GMR-Water

GNSS **M**ultipath **R**eflectometry - **W**ater level

User Manual

Written by Fei Yang, Ruixian Hao

China university of mining and technology-Beijing

Apr 2025

CONTENTS

1 Introduction

With the continuous development of Global Navigation Satellite Systems (GNSS), contemporary receivers can capture multi-frequency signals from multiple GNSS systems, including GPS, Galileo, GLONASS, and BDS. This progress is undoubtedly conducive to the development of GNSS-MR technology. Moreover, traditional GNSS-MR water level retrieval techniques predominantly rely on Signal-to-Noise Ratio (SNR) observations. Nevertheless, it has been demonstrated that both carrier phase and pseudo range observations also possess the potential for water level retrieval. Incidentally, there have been certain advancements in the combination methods for multi-frequency results as well. In response to these developments, we have developed a software named GMR-Water, which is based on MATLAB, for water level retrieval. This software accomplishes water level retrieval by utilizing multi-frequency, multi-system, and multi-observation values, thereby significantly augmenting the retrieval accuracy and temporal resolution in comparison to traditional techniques. Additionally, the software incorporates two multi-frequency fusion algorithms, generating water level retrieval information at uniform sampling intervals, thus further promoting the development of GNSS-MR water level retrieval technology.

2 Mathematical methods

2.1 Water level retrieval models

The GMR-Water software can conduct water level retrieval based on three types of observations: Signal-to-Noise Ratio (SNR), carrier phase, and pseudo range, and it includes a total of five retrieval models. Among them, based on the SNR, two retrieval models are provided, namely the spectral analysis model and the inverse modeling model. For carrier and pseudo range observations, the dual/triple frequency combination retrieval model is provided, and at the same time, a single-frequency retrieval model is also provided.

2.1.1 SNR Spectral analysis model

The SNR value observed by GNSS is the result of the interference between the direct signal and the reflected signal, and it can usually be expressed as:

$$SNR = \sqrt{A_d^2 + A_r^2 + 2A_dA_r \cos \Delta\varphi} \tag{1}$$

where A_d and A_r are the amplitudes of the direct signal and the reflected signal respectively. $\Delta\varphi$ is the phase difference, and according to the geometric relationship, it can be expressed as:

$$\Delta\varphi = \frac{4\pi h}{\lambda} \sin\theta \tag{2}$$

where θ is the elevation angle, h is the reflected height and λ is the wavelength of single, respectively.

For the direct signal part, it can be removed by quadratic polynomial fitting. As for the reflected signal part, due to the reflection and antenna gain, the signal strength is relatively small and can be ignored. Then the SNR residual sequence δSNR can be expressed as:

$$\delta SNR = 2A_d A_r \cos \Delta \varphi \tag{3}$$

The angular frequency for the residual sequence can be got by calculating the differential of $\Delta\varphi$:

$$\omega = \frac{d\Delta\varphi}{d(\sin\theta)} = \frac{4\pi h}{\lambda} = 2\pi f \tag{4}$$

and according to formula (4), reflected height h can be expressed as:

$$h = \frac{\lambda f}{2} \tag{5}$$

Therefore, once the frequency of the SNR residual sequence f is determined, the reflection height h of the signal can be obtained. Furthermore, by combining it with the antenna height of GNSS, the height of the reflecting surface (water surface) can be inverted. Since the $\sin\theta$ is non - uniformly sampled, the LSP analysis method is usually used to obtain the frequency of the SNR residual sequence.

2.1.2 SNR Inverse modeling model

The power of the reflected signal undergoes attenuation due to the roughness of the reflecting surface. The attenuation coefficient can be modeled as:

$$E^2 = e^{-4\kappa^2 s^2 \sin^2(e)} \tag{6}$$

where $\kappa = \frac{2\pi}{\lambda}$ is the wave number, and s represents the standard deviation of the surface height irregularities, indicating the sea surface roughness caused by wind-driven waves. The attenuation coefficient decreases as the elevation angle increases, which corresponds to the signal amplitude variation in δSNR data with the elevation angle. Neglecting the effects of the Fresnel reflection coefficient X and the antenna gain G, the attenuation is expressed in terms of δSNR, which represents the amplitude modulation of δSNR decreasing with increasing elevation angle. The model is established as:

$$\delta SNR = A_e E \cos\left(\frac{4\pi H \sin e}{\lambda} + \varphi_0\right) \tag{7}$$

where A_e represents the amplitude independent of the elevation angle.

The SNR interference patterns contain information about sea surface height. Although some methods like arc-based inversion perform well, they don't take full advantage of the continuity of the parameters. B-spline functions are powerful tools that can smoothly fit changes in sea surface height, especially useful in GNSS-MR applications. The sea surface height can be modeled as:

$$h(t) = \sum_{j=0}^{N} h_j N_j^r(t) \tag{8}$$

where h_j are the coefficients (to be estimated) and $N_j^r(t)$ are B-spline basis functions of order r. In practice, quadratic or cubic B-splines are commonly used, as they ensure first- or second-order continuity, which is suitable for modeling sea surface height variations including tides.

During the estimation, certain parameters (like reflection amplitude and damping factor) are assumed to remain constant over the analysis period. The total number of parameters to estimate is:

$$M_T = M_B + 2 \cdot M_f + 1 \tag{9}$$

where M_B is the number of B-spline nodes (coefficients), M_f is the number of GNSS frequencies and M_T is total number of parameters. Since the number of observations is much larger than the number of parameters, this forms an overdetermined problem. The solution involves nonlinear least squares optimization, which finds the best-fit parameters by minimizing the squared residuals:

$$\min \sum_N \left(y_i - f\left(x_0, x_1, \dots, x_{M_T}\right)\right)^2 \tag{10}$$

where N is the total number of observations and y_i are SNR measurements.

For details on the inverse modeling, please refer to **Strandberg et al. (2016)**: *Strandberg, J., Hobiger, T., Haas, R. (2016). Improving GNSS-R sea level determination through inverse modeling of SNR data. **Radio Science***, 51, 1286–1296. https://doi.org/10.1002/2016RS006057.

2.1.3 Carrier dual/triple frequency combination model

The carrier can be expressed as:

$$L_i = \rho - I(f_i) + T + M_{Li} - \lambda_i N_i \tag{11}$$

where, L_i represents the carrier phase observation for frequency band i, ρ is the geometric distance between the satellite and the receiver, $I(f_i)$ denotes the ionospheric delay for frequency band i, T is the tropospheric delay, M_{Li} represents the carrier phase multipath error for frequency band i, and N_i is the integer ambiguity.

For reflectometry, the reflection height information is embedded in the multipath error; therefore, it is necessary to extract the multipath component from the carrier phase observations. The dual-frequency carrier phase combination can be expressed as:

$$M_{L1,2} = L_1 - L_2 = I(f_2) - I(f_1) + M_{L1} - M_{L2} + \lambda_2 N_2 - \lambda_1 N_1 \tag{12}$$

and the triple-frequency carrier phase combination can be expressed as:

$$M_{L1,2,3} = (\lambda_3^2(L_1 - L_2) + \lambda_1^2(L_2 - L_3) + \lambda_2^2(L_3 - L_1))/1m^2 \tag{13}$$

Both dual-frequency and triple-frequency carrier phase combinations can separate the multipath information. Similar to SNR-based methods, the multipath information can undergo Lomb-Scargle Periodogram (LSP) analysis to determine the dominant frequency f. The dominant frequency exhibits a linear relationship with the reflection height h, expressed as:

$$h = a \times f + b \tag{14}$$

where a and b are coefficients typically obtained through simulation or fitting.

2.1.4 Pseudo range dual/triple frequency combination model

The pseudo range can be expressed as:

$$C_i = \rho + I(f_i) + T + M_{Ci} \tag{15}$$

where P_i represent pseudo range for frequency band i. Similar to carrier phase, pseudo range also allow the extraction of multipath information through dual-frequency and triple-frequency combinations:

$$M_{C1,2} = C_1 - C_2 = I(f_2) - I(f_1) + M_{C1} - M_{C2} \tag{16}$$

$$M_{C1,2,3} = (\lambda_3^2(C_1 - C_2) + \lambda_1^2(C_2 - C_3) + \lambda_2^2(C_3 - C_1))/1m^2 \tag{17}$$

Similarly, the dominant frequency of the signal can be obtained through LSP spectral analysis, which is then used to determine the reflection height.

2.1.5 Carrier & pseudo range combination model

Additionally, for single-frequency scenarios, multipath information can also be extracted through the combination of carrier phase and pseudo range observations:

$$M_i = C_i - L_i = 2I(f_i) + M_{Ci} - M_{Li} + \lambda_i N_i \tag{18}$$

where $I(f_i)$ represents the ionospheric delay, which can be removed using high-order polynomial filtering. Similar to other multi-frequency combinations, the dominant frequency of the signal is obtained through LSP spectral analysis, which is then used to derive the reflection height.

2.2 Error estimation

2.2.1 Tropospheric correction

Santamaría-Gómez and Watson (2016) used the experiential astronomical refraction model based on atmospheric pressure and temperature to remove this effect in GPS-MR sea level estimations. The elevation angle was corrected as follows:

$$\Delta e = \frac{510}{\frac{9}{5}T+492}\frac{P}{1010.16}cot\left(e_T + \frac{7.31}{e_T+4.4}\right) \tag{19}$$

where T is the temperature in °C, P the pressure in hPa at the antenna, and e_T is the true elevation angle.

Williams and Nievinski (2017) used the Global Temperature and Pressure (GPT2w) model together with the VMF1 mapping function to remove this effect. The tropospheric delay τ_T is calculated by

$$\tau_T = 2\Delta\tau_h^z \times m_h(e) + 2\Delta\tau_w^z \times m_w(e) \tag{20}$$

where $\Delta\tau_h^z = \tau_h^z(\text{-}h) - \tau_h^z(0)$ is the zenith delay difference across antenna and surface positions, and m is the mapping function, which is indicated separately for the hydrostatic and wet components.

2.2.2 Dynamic tide level correction

The traditional GNSS-MR water level retrieval technique operates under the assumption that the reflecting surface remains stationary. However, because the sea surface is constantly changing, this static assumption becomes unreasonable and inaccurate. To address this issue, this study introduces a dynamic correction algorithm based on traditional SNR retrieval methods. By considering the dynamic nature of sea level variations, the satellite elevation angle rate $\dot{\theta} = \frac{d\theta}{dt}$ and the reflection path change rate $\dot{h} = \frac{dh}{dt}$ are introduced. Consequently, the corrected dynamic vertical reflection distance h can be expressed as:

$$h = \bar{h} - \frac{\tan\theta}{\dot{\theta}}\dot{h} \tag{21}$$

where $\bar{h}$ represents the vertical reflection distance under the static assumption, and h represents the dynamic vertical reflection distance. To further refine this, the $\bar{h}$ sequence is used to fit a tidal wave curve based on harmonic analysis. The curve fitting equation is:

$$\bar{h} = \sum_{i=1}^{N} C_i f_i \sin(\omega_i t + v_i) + S_i f_i \cos(\omega_i t + v_i) \tag{22}$$

where C_i and S_i are the sine and cosine coefficients, ω_i is the tidal frequency, v_i is the tidal phase, and f_i adjusts the amplitude. Once the parameters are obtained, the derivative of the fitted curve yields the reflection path change rate can be expressed as:

$$\dot{h} \approx -\sum_{i=1}^{N} \omega_i C_i f_i \sin(\omega_i t + v_i) + \omega_i S_i f_i \cos(\omega_i t + v_i) \tag{23}$$

Finally, the dynamic reflection distance h is corrected by subtracting $\frac{\tan\theta}{\dot{\theta}}\dot{h}$ from $\bar{h}$.

2.3 Combination algorithm

2.3.1 Robust regression strategy

For jth signal of track l in the ith time window, the state transition equation can be expressed as:

$$\bar{h}_{i,j,l}(t_{i,j,l}) - \Delta h_{T_{i,j,l}}(t_{i,j,l}) = \frac{\tan(e_{i,j,l})}{e_{i,j,l}}\dot{h}_i(t_i) + \dot{h}_i(t_i) \times (t_{i,j,l} - t_i) + h_i(t_i) \tag{24}$$

If this equation is established for each signal of the tracks in the set $L_i = \{\dots, l-1, l, l+1, \dots : |t_s - t_i| < T/2\}$, a system of equations of the ith window can be obtained as follows:

$$\begin{pmatrix} \bar{h}_{i,j,l-1}(t_{i,j,l-1}) - \Delta h_{T_{i,j,l-1}}(t_{i,j,l-1}) & = [\frac{\tan(e_{i,j,l-1})}{e_{i,j,l-1}} + (t_{i,j,l-1} - t_i)]\dot{h}_i(t_i) + h_i(t_i) \\ \bar{h}_{i,j,l}(t_{i,j,l}) - \Delta h_{T_{i,j,l}}(t_{i,j,l}) & = [\frac{\tan(e_{i,j,l})}{e_{i,j,l}} + (t_{i,j,l} - t_i)]\dot{h}_i(t_i) + h_i(t_i) \\ \bar{h}_{i,j,l+1}(t_{i,j,l+1}) - \Delta h_{T_{i,j,l+1}}(t_{i,j,l+1}) & = [\frac{\tan(e_{i,j,l+1})}{e_{i,j,l+1}} + (t_{i,j,l+1} - t_i)]\dot{h}_i(t_i) + h_i(t_i) \end{pmatrix} \tag{25}$$

or equivalently in terms of a matrix:

$$\overline{H}_i = M_i \dot{h}_i + h_i = A_i X_i \tag{26}$$

with $\overline{H}_i = \begin{pmatrix} \vdots \\ \bar{h}_{i,j,l}(t_{i,j,l}) - \Delta h_{T_{i,j,l}}(t_{i,j,l}) \\ \vdots \end{pmatrix}$, $M_i = \begin{bmatrix} \left[\frac{\tan(e_{i,j,l})}{e_{i,j,l}} + (t_{i,j,l} - t_i)\right] \\ \vdots \end{bmatrix}$, and $A_i = (M_i 1)$, $X = \begin{pmatrix} \dot{h}_i \\ h_i \end{pmatrix}$.

