle is as follows: When the average sliding signal-to-noise ratio of a certain epoch is less than the set signal-to-noise ratio threshold, the data after that epoch is discarded and implemented in the python program using the SNR_filter() function. For the closed-loop and open-loop periods in the ROEX-A file, the sliding window sizes are 30 and 60 respectively, corresponding to a duration of 1 second. For the ROEX-I file, the sliding window size is 60 and the corresponding duration is 1 minute.

The default signal-to-noise ratio threshold for all is 40 V/V. The ROEX-A file contains the signal data of the reference star that does not participate in the occultation event. Since these data do not include the signal-to-noise ratio, the data of the reference star has not been filtered.

After setting the SNR threshold, the user determines the data type to be plotted by selecting "Original Observation" or "Rate of Change". In the "Original Observation" interface, the visualization of various observations in the ROEX file is implemented, while in the "Rate of Change" interface, the plotting of the rate of change of observations calculated by difference is achieved.

For the meanings of the identifiers in the interface, please refer to the ROEX documentation. Figure 4 shows the basic steps for drawing the timing diagram of the signal-to-noise ratio of different frequencies in the reference ROEX-A file. The red arrows point to the channels of the custom image's horizontal and vertical coordinate labels is implemented.

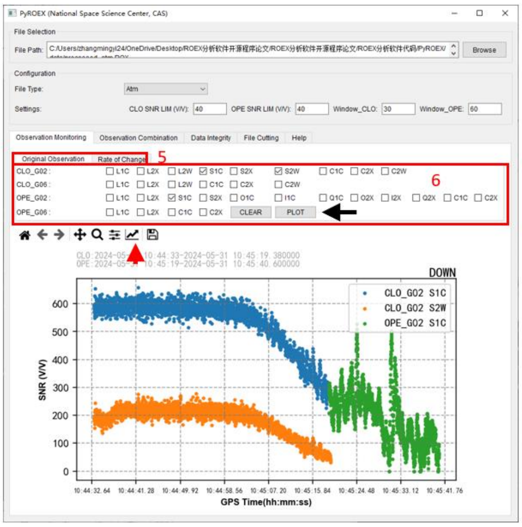

Fig.4. Observation monitoring interface and SNR results

3.3 Observation Combination

In the "Observation Combination" module, users can draw combination time series diagrams among observations, including MW combination, GF combination, IF combination and DIF combination. The specific operation steps and results are shown in Figure 5. First, select the File Type under "File Type", and then the user still needs to set the signal-to-noise ratio threshold.

After selecting the file type, users can draw a combination of data from the specified satellite under the specified observation mode in the four tabs of "MW", "GF", "IF" and "DIF". This function can be used to detect data quality. Please note:

- The image drawn here is not a combined time sequence diagram, but a combined first-order difference time sequence diagram. For specific combination principles, please refer to Section Four.
- When combining data from ROEX-I files, there is no "DIF" option.

Figure 5 shows the result of drawing the GF combination of individual data in the sample ROEX-A file. It should be noted that the naming of the data identiators in this module is different from that in the data monitoring module. For example, "1C2X" in the "MW" interface represents the combined result obtained from the carrier phase information "L1C", "L2X" and the pseudo-range information "C1C", "C2X" in accordance with the MW combination operation rules. In the "GF" interface, "1C2X" represents the combination result obtained from the carrier phase information "L1C" and "L2X" in accordance with the GF combination operation rules. The meaning of the identifier in the "IF" interface is the same as that in "GF".

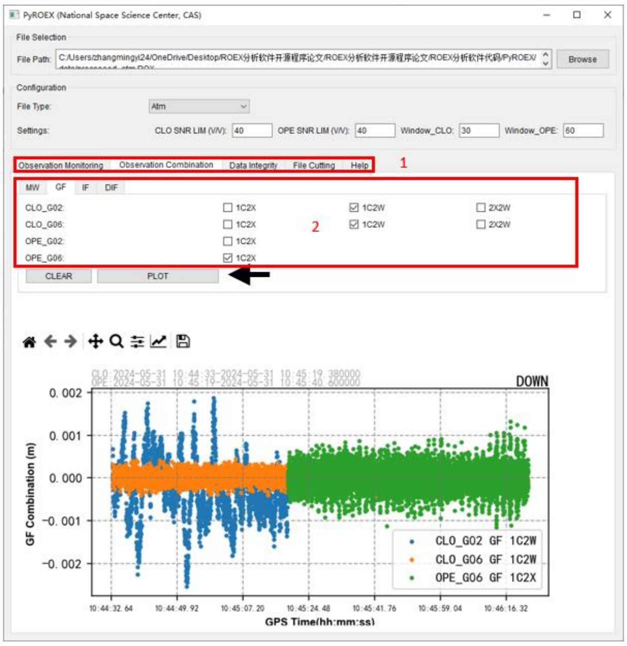

Fig.5. Steps of view observation combinations and results

3.4 Data Integrity

Users can utilize the "Data Integrity" module to view the completeness of data and locate the identifiers and times where the data is missing. The completeness of the calculated data is not the completeness of the entire file, but the completeness of the valid data after screening based on the signal-to-noise ratio threshold. Users can freely choose the data they need to view and then obtain the completeness of their data. This module will also write LOG files in the data folder of the parent folder in the running directory for users to view. The specific operation process and results are shown in Figure 6.

