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 and Tools Instructions

The PyROEX software mainly includes three functional modules: ROEX file data monitoring, quality control, and file editing. The overall structure is shown in Figure 1.

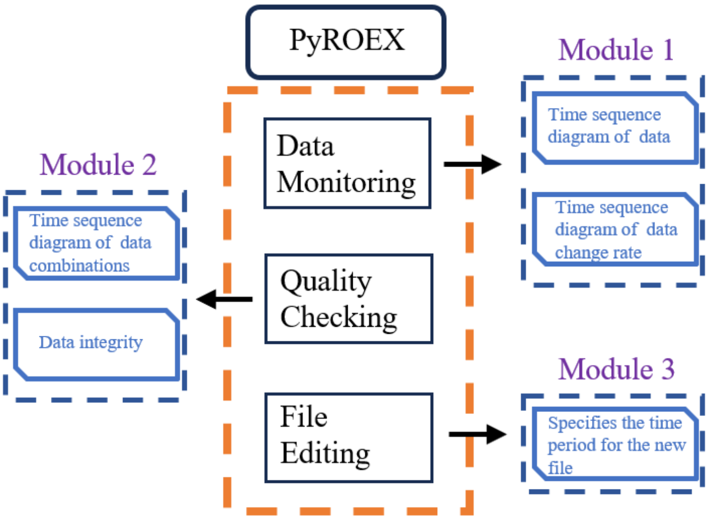

Fig.1. Software modules and functions

3.1 Data monitoring

Through the data 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 2.

Fig.2. File selection interface

After selecting the file type, users can input the signal-to-noise ratio (SNR) thresholds in the text boxes "CLO SNR LIM" and "OPE SNR LIM" to filter the data. For atmospheric occultation observation files, the software defaults the closed-loop SNR threshold to 40 dB and the open-loop SNR threshold to 0.1 dB. For ionospheric occultation observation files, users only need to set the "CLO SNR LIM". The atmospheric occultation observation files include signal data from reference satellites that are not involved in the occultation events, and since these data do not include SNR, no filtering is performed on the reference satellites data. After setting the SNR thresholds, users can select either "Original Data" or "Data Changing Rate" to determine the type of data to be plotted. The "Original Data" interface visualizes the various data in the ROEX files, while the "Data Changing Rate" interface plots the rates of change of data calculated through differential computation. It is important to note that the rate of change here is not the difference of data after SNR filtering, but the difference of the original data before filtering. The epochs in the rate of change time series are the epochs after filtering.

The meaning of the labels in the interface can be referred to in the ROEX documentation. Figure 3 illustrates the basic steps for plotting the time series of signal-to-noise ratios at different frequencies in the reference atmospheric occultation observation file. The red arrow points to the channel for customizing the horizontal and vertical axis labels of the image.

Fig.3. Data monitoring interface and SNR results

3.2 Quality checking

In the "Quality Checking" module, users can analyze the data completeness of the file and plot combination time series of the data, including MW combination, GF combination, and IF combination. The specific operation steps and results are shown in Figure 4. 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 plot the data combinations of the specified satellite under the specified observation mode in the "MW," "GF," and "IF" tabs. This function can be used to check the data quality. Please note:

- The images plotted here are not combination time series, but the first-order difference time series of the combinations. For the specific combination principles, please refer to Section 4.
- In the Python program, values that do not meet the signal-to-noise ratio threshold are set to numpy.nan, and numpy.nan will propagate through calculations, making the result also numpy.nan. When combining dual-frequency signals, the epoch range of the resulting data may differ from the epoch range of the image obtained from data monitoring.

Figure 4 shows the results of the GF combination of individual data from the example atmospheric occultation observation file. It is important to note that the data identifiers in this module are different from those in the data monitoring module. For example, "1C2X" in the "MW" interface represents the combined result obtained by applying the MW combination algorithm to the carrier phase information "L1C", "L2X", and pseudorange information "C1C", "C2X"; "1C2X" in the "GF" interface represents the combined result obtained by applying the GF combination algorithm to the carrier phase information "L1C", "L2X"; the identifiers in the "IF" interface have the same meaning as those in the "GF" interface.

Fig.4. Steps of view data combinations and results

By clicking the "Data Integrity" button, a visual representation of data integrity will be displayed. The software will automatically generate a LOG folder in the current working directory to store the data integrity files. These data integrity files are named as "LOG_DI" + ROEX file name, and they contain the data integrity for each satellite during different time periods, along with the data identifiers for missing data at specific times. When calculating the data integrity of the occultation, the theoretical total amount of data used is the theoretical total of valid data, meaning the data from epochs with a signal-to-noise ratio greater than or equal to the SNR threshold, rather than the theoretical total of data for all epochs in the file. Figure 5 presents the analysis results of data integrity for the example ROEX file.