Finally, it can be solved using the least squares method:

$$X_i = (A_i^T P_i A_i)^{-1}(A_i^T P_i)\bar{H}_i \tag{27}$$

The standard least squares method assumes normally distributed errors in the data. When this assumption is violated—such as due to asymmetric errors or outliers—model estimates become unreliable. To address this, the robust regression method is introduced, offering more stable estimates by minimizing the impact of outliers.

Robust regression uses iteratively reweighted least squares, where:

1. Initial weights are equal, and parameters are estimated via standard least squares:

$$\hat{X}_i^{(1)} = (A_i^T A_i)^{-1} A_i^T H_i \tag{28}$$

2. Residuals are computed:

$$v_i^{(1)} = A_i \hat{X}_i^{(1)} - H_i \tag{29}$$

3. In each iteration, weights are updated using the bisquare function:

$$P_i^{(k+1)} = \begin{cases} (1 - (v_i^{(k)})^2)^2, & |v_i^{(k)}| < 1 \\ 0, & |v_i^{(k)}| \geq 1 \end{cases} \tag{30}$$

4. Updated residuals and parameters are recalculated:

$$v_i^{(k)} = A_i \hat{X}_i^{(k)} - H_i \tag{31}$$

$$\hat{X}_i^{(k+1)} = (A_i^T P_i^{(k+1)} A_i)^{-1} A_i^T P_i^{(k+1)} H_i \tag{32}$$

The process continues until the solution converges within a small threshold ε:

$$|\hat{X}_i^{(k+1)} - \hat{X}_i^{(k)}| < \varepsilon \tag{33}$$

For details, please refer to **Wang et al. (2019)**: *Wang, X. L., He, X. F., Zhang, Q. (2019). Evaluation and combination of quad-constellation multi-GNSS multipath reflectometry applied to sea level retrieval.* ***Remote Sensing of Environment****, 231(2): 111229.* https://doi.org/10.1016/j.rse.2019.111229.

2.3.2 B-spline curve estimation

Similar to the inverse modeling approach, the characteristics of B-spline curves are utilized to fit the multi-frequency inversion results, enabling the combination of results from multiple frequencies. The sea surface height can be modeled as:

$$h(t) = \sum_{j=0}^{N} h_j N_j^r(t) \tag{34}$$

where h_j are the coefficients (to be estimated) and $N_j^r(t)$ are B-spline basis functions of order r. In practice, quadratic or cubic B-splines are commonly used, as they ensure first- or second-order continuity, which is suitable for modeling sea surface height variations including tides.

During the estimation, certain parameters (like reflection amplitude and damping factor) are assumed to remain constant over the analysis period. The total number of parameters to estimate is:

$$M_T = M_B + 2 \cdot M_f + 1 \tag{35}$$

where M_B is the number of B-spline nodes (coefficients), M_f is the number of GNSS frequencies and M_T is total number of parameters. Since the number of observations is much larger than the number of parameters, this forms an overdetermined problem. The solution involves nonlinear least squares optimization, which finds the best-fit parameters by minimizing the squared residuals:

$$\min \sum_N \left(y_i - f\left(x_0, x_1, \dots, x_{M_T}\right) \right)^2 \tag{36}$$

where *N* is the total number of observations and y_i are SNR measurements.

3 Operation

3.1 Dependencies

GMR-Water requires **MATLAB 2022a or later versions** to operate.

3.2 Starting

The GMR-Water software offers two initialization modes: it can be executed either through an interactive graphical user interface (GUI) or via function calls in batch processing mode using scripts. The main structure of GMR-Water is shown in Figure 3-1:

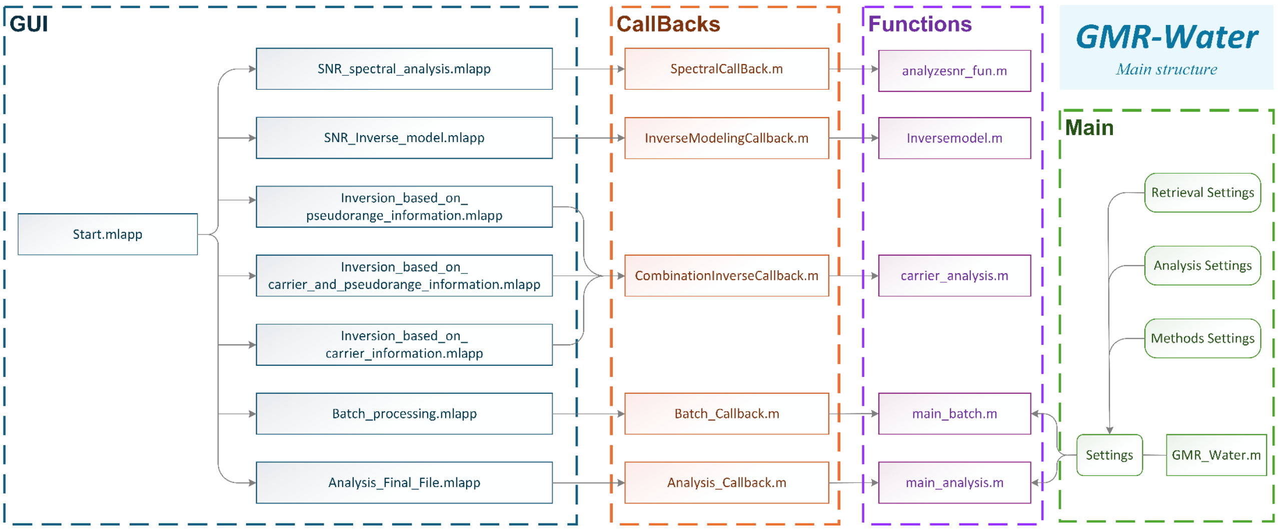

Figure 3-1. Main structure of GMR-Water

<u>Run in GUI mode</u>

The main interface of the software is "**Start.mlapp**", as shown in Figure 3-2. The left side of the software features the functional area, which includes modules for five water level retrieval methods using multi-observation (SNR, carrier phase and pseudo range), along with the "*batch processing*" and the "*analysis final file*". On the right side of the start interface, there is an introduction and schematic diagram of GNSS-MR technology.

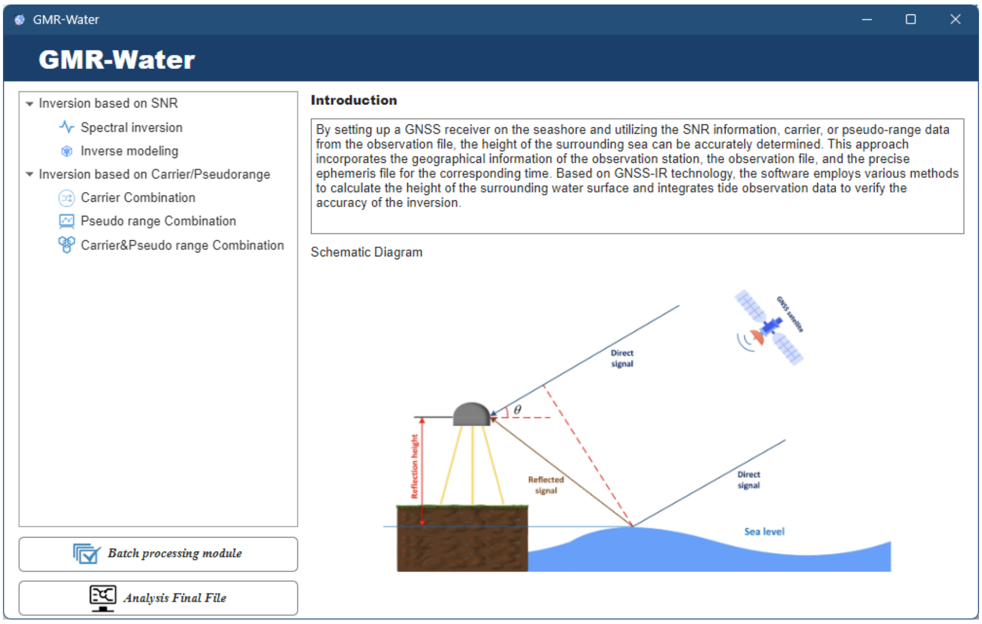

Figure 3-2. The start interface of GMR-Water

<u>Run as a function</u>

The main function of GMR-Water is "**GMR_Water.m**", which can be executed by providing an input configuration structure "Settings" or by simply typing "GMR_Water" in the command line to run a demo case. The GMR_Water function integrates water level retrieval, result analysis, customizable parameter configuration, and parallel processing capabilities.

4 Graphical User Interface (GUI) Operation Manual

4.1 Starting interfaces

To select an inversion method, right-click and choose "**Turn to this**" to access the corresponding module. Then the right side of the boot interface will change to show the introduction and the process flow of the corresponding method, as illustrated in Figure 4-1. Click the "**Run this module**" button in the lower right corner to start the current module. You can also click on "**Batch Processing Module**" and "**Analysis Final File**" to access their respective functional interfaces.

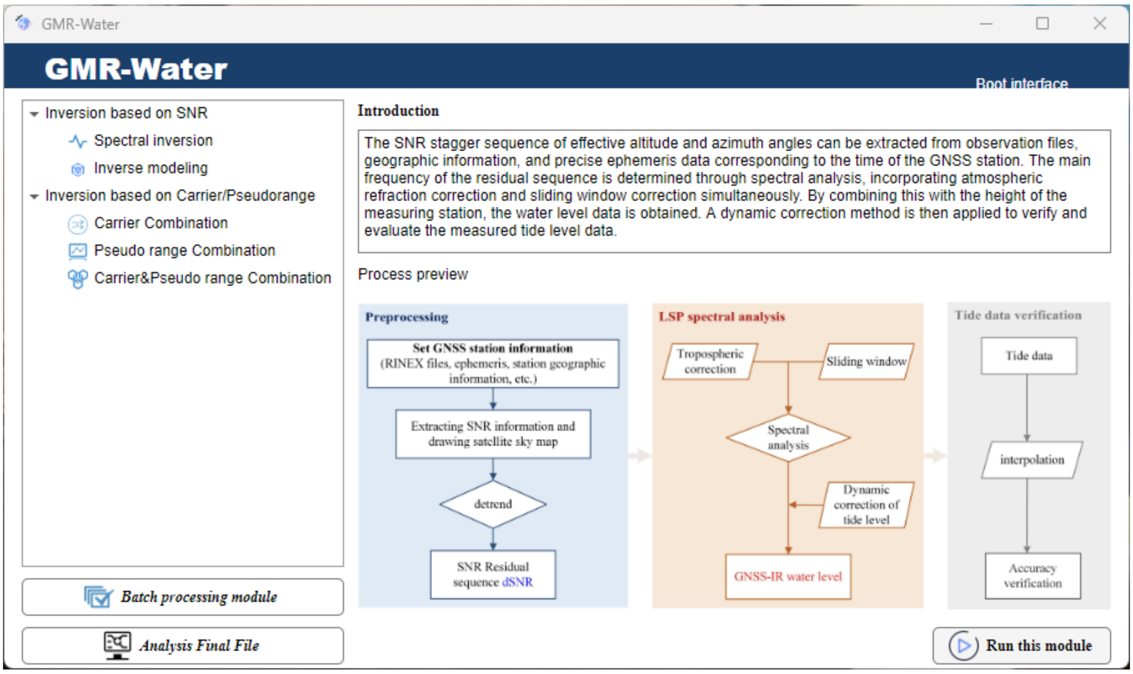

Figure 4-1. Interface when turn to one method

4.2 Operation settings

Regardless of which inversion module is selected, it is necessary to complete the settings including information about the GNSS station, the GNSS observation file and the precision ephemeris file. The station information setting interface is shown in Figure 4-2.

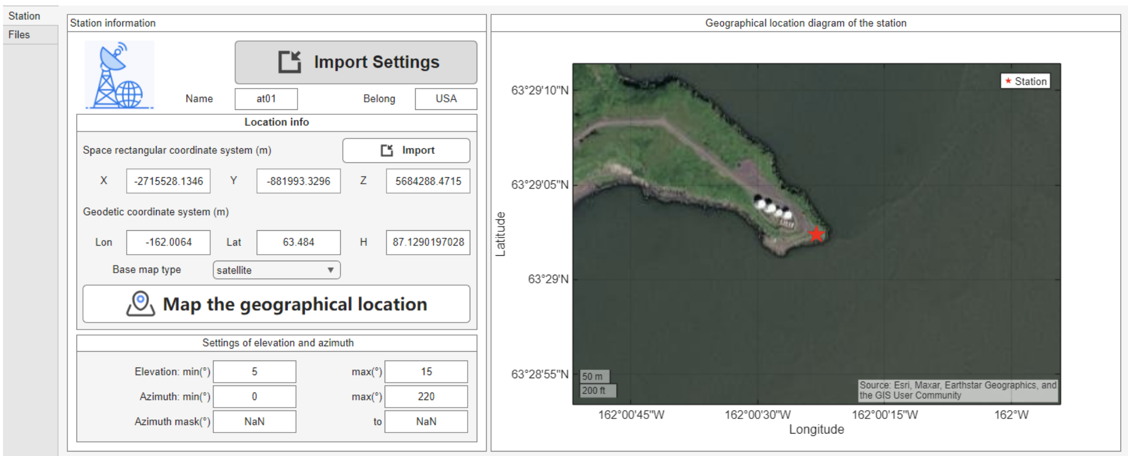

Figure 4-2. station information setting interface

Import Settings: Import the configuration file, including all the information required for retrieval: station related information, RIENX file and precision ephemeris file path, inversion time, etc. This file is often saved from the last retrieval. The example is shown in Figure 4-3.

Name	Value	Size	Class
azi	[0,220]	1×2	double
azimask	[NaN,NaN]	1×2	double
dt	'15'	1×2	char
elv	[5,15]	1×2	double
obs_file	*2×1 cell*	2×1	cell
obs_path	'E:\GMR-water\data\at01\o...	1×27	char
rinex_version	'rinex 3'	1×7	char
sp3_file	*2×1 cell*	2×1	cell
sp3_path	'E:\GMR-water\data\sp3\C...	1×26	char
sp3_type	'COD'	1×3	char
station_belong	'USA'	1×3	char
station_l	[-162.0064,63.4840]	1×2	double
station_name	'at01'	1×4	char
station_xyz	[-2.7155e+06,-8.8199e+05,...	1×3	double
time	*1×2 datetime*	1×2	datetime

Figure 4-3. configuration file : "station_setting.mat"

Import: Import a file containing station related information, often including station name, geographic location information, etc. This is a standard format file; the required information is shown in Figure 4-4.