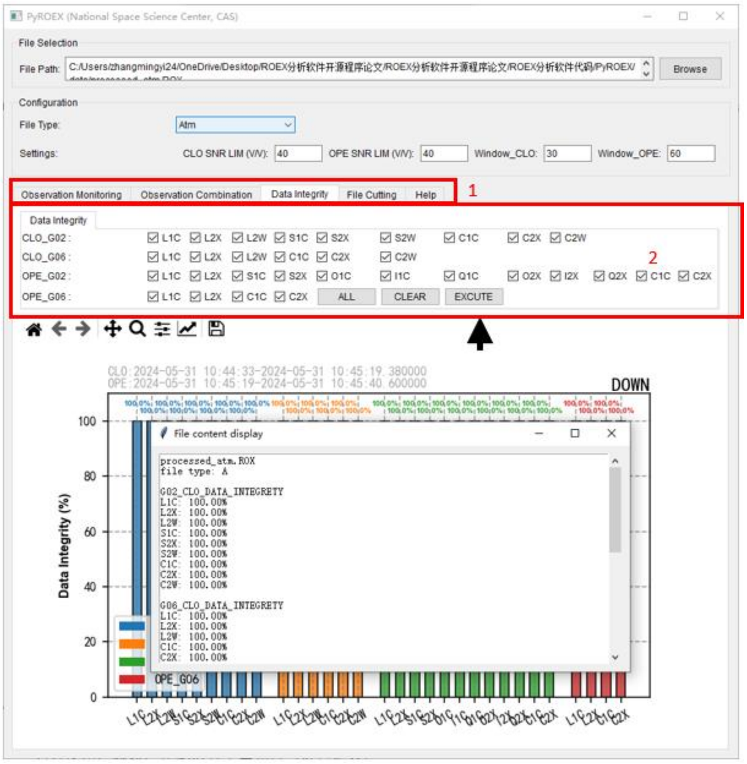

Fig.6. Steps of view data integrities and results

4 Theory

To facilitate user understanding, the following definitions are provided for each letter and symbol:

c: the speed of light in a vacuum, 299,792,458 meters per second (m/s)

L: The phase measurement value, with the unit in meters (m)

P: The phase measurement value, with the unit in cycles

C: The pseudorange measurement value, with the unit in meters (m)

f: The signal frequency, with the unit in hertz (Hz)

λ: The signal wavelength, with the unit in meters (m)

N: The integer ambiguity of the corresponding carrier phase, with the unit in cycles

4.1 Data integrity

The formula for evaluating the data completeness of the ROEX file is as follows:

$$DI = \left(\sum_i B_i \Big/ \sum_i A_i \right) \times 100\% \tag{1}$$

Among them, B_i represents the total theoretical data of the epoch that meets the signal-to-noise ratio threshold, and A_i represents the total actual observed data of this epoch. If the observed data is interrupted, the completeness of the data will decline.

4.2 MW combination

The MW (Melbourne-Wübbena) combination involves the use of dual-frequency pseudorange and carrier phase observations. Specifically, it is the difference between the wide-lane combination of carrier phase values and the narrow-lane combination of pseudorange values:

$$P_{MW} = (P_1 - P_2) - \frac{f_1 - f_2}{f_1 + f_2}\left(\frac{C_1}{\lambda_1} + \frac{C_2}{\lambda_2}\right) \tag{2}$$

The MW combination can eliminate ionospheric errors, geometric errors, tropospheric errors, and clock bias, but it is still affected by pseudorange observation noise and multipath effects. Users can combine it with the GF combination and use the TurboEdit algorithm to check for cycle slips. The first-order difference of P_{MW} is plotted in the "Quality Checking" module.

4.3 GF combination

The Geometry-Free Combination is defined as follows:

$$L_{GF} = \lambda_1 P_1 - \lambda_2 P_2 = \lambda_1 N_1 - \lambda_2 N_2 + (\gamma - 1)I \tag{3}$$

Where $\gamma = f_1^2/f_2^2$, and I represents the ionospheric delay on L_1. By performing a first-order difference on the above equation at adjacent epochs t and $t+1$, we obtain:

$$\Delta L_{GF} = L_{GF}(t+1) - L_{GF}(t) = \lambda_1 \Delta N_1 - \lambda_2 \Delta N_2 + (\gamma - 1)\Delta I \tag{4}$$

ΔI represents the ionospheric delay variation, and ΔN_1 and ΔN_2 represent the corresponding changes in the integer ambiguity. The GF combination eliminates error terms related to geometric distance, such as satellite clock bias and tropospheric delay. Since the ionospheric variations are relatively flat, ΔL_{GF} can be used to detect cycle slips. ΔL_{GF} is plotted in the "Quality Checking" module.

4.4 IF combination

By utilizing the form where the first-order ionospheric term is inversely proportional to the square of the signal frequency, the ionospheric-free combination is obtained as follows:

$$L_{IF} = \frac{f_1^2}{f_1^2 - f_2^2} L_1 - \frac{f_2^2}{f_1^2 - f_2^2} L_2 \tag{6}$$

The IF combination eliminates the first-order ionospheric term and can be used in conjunction with the GF combination results for a comprehensive analysis of the data. The first-order difference of L_{IF} is plotted in the 'Quality Checking' module. The IF combination does not eliminate the satellite clock bias. When the two received signal frequencies are the same, such as the L2X and L2W signals in the GPS system, PyROEX automatically replaces the IF combination with the GF combination.

4.5 DIF combination

The DIF combination is the single difference (at the same frequency point) of the OCC carrier phase and the REF carrier phase:

$$P_{DIF} = P_{OCC} - P_{REF} \tag{7}$$

The DIF combination eliminates the influence of satellite clock errors and other factors. The first-order difference of P_{DIF} is plotted in the "Data Combination" module.