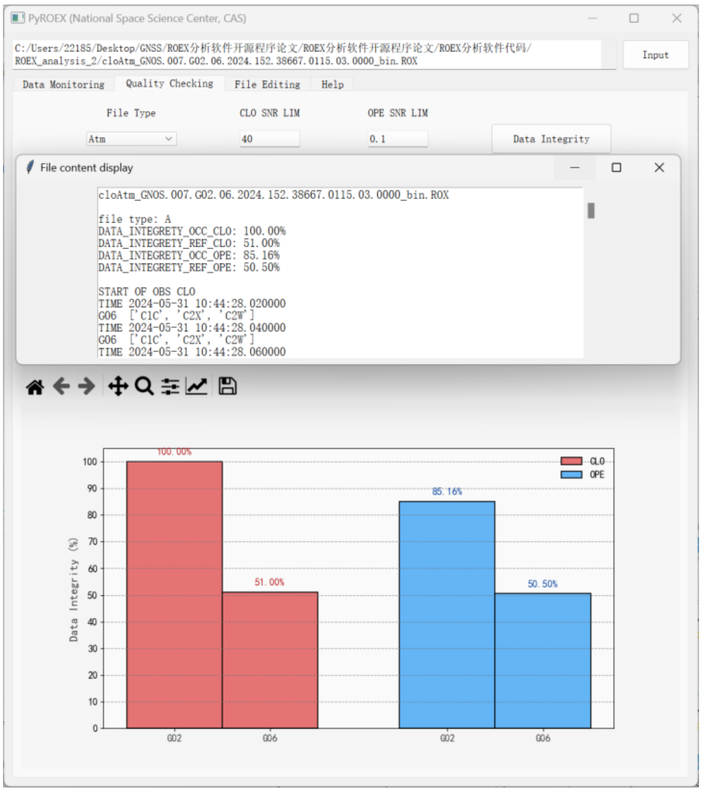

Fig.5. Results of data integrity

3.3 File editing

After using the data monitoring and quality detection modules, users can clearly assess the data quality status of each time period in the current file. If users need to extract data from time periods with good quality, they can use the file editing module to generate a new ROEX file within the user-defined time range. The file editing module interface is shown in Figure 5.

For atmospheric occultation observation files, the user needs to input five parameters: closed-loop observation start time, closed-loop observation end time, open-loop observation start time, open-loop observation end time, and output file name. For ionospheric occultation observation files, the user only needs to input the observation start time, observation end time, and output file name. After inputting the information, click the

"Execute" button. The input time format should be "YY-MM-DD HH:MM:SS", and it is important to ensure that the seconds include a decimal point. For example, "2024-2-21 18:42:56.00" is valid, while "2024-2-21 18:42:56" will cause an error.

The new ROEX file will be named with the inputted output file name followed by the ".ROX" extension, and it will be saved in the current execution folder. The newly generated ROEX file strictly adheres to the ROEX format.

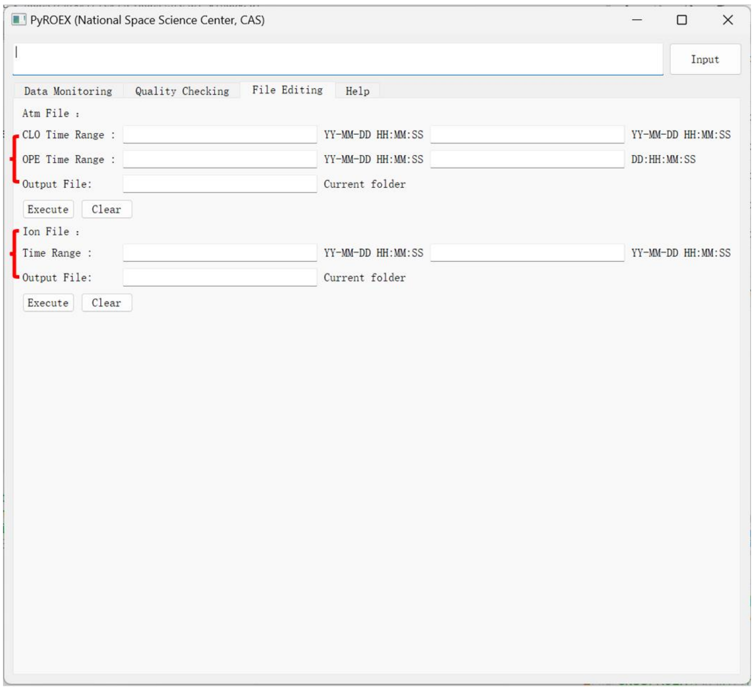

Fig.6. File editing interface

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:

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

- B_i: The theoretical total data amount for epoch i, which meets the signal-to-noise ratio (SNR) threshold. If the SNR of a specific frequency point for an epoch is below the threshold, the data for that frequency point is not included in the theoretical total data amount.
- A_i: The actual total data amount for epoch i, If the data at a specific point is 0, it is considered that no data is provided for that point.
- For atmospheric occultation files, the program computes four completeness indices：DI_OCC_CLO、DI_REF_CLO、DI_OCC_OPE和DI_REF_OPE， For ionospheric occultation files, only one completeness index DI is calculated

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.

4.3 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.