Name	Value	Size	Class
azi_lim	[0,220]	1×2	double
azi_mask	[NaN,NaN]	1×2	double
elv_lim	[5,15]	1×2	double
sta_asl	12	1×1	double
sta_lat	63.4840	1×1	double
sta_lon	-162.0064	1×1	double
station_belong	'USA'	1×3	char
station_name	'AT01'	1×4	char
staxyz	[-2.7155e+06,-8.8199e+05,...	1×3	double
tide_range	[-2,2.5000]	1×2	double

Figure 4-4. at01_info.mat file

Map the geographical location: According to the coordinate information of the station, the geographical position diagram of the station is drawn (on the right side of the interface), and sometimes the azimuth range can be roughly estimated according to the geographical map.

Observation files and the Ephemeris files settings interface are shown in Figure 4-5.

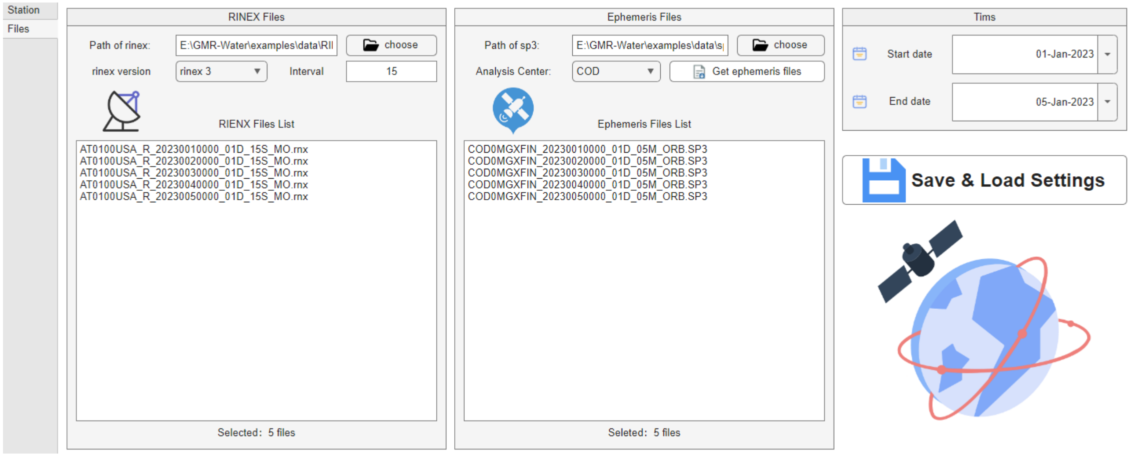

Figure 4-5. Files settings interface

The software supports selecting observation files for multiple consecutive days at the same station. Please ensure that the observation files are **stored and named according to the standard RINEX format**. For optimal performance, we recommend using RINEX 3.0 or higher. The software remains backward compatible with RINEX 2.11 for legacy data processing. The software automatically adjusts the inversion time based on the selected RINEX file. For precise ephemeris files, you need to select the file corresponding to the inversion time for the target period. The software also offers a "**Get Ephemeris Files**" function, which automatically downloads the necessary ephemeris files based on the current time and the selected Analysis Center.

4.3 Observation extraction

In this part, the main task is to extract the observation information (SNR, carrier phase and pseudo range) in the RINEX file, to calculate the elevation and azimuth angle at each epoch by using the precision ephemeris, and to screen the observation information by using the station information (limited elevation and azimuth range).

<u>SNR extraction</u>

For spectral analysis and inverse modeling methods, the first step is to extract the SNR data. The software interface for SNR extraction and analysis is displayed in Figure 4-6 and Figure 4-7.

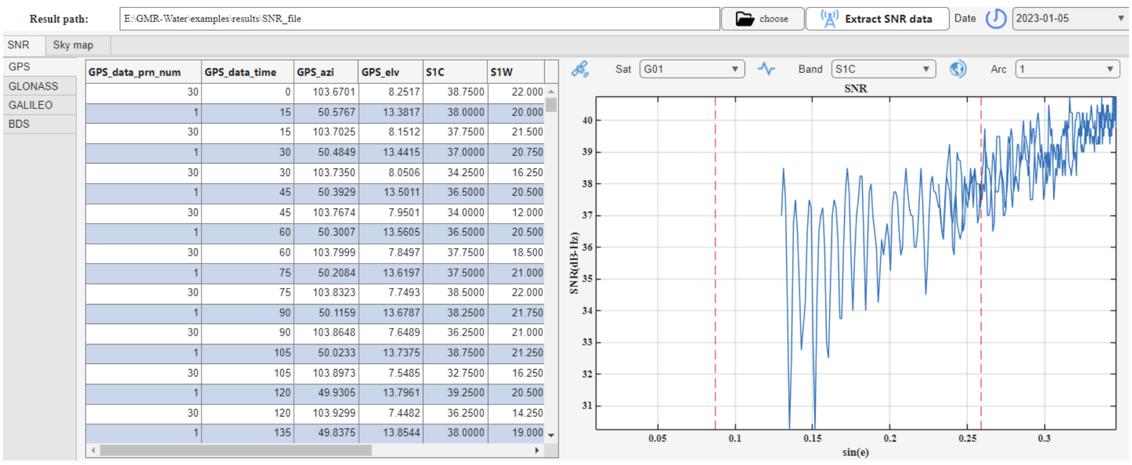

Figure 4-6. Interface of SNR information extraction viewing

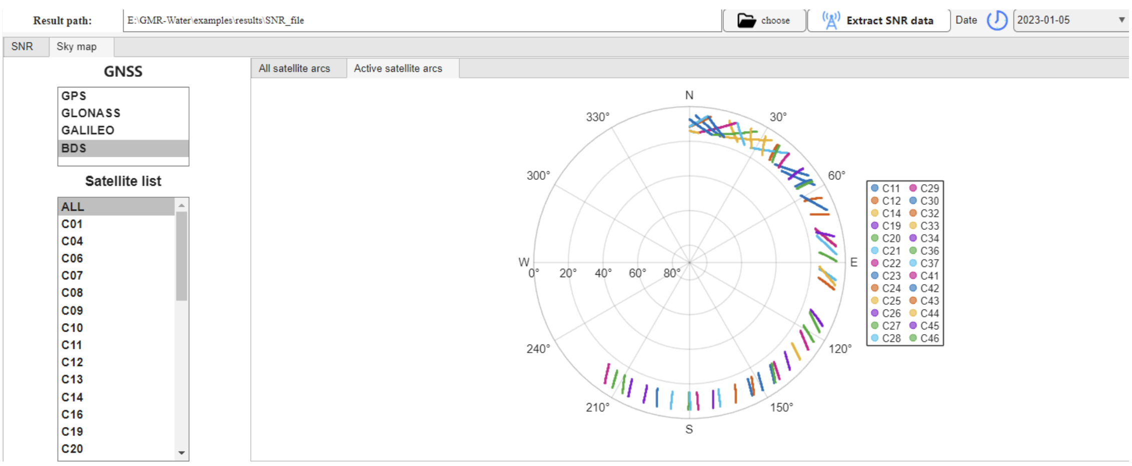

Figure 4-7. Interface of Sky map viewing

In Figure 4-6 and Figure 4-7, this section allows you to view the SNR series with elevation angle and the sky map. On the left side of the interface, the SNR viewing section displays the extracted results for each system, and related .mat files can be accessed in the Result path. The right side shows the SNR series, where the red dashed line indicates the range of limited elevation angle. At the top of the figure, you can switch between satellite PRN, frequency bands, and arc segments. On the far left of the interface, you can switch between different GNSS constellations.

For the sky map, you can specify a specific satellite system and display sky map of all satellites or a specific satellite. Note that by selecting “All satellite arcs” and “Active satellite arcs”, the sky map can switch to display all the arcs of corresponding satellite and the arcs only used for subsequent inversion. In addition, the top-right corner of the interface allows you to select the date to display.

<u>Carrier/Pseudo range extraction</u>

The extraction of carrier phase and pseudo range has been integrated into the inversion module, as shown in Figure 4-8. Simply select the target folder and click "**Extraction**" to begin extracting the relevant information. Please note that when extracting carrier phase or pseudo range data separately, you only need to select the target folder for storage. However, when extracting both simultaneously, the software will create two separate folders within the selected directory to store the carrier phase and pseudo range independently.

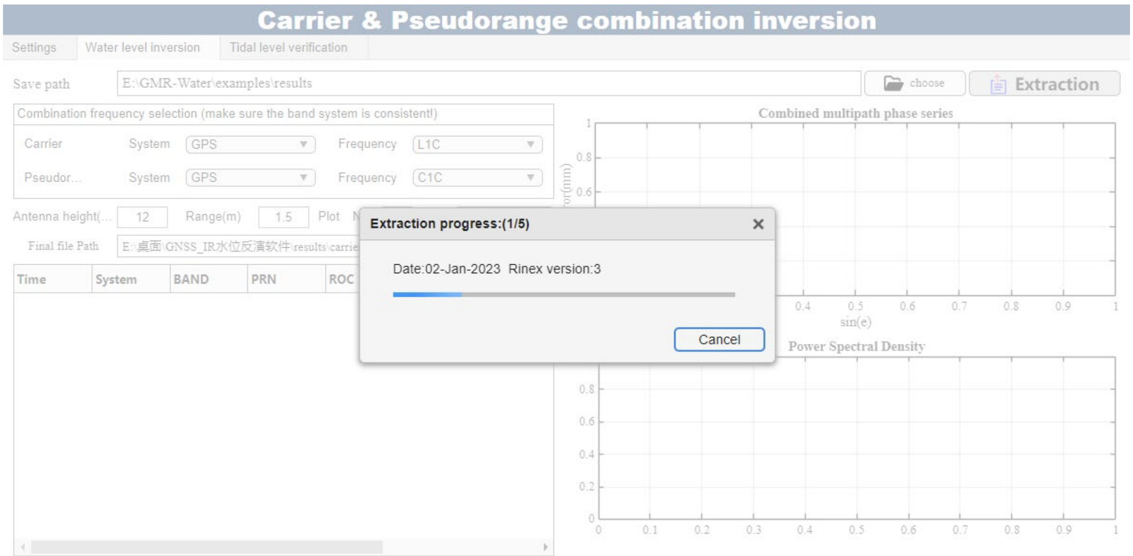

Figure 4-8. Interface of carrier and pseudo range extraction

4.4 Water level retrieval methods

<u>SNR Spectral analysis method</u>

The process of water level retrieval using the SNR spectral analysis method consists of three main parts: preprocessing, LSP spectral analysis, and tide data validation. The retrieval process is illustrated in Figure 4-9. The preprocessing step yields a detrended SNR residual sequence, which is analyzed using LSP spectral analysis to retrieve water levels. This process accounts for tropospheric delays and dynamic sea surface height variations, and also incorporates sliding window retrieval (WinLSP). Finally, tide data is used as a reference for accuracy validation.

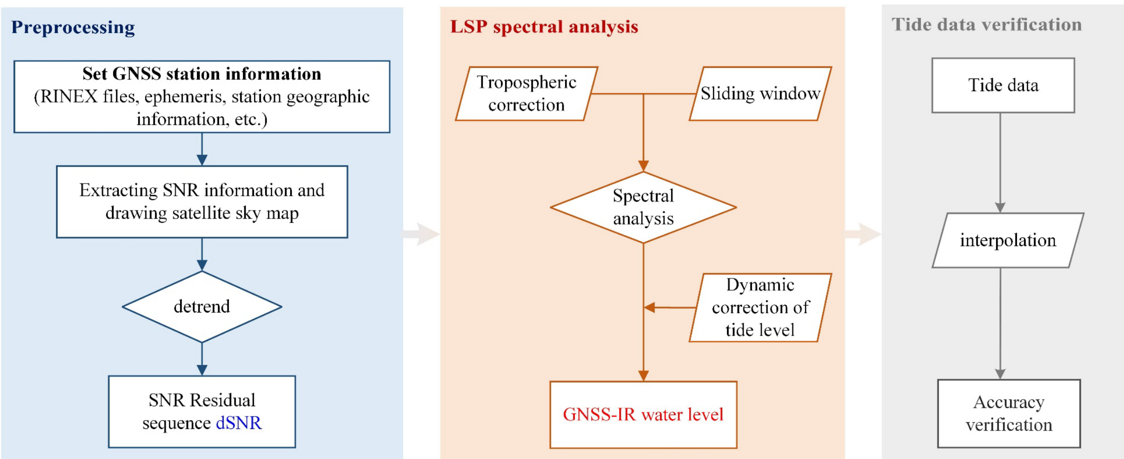

Figure 4-9. Retrieval process of SNR Spectral analysis method

In GMR-Water, the spectral analysis method is divided into three parts: water level inversion, tidal verification, and tidal correction. The first part is water level inversion, as shown in Figure 4-10. Before performing the inversion, you need to set the SNR data path (from the SNR extraction results) and the path to save the inversion results (note that this is not the final output file). Then, configure the basic inversion settings in the upper left corner of the interface. Choose whether to use the tropospheric correction model, set the SNR sampling interval (usually matching the RINEX file), the average distance from the station to the water surface, and the tidal monitoring range. You can also choose whether to enable real-time plotting (disabling this will speed up the inversion process) and whether to output the reflection area file (the first Fresnel reflection zone, in .kml format). Additionally, you can enable the sliding window function for this inversion, which could increase the number of inversions when the available arc segments are limited.

After completing the settings, click the "**Inverse**" button to start the inversion process. A progress bar and status indicator will be displayed in the lower left corner. A green status light indicates the inversion is in progress. During the inversion, the upper right corner will display δSNR arcs and the power spectral density graph from the spectral analysis. After each frequency is processed, its inversion result will be shown in the graph in the lower right corner.

Once the inversion is complete, the results will be displayed on the table on the left side, with detailed descriptions available in the table header. Additionally, we provide the "**View Reflection Area**" function, which allows users to easily view the generated .kml files. For detailed instructions, please refer to Section 6.2.

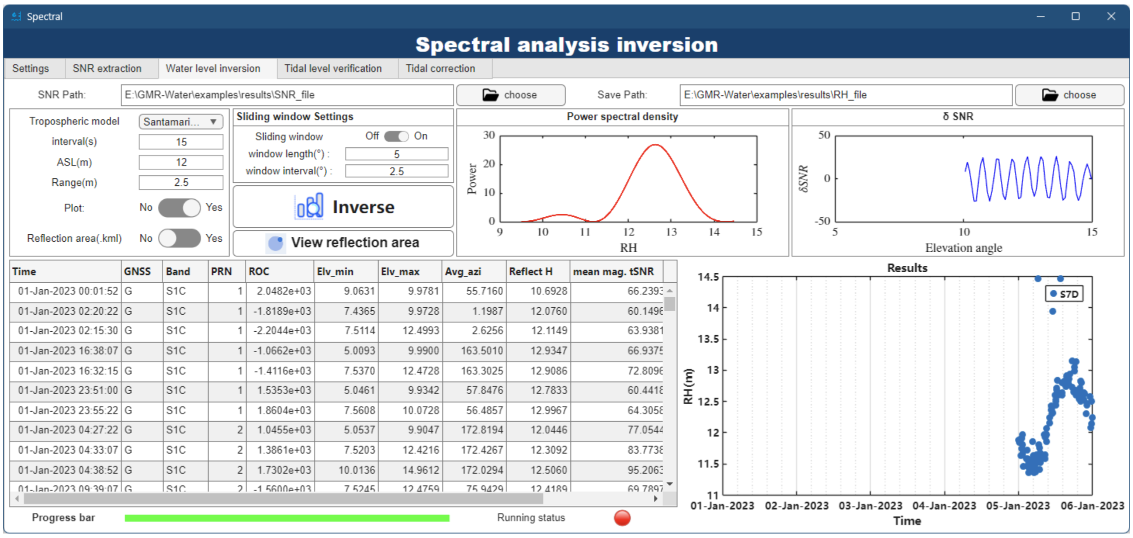

Figure 4-10. Interface of water level inversion part

In the interface of "Tidal level verification", it validates the accuracy of the inversion results with the tidal data and gives a comparison with different GNSS systems and different SNR data. First, you need to select a standard tidal level file (for specific details, see Section 6.1). Click the "Plot" button, and the results for all frequency along with the tidal results will be displayed on the coordinate axes below. A new window will also pop up to display the results for each system, as well as the RMSE and daily average number of inversion points for each frequency, as illustrated in Figure 4-11.

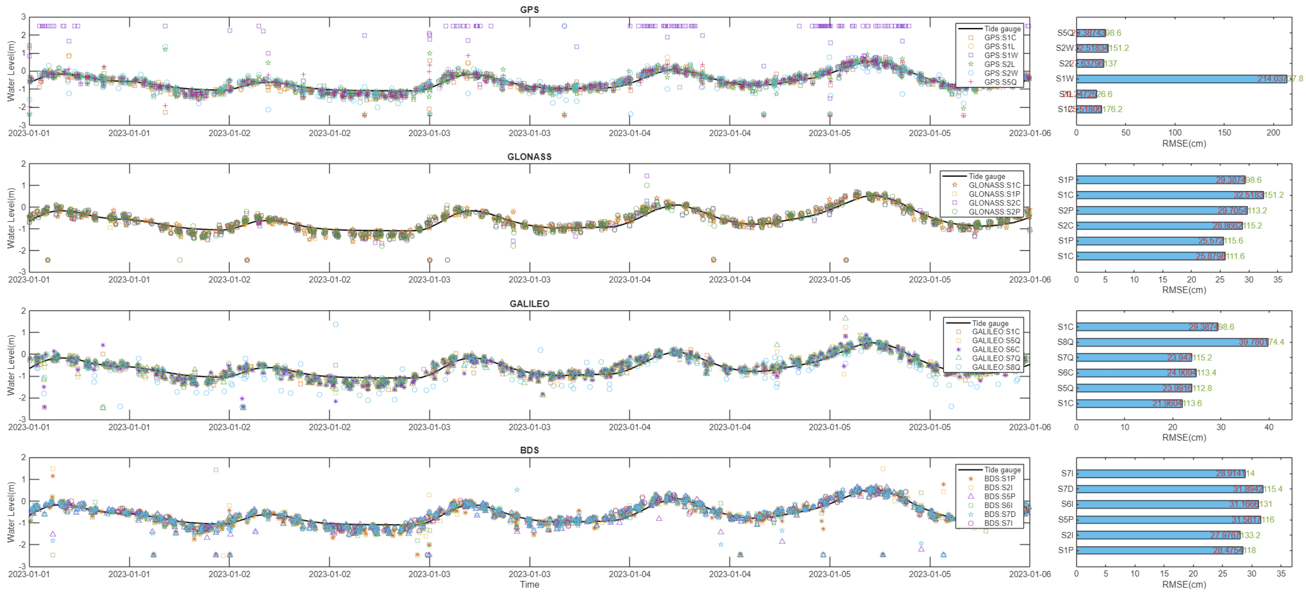

Figure 4-11. Results for each GNSS constellation and frequency point, including RMSE and daily average inversion points

The third part is the tidal correction, which mainly applies tidal correction to the inverted water levels and outputs the final file, as shown in Figure 4-12. Additionally, various visualization results are provided, including comparison charts of water levels before and after tidal correction, comparison charts of results across different frequency, correlation scatter plots, residual histograms for each frequency point, and Fresnel reflection zones.

You can configure the tidal components for correction, which is generally recommended to use O1, K1, N2, M2, and S2 tidal components. For longer time series (typically over a year), all 145 tidal components are suggested to be used for the fitting and correction. Once the settings are completed, click the "**Correct & Plot**" button, and the software will output the corrected final file to the specified path and generate the selected plots. The RMSE before and after correction, as well as the daily average number of inversions, will also be displayed.

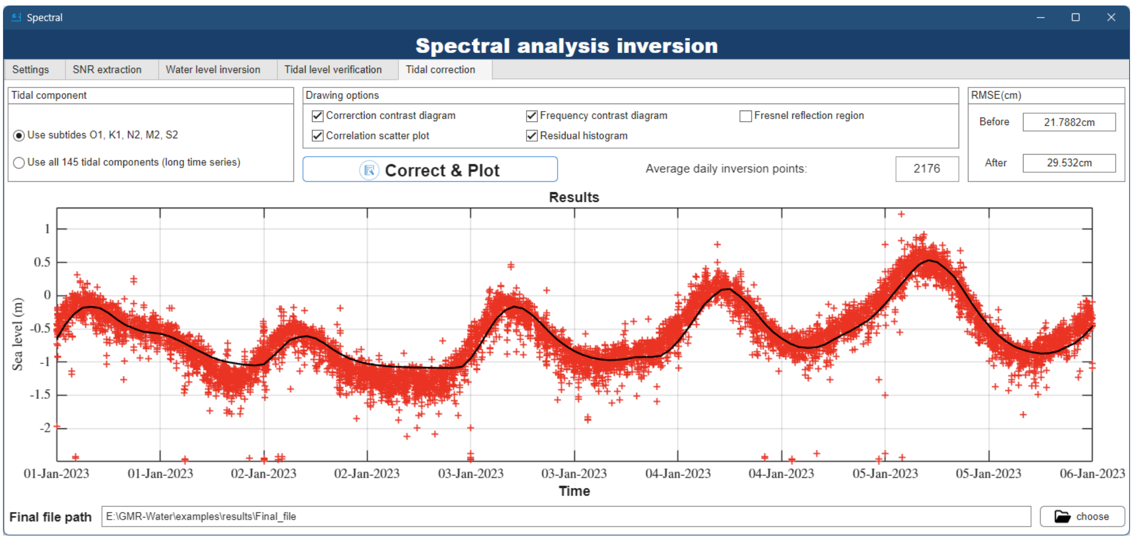

Figure 4-12. Interface of Tidal correction

SNR Inverse modeling method

The SNR inverse modeling method for water level retrieval involves three main steps: preprocessing, inverse modeling, and tide data validation, as illustrated in Figure 4-13. Preprocessing produces a detrended SNR residual sequence. Using the inverse modeling method, least-squares fitting is applied to derive the B-spline control points for the retrieved tidal levels, which are then used to calculate the water levels. Finally, tide data is employed as a reference for accuracy validation.

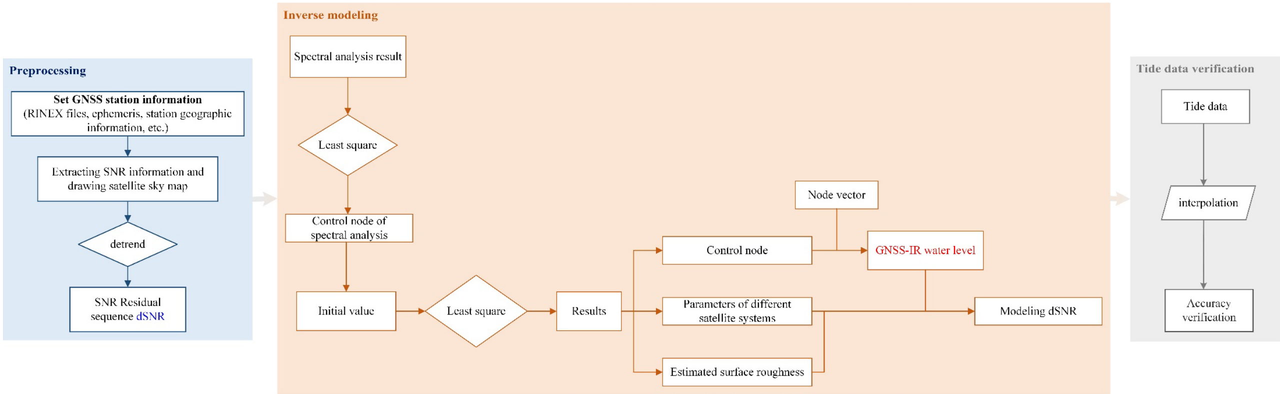

Figure 4-13. Retrieval process of SNR Inverse modeling method

Figure 16 presents the interface for inverse modeling. First, you need to select the path for the SNR data (obtained from the previous extraction step) and specify the save path for the results of the inverse modeling. In the modeling options, you can specify the estimated surface roughness (typically set the initial value to 0.01). The modeling employs a sliding window strategy, allowing you to choose the window length and interval (usually set to 18 hours and 3 hours, respectively). Additionally, you must provide the average height of the antenna above the water level and the water level detection range.

When "**Plot**" is on, a comparison graph of modeled values and true values will be plotted in the coordinate system below after completing modeling for one window. After finishing the above settings, click the "**Modeling**" button to start the modeling process. The status indicator will turn green until the inversion is completed.

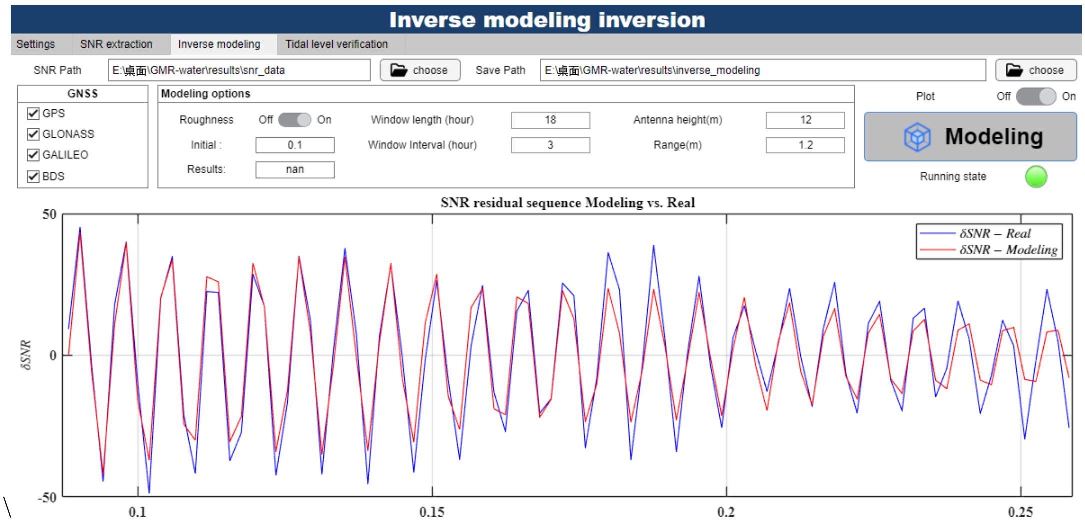

Figure 4-14. Interface of Inverse modeling

The second part is the tidal verification module, which aims to validate the accuracy of the inverse modeling, compare them with the results of the spectral analysis, and output the final files. First, you need to select a standard format tidal level file (see Section 6.1) and specify the path for the final output file. By clicking the "**Plot**" button, the results will be plotted according to the selected options, and the accuracy information will be provided. The final file will be generated in the specified path.

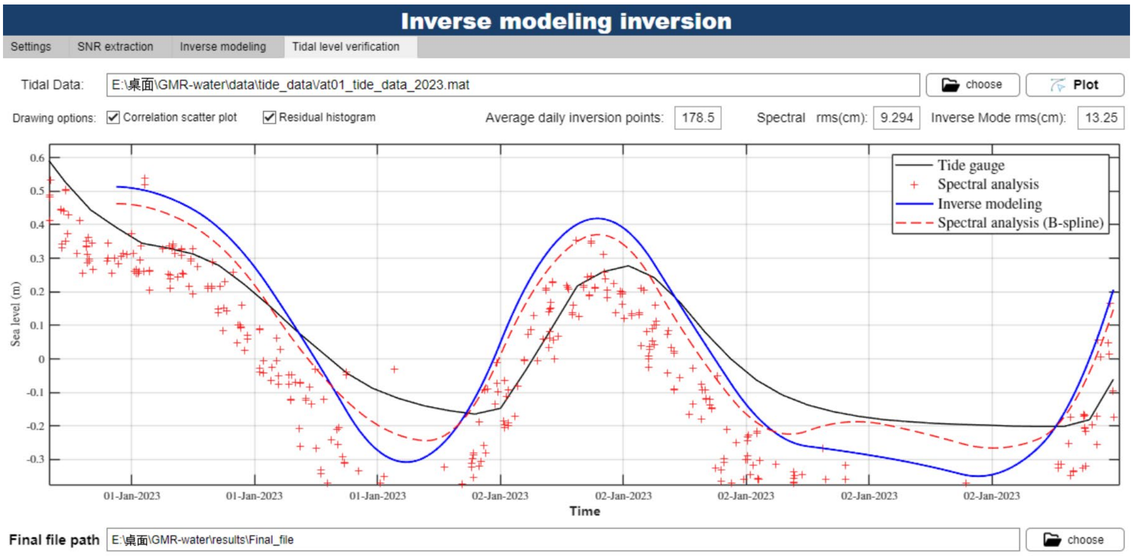

Figure 4-15. Result of inverse modeling

<u>Carrier dual/triple frequency combination method</u>

The carrier dual /triple-frequency combination method for water level retrieval consists of three main steps: preprocessing, combination inversion, and tide data validation, as shown in Figure 4-16. In the preprocessing step, combination observations (multipath components) are derived using dual-frequency or triple-frequency carrier combinations. These combined observations are then analyzed through spectral analysis to retrieve water levels. Finally, tide gauge data is used as a reference for accuracy validation.

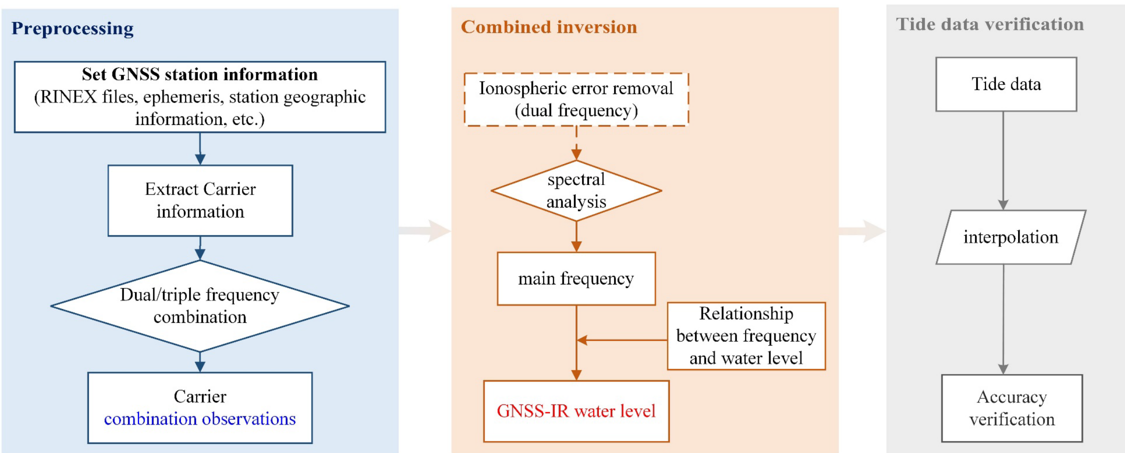

Figure 4-16. Retrieval process of Carrier dual/triple frequency combination method

The first step in carrier-based inversion is to extract carrier information from the RINEX files by selecting the carrier path for extraction. Choose the combination type for either dual-frequency or triple-frequency combinations, ensuring that the selected frequency belong to the same GNSS constellation. Next, set the antenna height and range, along with the path for saving the final file. You can also provide plotting options (disabling this may enhance inversion speed). Click "**Inverse**" to start the inversion process. Multipath signals and power spectral density will be displayed on the right. Upon completion, the results will be displayed in the table at the lower left corner of the interface.

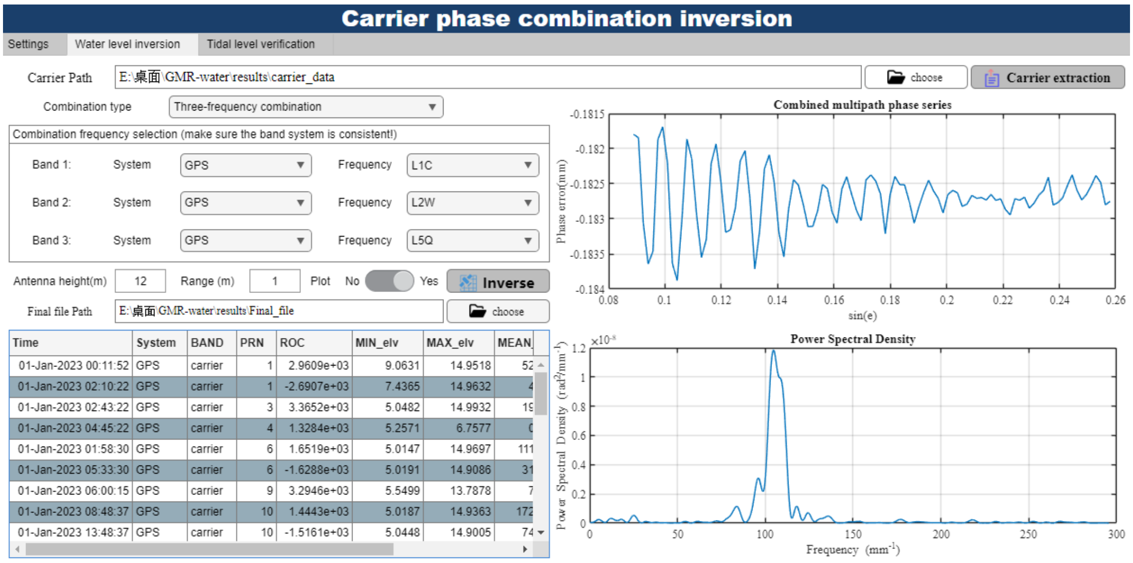

Figure 4-17. Interface of carrier combination inversion

To enter the tidal verification module, select a standard format tidal station data file and choose your plotting options. Click "**Plot**," and the inversion results will be displayed in the coordinate system below. A pop-up window will also generate the corresponding graphs based on your selected plotting options.

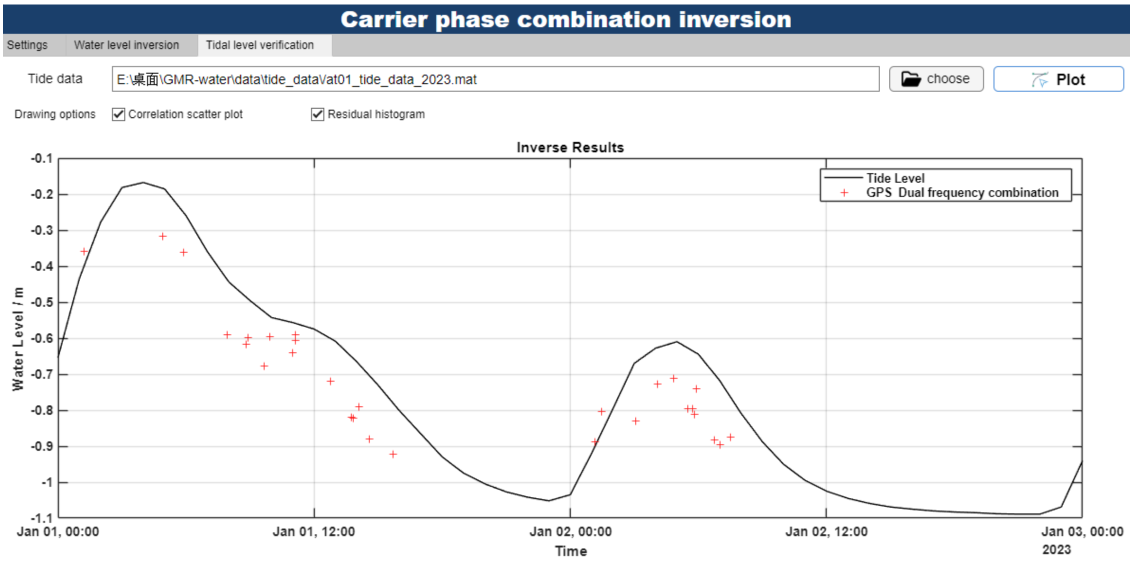

Figure 4-18. Interface of Tidal level verification

Carrier & pseudo range combination method

Similar to the carrier combination inversion, it is important to ensure that the selected combinations are from the same system and frequency band. For example, L1C and C1C within the GPS system should be chosen together.

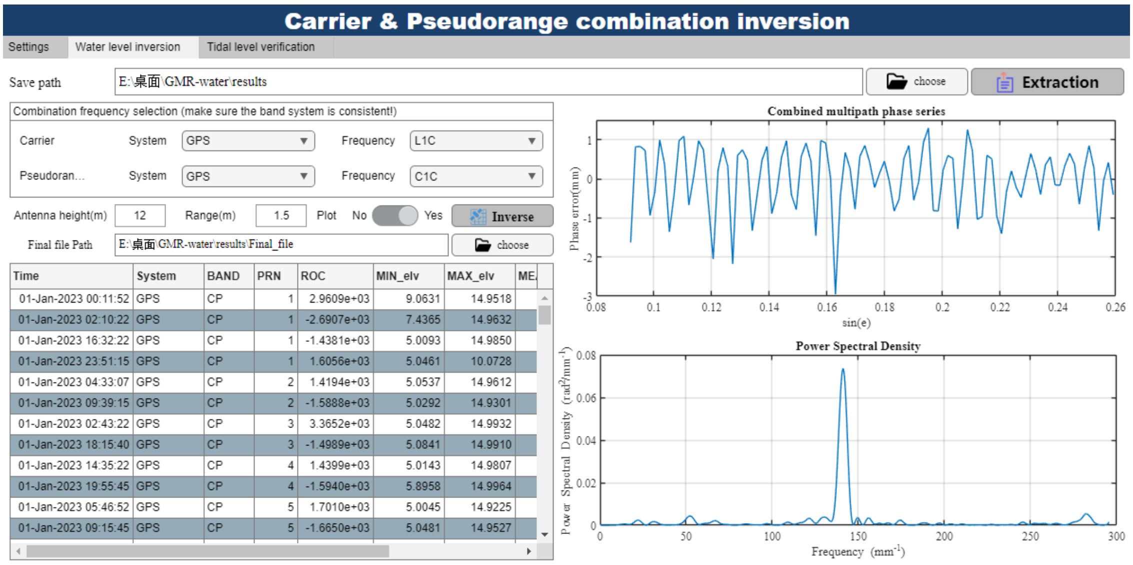

Figure 4-19. Interface of Carrier & pseudo range combination

3.5.5 Pseudo range dual/triple frequency combination method

All processes are consistent with the carrier combination inversion, with the only difference being that carrier phase is replaced by pseudo range.

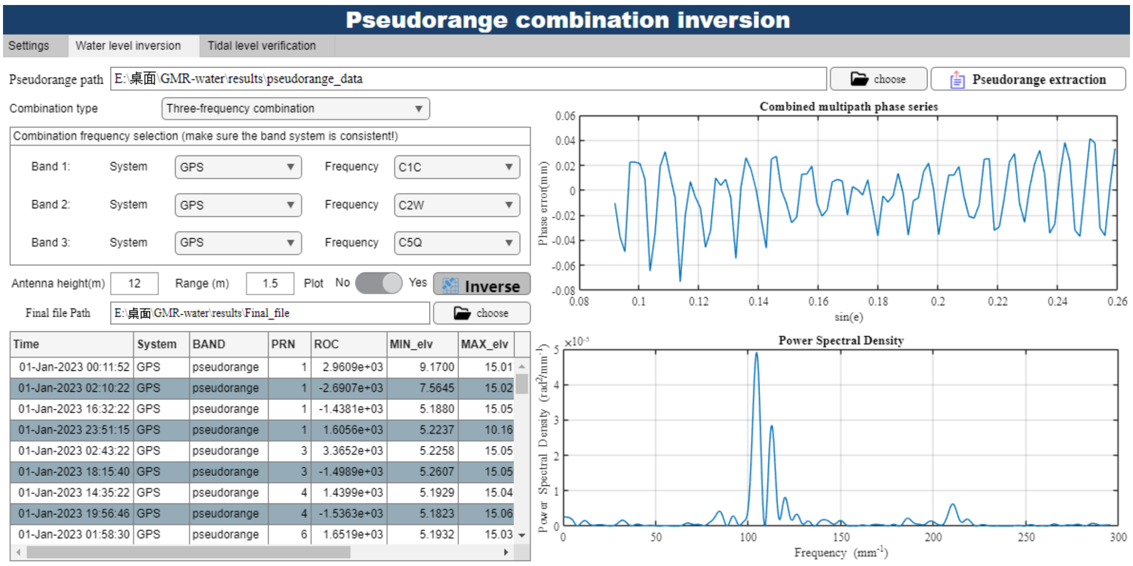

Figure 4-20. Interface of Pseudo range combination inversion

4.5 Batch processing

To improve data processing efficiency of this software, we provide a batch processing module that simplifies the workflows, reduces selection operations and also offers various visualization results. The overall process is divided into three parts: data extraction from RINEX files, inversion of reflection height, tidal correction, and output of the final file.

You only need to input three files/folders: a standard format station file, a folder containing the RINEX files, and a folder for the ephemeris files. Here you must provide information such as the RINEX version and the analysis center of the ephemeris. The RINEX and ephemeris files must be named according to the standard format. Furthermore, select the water level retrieval method and click the "Methods Settings" button to open "genMethodsSettings.m" for configuring the inversion process parameters. Alternatively, you can directly modify the content of "genMethodsSettings.m," and the program will execute the specified parameter settings during runtime. The execution process is divided into three parts, each of which can run independently. For the first execution, all parts must be selected to ensure the existence of the required files for each stage. Once the settings are complete, click "RUN," and the status indicator will turn green to show that the process is running. Note that the batch processing module supports parallel computation for multiple days of data to enhance processing speed, and the number of parallel processes can be adjusted based on your computer's performance.

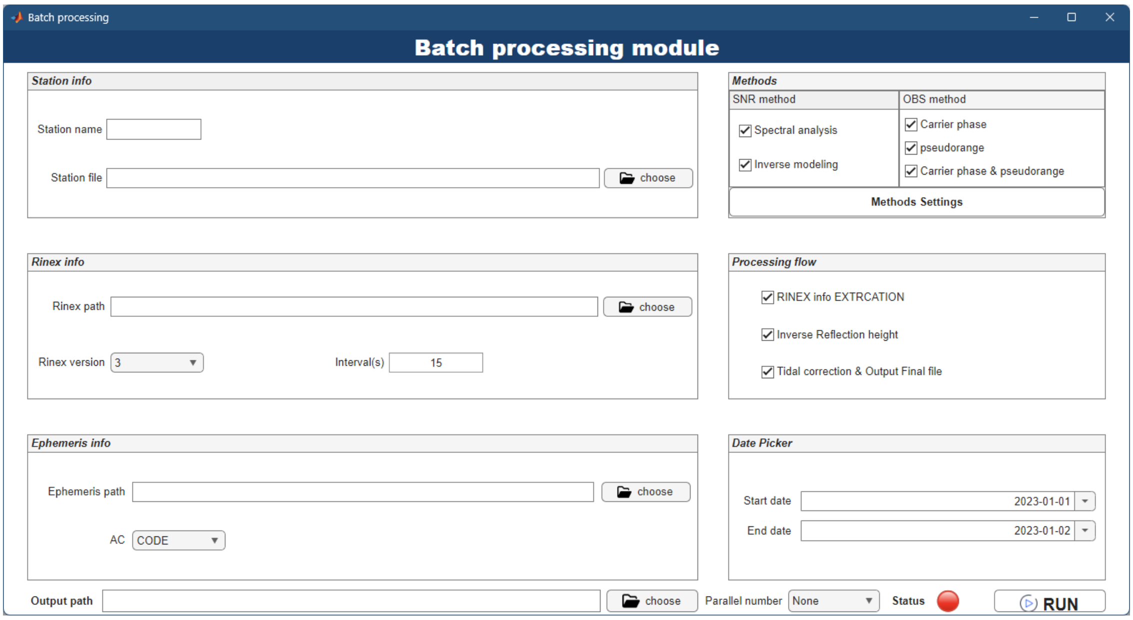

Figure 4-21. Interface of Batch processing module

4.6 Analysis Final File

All results produced by the former modules are considered as Final files. The "Analysis Final File" module is designed to process, analyze, combine, and visualize these results. This module supports selecting multiple final files from the same station. Optionally, standard tide gauge data can be provided to offer accuracy information. The analysis results will ultimately be output to the specified save path.

First, quality control can be performed on the raw final file data, allowing the exclusion of frequency points with poor accuracy. The 3-sigma rule can be applied for noise removal, and the quality-controlled file will be output. Furthermore, data in the final file can be processed using composite estimation, with the option to apply either a robust regression strategy or a B-spline strategy to obtain an evenly sampled water level series. Additionally, the module supports displaying correlation scatter plots, daily number, and other relevant information. Once you have completed these settings, click "**Analysis**", and the status indicator will turn green to show that the process is running. The module will generate images based on the settings and save them in .fig format, while the combined data will be saved in .mat format according to the selected combination options.

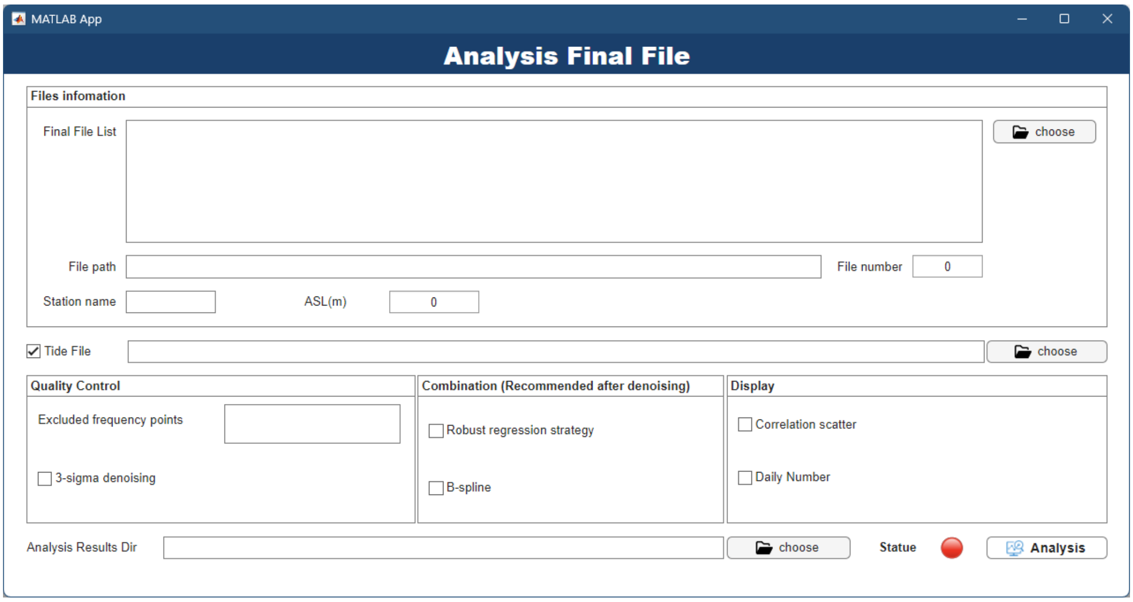

Figure 4-22. Interface of Analysis Final File

5 Usage of "GMR_Water.m"

The **GMR_Water** function provides a more convenient way to start the program, essentially executing the entire water level retrieval process by invoking the *"Batch Processing"* and *"Analysis Final File"* modules. It is driven by the input **"settings"** structure variable, which contains the necessary parameters for both modules. The specific parameter definitions are as follows:

settings

.Station_info_file

string - Path to the station information file (.mat format).

Example: 'E:\GMR-water\examples\data\station_info\scoa_info.mat'

.Rinex_path

string - Path to RINEX observation data folder.

Example: 'E:\GMR-water\examples\data\RINEX\SCOA'

.Rinex_version

string - Version of the RINEX files. Supported version: Rinex 2.x / Rinex 3.x.

Example: '3'

.Rinex_dt

numeric - Time interval of observation data (unit: seconds).

Example: 30 (30-second sampling rate)

.Eph_path

string - Path to precise ephemeris folder (SP3 files).

Example: 'E:\GMR-water\examples\data\sp3\COD'

.AC

string - Analysis center name for ephemeris data. Supported AC: CODE / GFZ.

Example: "CODE"

.Time

datetime array - Time range for data processing.

Format: [start_time, end_time].

Example: [datetime(2023, 1, 1, 'Format', 'yyyy-MM-dd'), datetime(2023, 1, 3, 'Format', 'yyyy-MM-dd')]

.Out_path

string - Output directory for processing results.

Example: 'E:\GMR-water\examples\results\Batch_result'

.station_name

string - Name of the station (used for labeling results).

Example: 'SCOA'

.methods

numeric array - Flags to enable/disable inversion methods.

Each position in the array corresponds to a specific method (1 = enabled, 0 = disabled).

Example: [1,1,1,1,1] (all methods enabled: Spectral analysis & Inverse modeling & Carrier phase & Pseudo range & Carrier phase and pseudo range)

.flow

numeric array - Flags to control processing steps (1 = enabled, 0 = disabled).

Example: [1 1 1 1] (all steps enabled: RINEX_info_EXTRCATION, Inverse_RH, Tidal_correction & save, Analysis & genReport)

.par

string - Parallel computing option.

Options: 'None' (disable) or '2' '4' '6' '8' (enable, Number of parallel).

.Tide_available

numeric - Tide data availability flag (0=disable, 1=enable tide data).

Example: 1

.tide_file

string - Full path to tide data file (.mat format).

Example: 'E:\GMR_Water\examples\data\tide_data\scoa_tide_data_2023.mat'

.efps

cell array of strings - excluded frequency points

Supported options: {'S2W','S8Q'} representing different frequency points.

Example: {'S2W'}

.sigma

numeric - Quality control threshold in standard deviations for outlier removal.

Example: 1

.rrs

numeric - robust regression strategy (0=disable, 1=enable).

Example: 1

.bspline

numeric - B-spline strategy (0=disable, 1=enable).

Example: 1

.cor

numeric - Correlation visualization flag (0=hide, 1=show correlation plots).

Example: 1

.day_num

numeric - display daily number (0=disable, 1=enable).

Example: 1

.results

string - Output directory path for analysis results.

Example: 'E:\GMR-water\examples\results\analysis_results'

We also provide a sample program, **"GMR_Water_AT01.m"**, which is designed for water level retrieval at the AT01 station. You can use this script as a reference to configure your own station settings accordingly.

6 Additional Notes

6.1 Tide data

Please note that the tide level data entered in this software should be in standard format, as shown in Figure 6-1. The tide level data is stored in .mat format, including two variables **xaxis** (timestamp stored in datenum format) and **slvl** (tide level data in meter). In addition, we provide a script named "**get_tide_data_from_unesco.m**" to help you download and save some tidal level data from https://www.ioc-sealevelmonitoring.org in standard format, as shown in Figure 6-2.

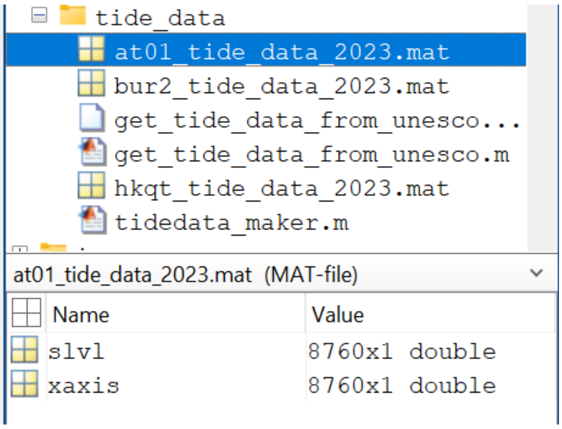

Figure 6-1. Example of standard tide data file

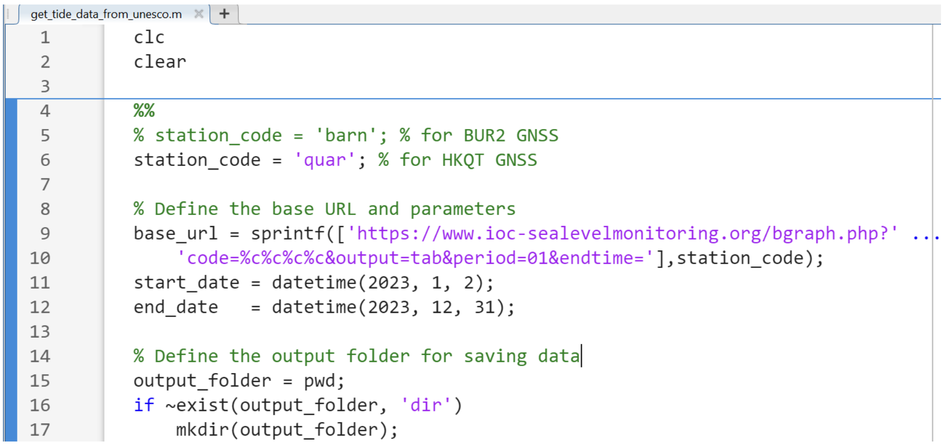

Figure 6-2. get_tide_data_from_unesco.m

6.2 Fresnel reflection region

The software provides a module to draw .kml files (of course, it can also be displayed through Google Earth), which is used to draw the Fresnel reflection area generated by the software. The interface is shown in Figure 6-3. Select the folder to store .kml files in path, please note that the .kml files should not exceed 250, which will reduce the efficiency of drawing.

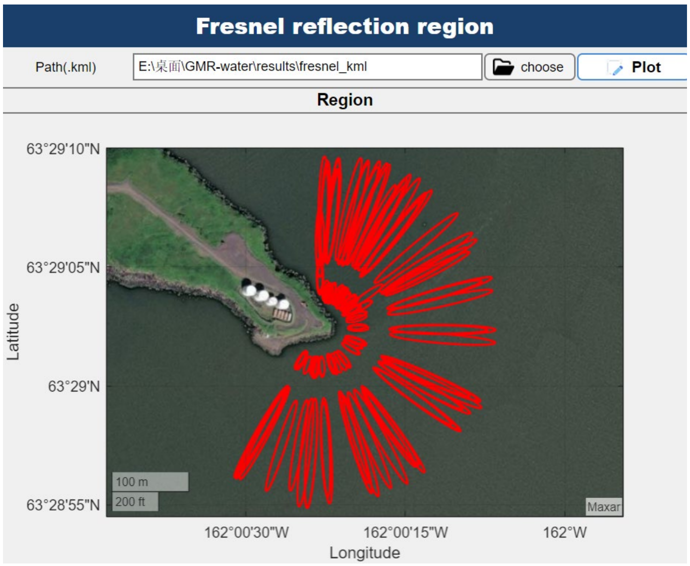

Figure 6-3. Interface of Fresnel reflection region

6.3 The format description of final file

All final outputs of GMR-Water adhere to a standard final file format. The unified naming format for the final files is as follows:

<StationName>_<StartDate>_<EndDate>_<Method>.mat

Where:

StationName: Four-digit GNSS station name

StartDate (EndDate): The start date (end date) of data processing in YYYY-MM-DD format

Method: The method used for inversion (e.g., SNR-Spectral, OBS-Carrier)

The file contains a struct variable "**Final_info**", the struct field name is "***<System><serise_type>***" (eg., GPSS1C, GALILEOCarrier), each field contains a table containing the field information, as shown in Figure 28:

Final_info | Final_info.GPSS1C

Final_info.GPSS1C

	1 Time	2 System	3 BAND	4 PRN	5 ROC	6 MIN_elv	7 MAX_elv	8 MEAN_AZI	9 RH	10 trop_c	11 tidal_cor
1	01-Feb-202...	"GPS"	"S1C"	3	3.357...	5.2455	15.0510	19.7511	11.9163	0.1386	0.1311
2	01-Feb-202...	"GPS"	"S1C"	11	1.598...	5.2436	15.0189	128.9479	11.9032	0.1326	0.0607
3	01-Feb-202...	"GPS"	"S1C"	17	-1.60...	5.2630	15.0522	44.2410	11.8107	0.1312	-0.0563
4	01-Feb-202...	"GPS"	"S1C"	29	1.437...	5.1878	15.0635	208.1436	11.6973	0.1372	0.0456
5	01-Feb-202...	"GPS"	"S1C"	24	-1.48...	5.2393	14.9660	162.3381	11.6914	0.1337	-0.0321
6	01-Feb-202...	"GPS"	"S1C"	19	-1.50...	5.2673	15.0631	52.7662	11.6596	0.1370	-0.0309
7	01-Feb-202...	"GPS"	"S1C"	2	1.422...	5.2566	15.0616	172.5802	11.8069	0.1303	0.0208
8	01-Feb-202...	"GPS"	"S1C"	4	1.327...	5.2073	7.0931	0.9445	11.7668	0.2629	0.0145
9	01-Feb-202...	"GPS"	"S1C"	20	1.881...	5.1907	14.9993	98.2253	11.7588	0.1367	0.0165
10	01-Feb-202...	"GPS"	"S1C"	6	-1.63...	5.1975	15.0464	31.2369	11.6779	0.1384	0.0036
11	01-Feb-202...	"GPS"	"S1C"	5	1.715...	5.2526	15.0341	110.2876	11.8001	0.1349	-0.0110
12	01-Feb-202...	"GPS"	"S1C"	18	1.427...	5.1887	14.9978	188.4573	11.7946	0.1365	-0.0126
13	01-Feb-202...	"GPS"	"S1C"	9	3.199...	5.2347	13.9181	8.1440	11.6650	0.1936	-0.0332

Figure 6-4. The header of the table in the Final file

7 Examples

7.1 AT01 Station

This section presents the water level retrieval results obtained by GMR-Water from processing GNSS observation data at the AT01 station from February 1 to February 28, 2023, **as a supplementary part of the manuscript.**

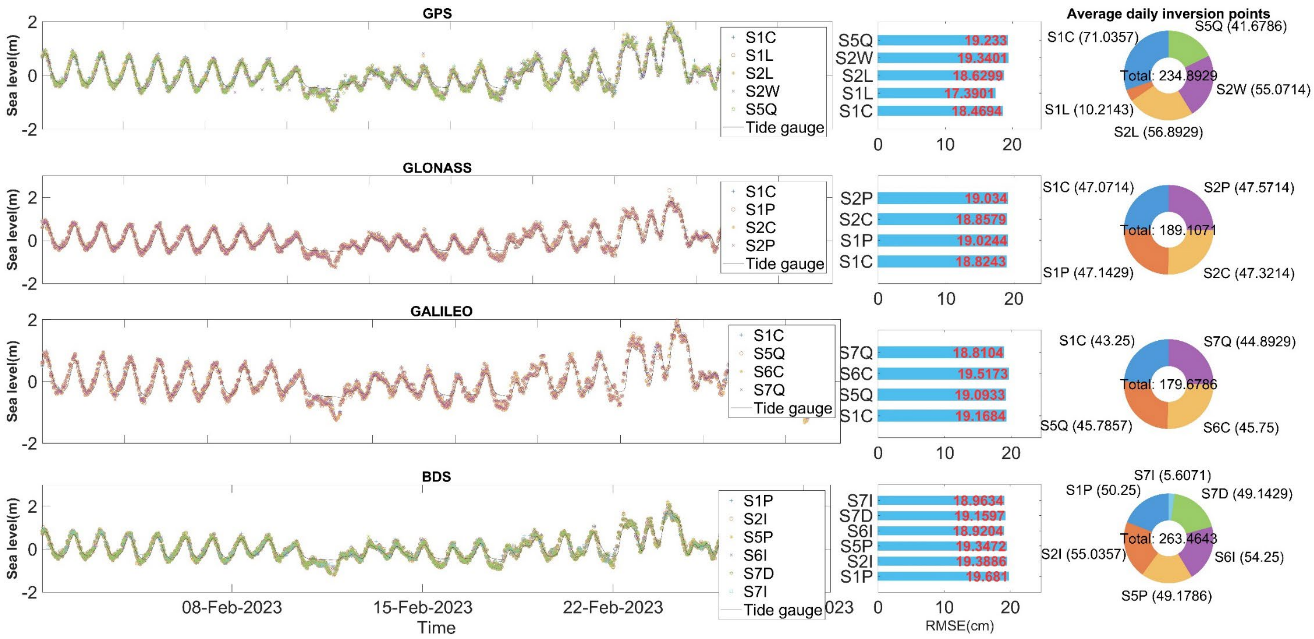

Figure 7-1. Retrieval results using SNR spectral analysis method, including multi-GNSS, multi-frequency time series, RMSE at each frequency, and the number of average daily inversion points.

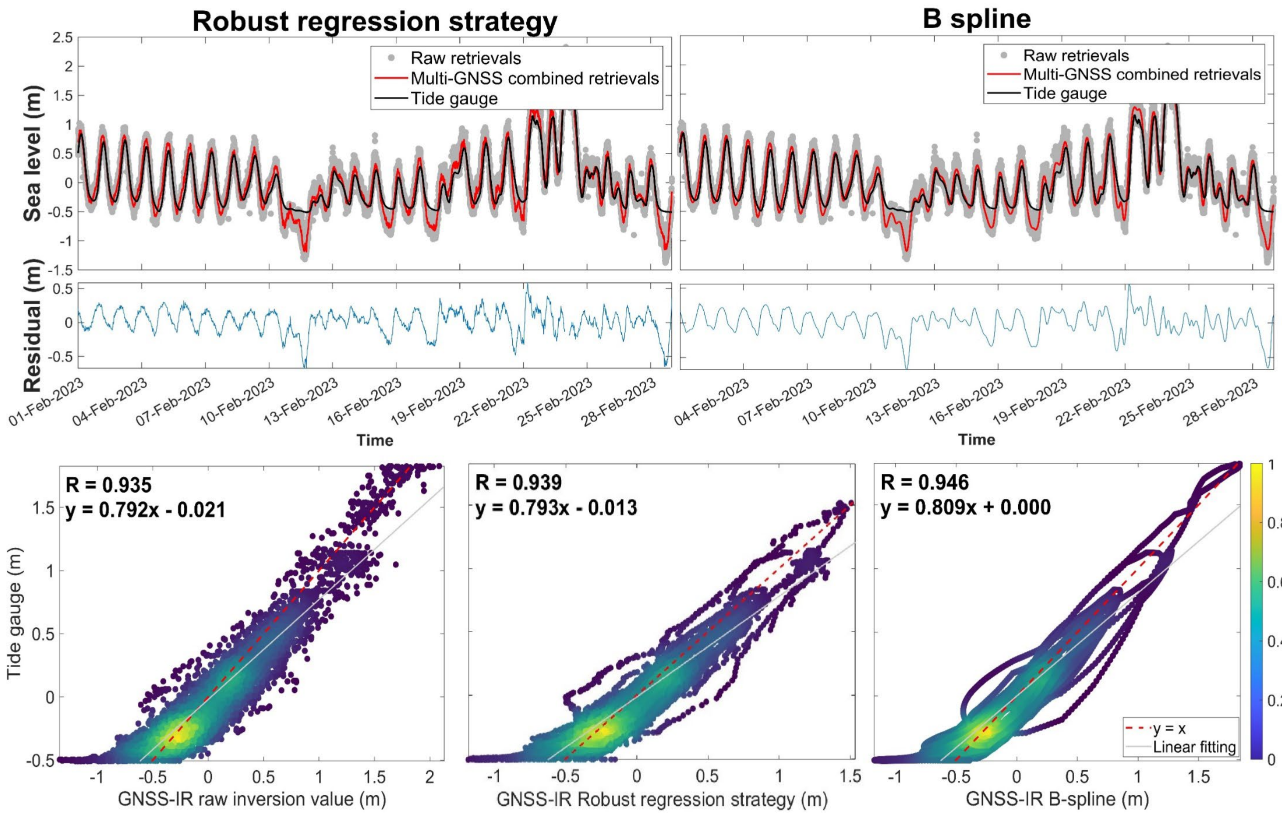

Figure 7-2. Time series and residual for two combination strategies (top), along with the scatter density plot (bottom)

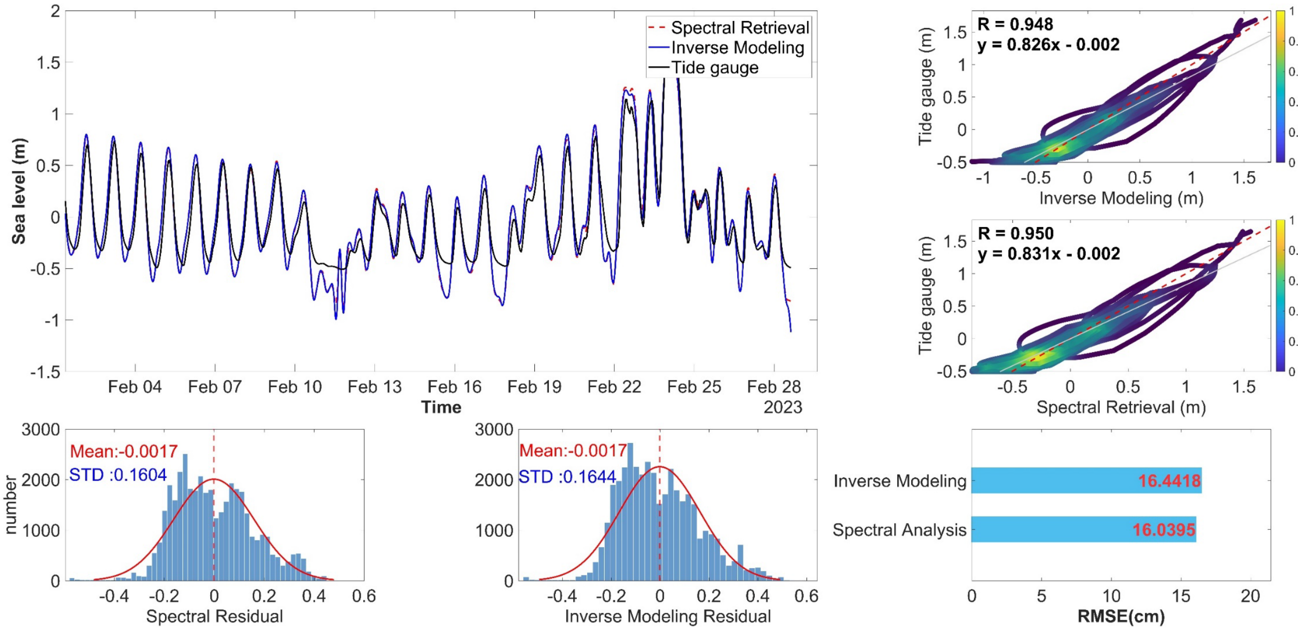

Figure 7-3. Time series of retrieval results, RMSE, scatter density plot, and residual histogram for the inverse modeling method and the spectral analysis method (B-spline estimation).

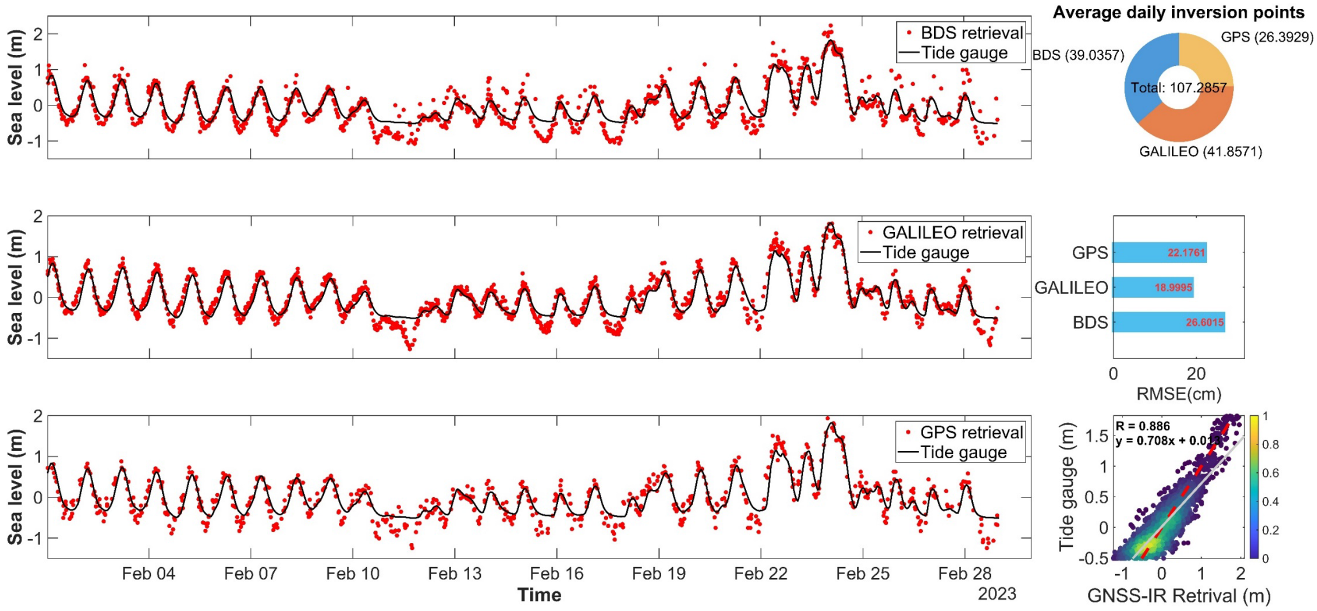

Figure 7-4. Water level retrieval results using the triple-frequency carrier phase combination method.

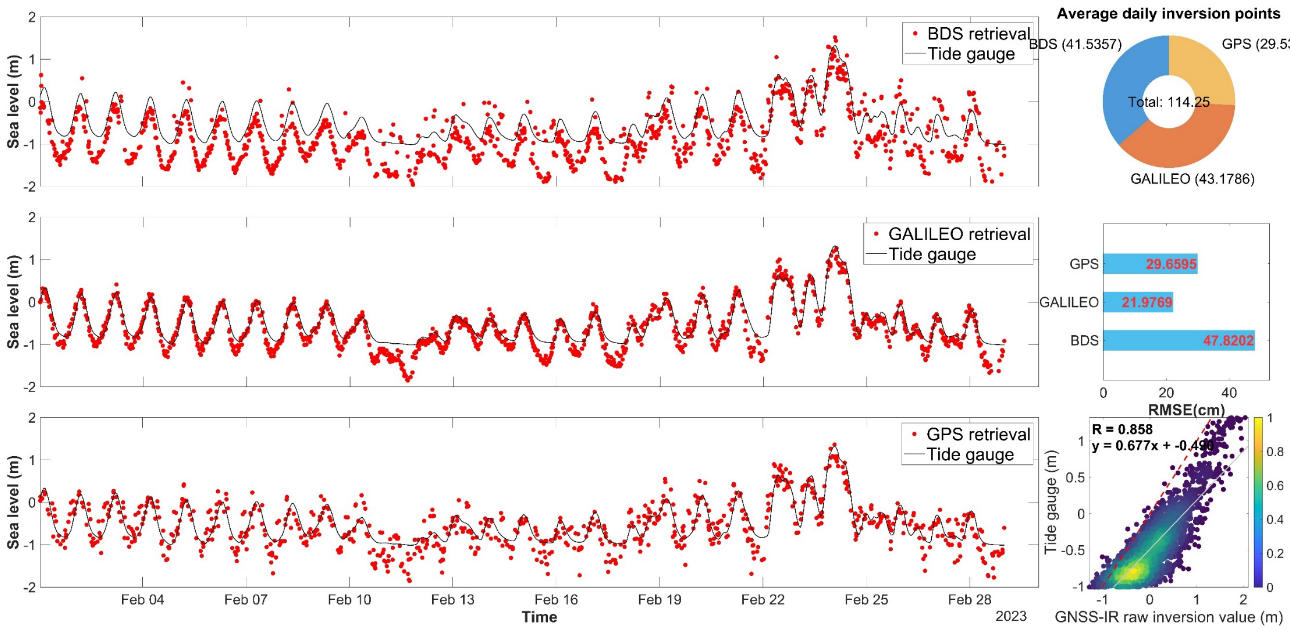

Figure 7-5. Water level retrieval results using the triple-frequency pseudo range combination method.

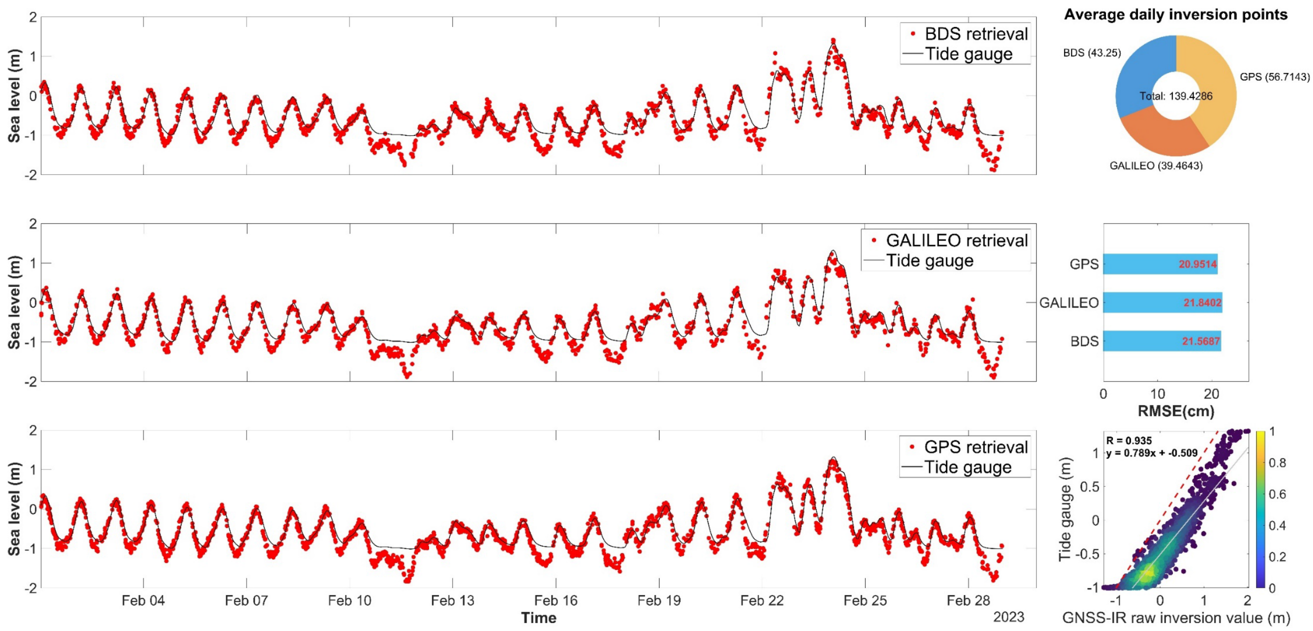

Figure 7-6. Water level retrieval results using the single frequency carrier phase and pseudo range combination method.

7.2 SCOA Station

The SCOA GNSS observation station is located in Socoa (43°23′42.83″N, 01°40′54.05″W), with the antenna positioned 10.66 meters above mean sea level. The station (SCOA) is part of the Réseau GNSS Permanent (RGP) operated by the French National Institute of Geographic and Forest Information (IGN) (http://rgp.ign.fr/). The station utilizes a Trimble NETR5 receiver and a Trimble Zephyr Model 2 (TRM55971.00) geodetic antenna to collect GNSS data.

Figure 7-7. Schematic Diagram of the Geographical Location and Surrounding Environment of the SCOA Station

In this example, we demonstrate the batch processing results using the "**GMR_Water"** function. Observational data from January 1 to January 5, 2023, were selected for inversion processing. The tide gauge data were obtained from a radar sensor located 2 meters away. The **WinLSP** strategy was adopted, and the results were analyzed. Some of the results are shown in the following figure:

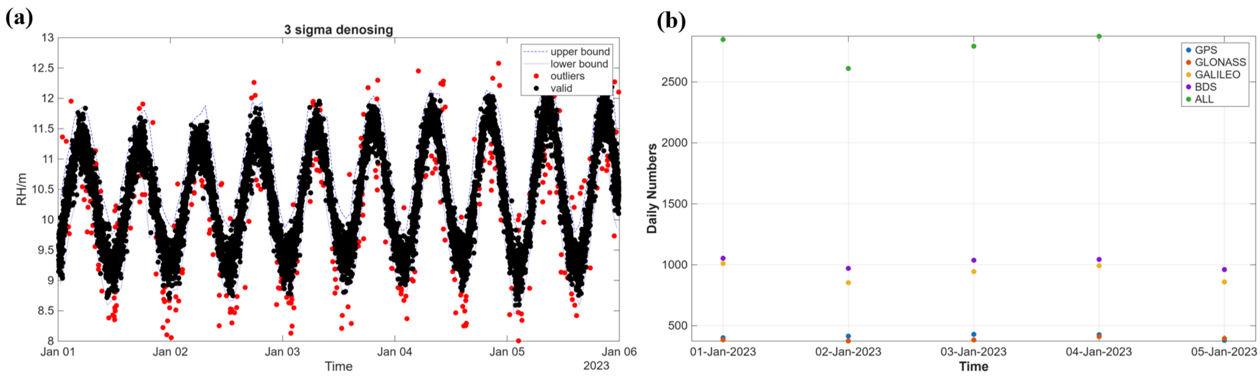

Figure 7-8. (a) Spectral analysis method 3-sigma denoising results. (b) Number of daily inversion points.

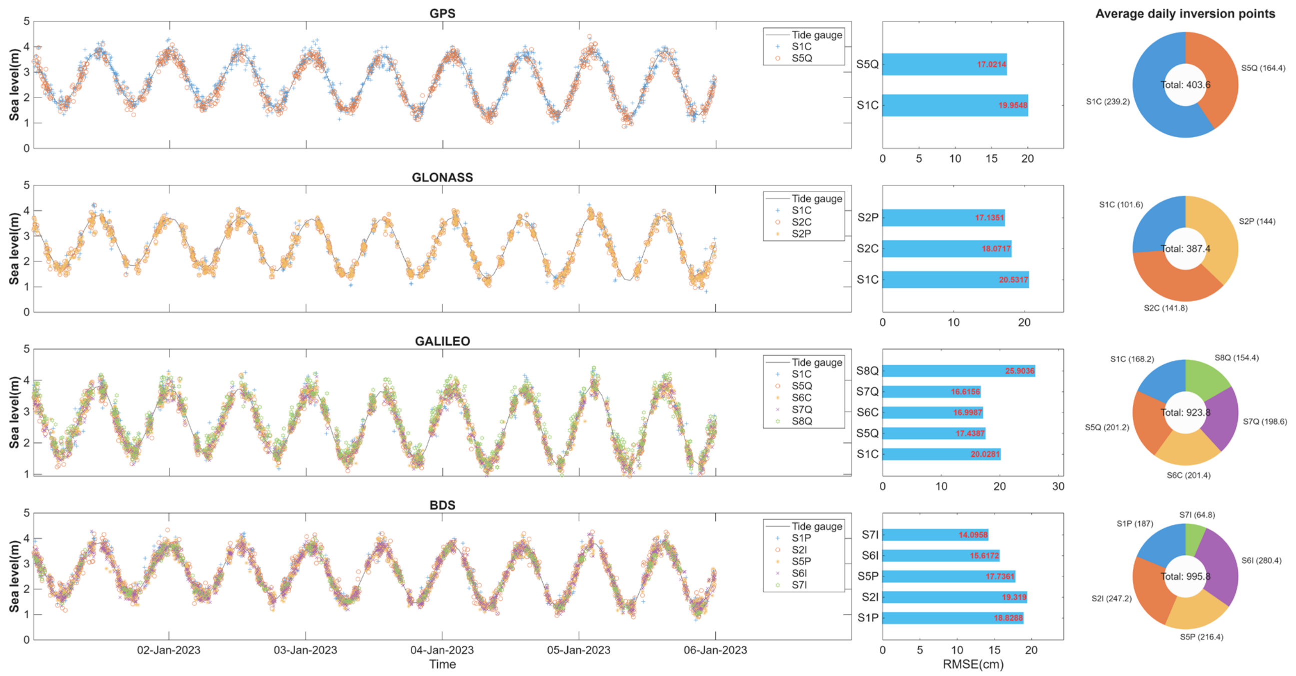

Figure 7-9. Spectral analysis method results, accuracy of each frequency point, and daily average number of inversion points.

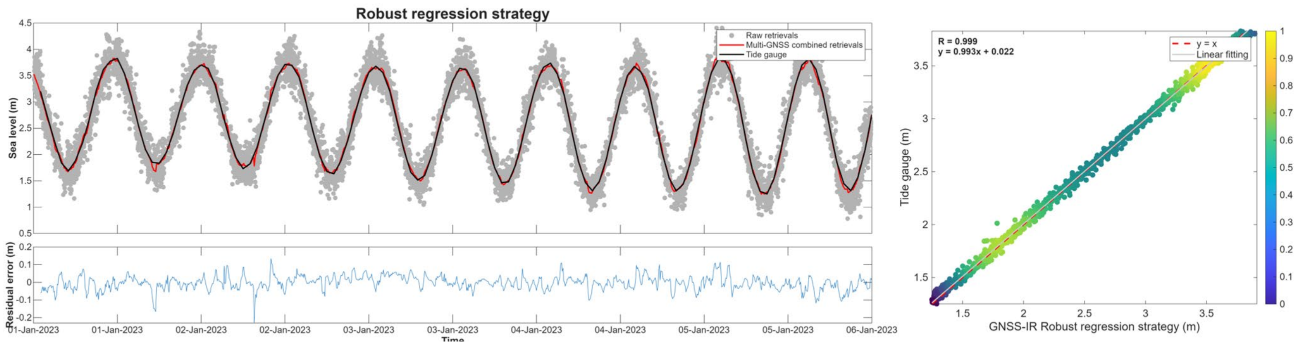

Figure 7-10. Spectral analysis method results, robust regression strategy estimation results, residuals, and correlation scatter plot.

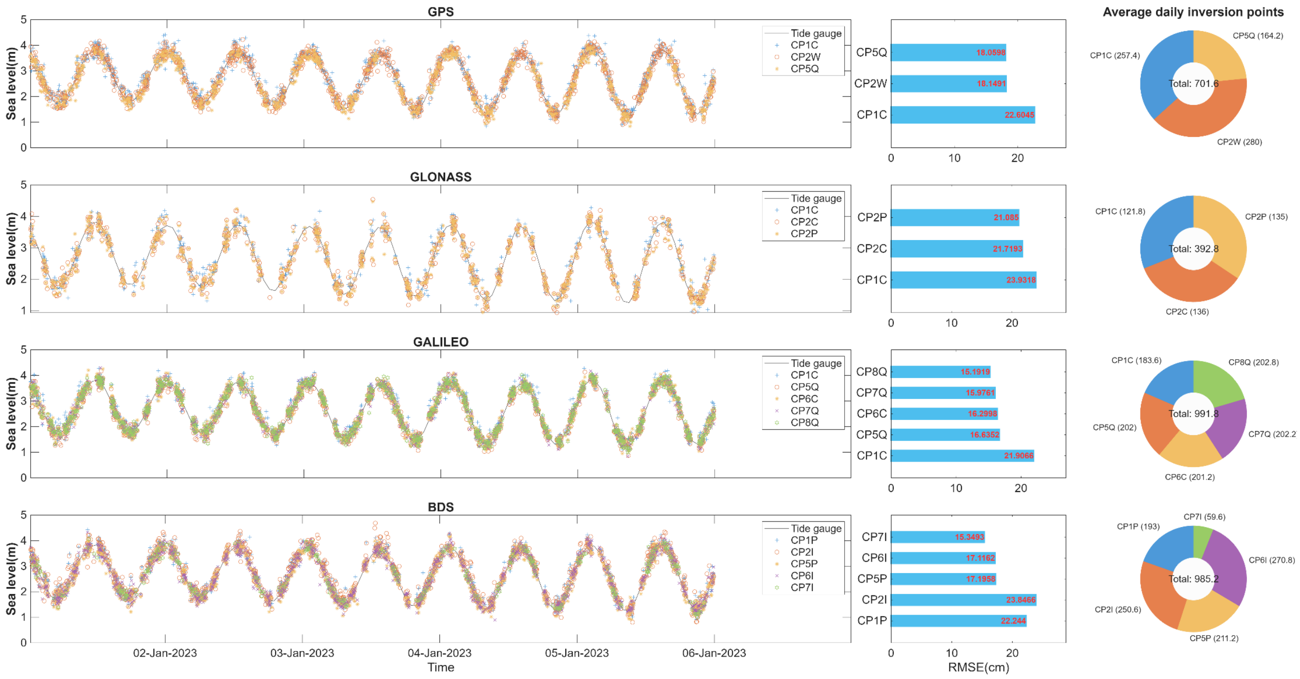

Figure 7-11. Single-frequency carrier phase & pseudo range combination results, accuracy of each frequency point, and daily average number of inversion points.

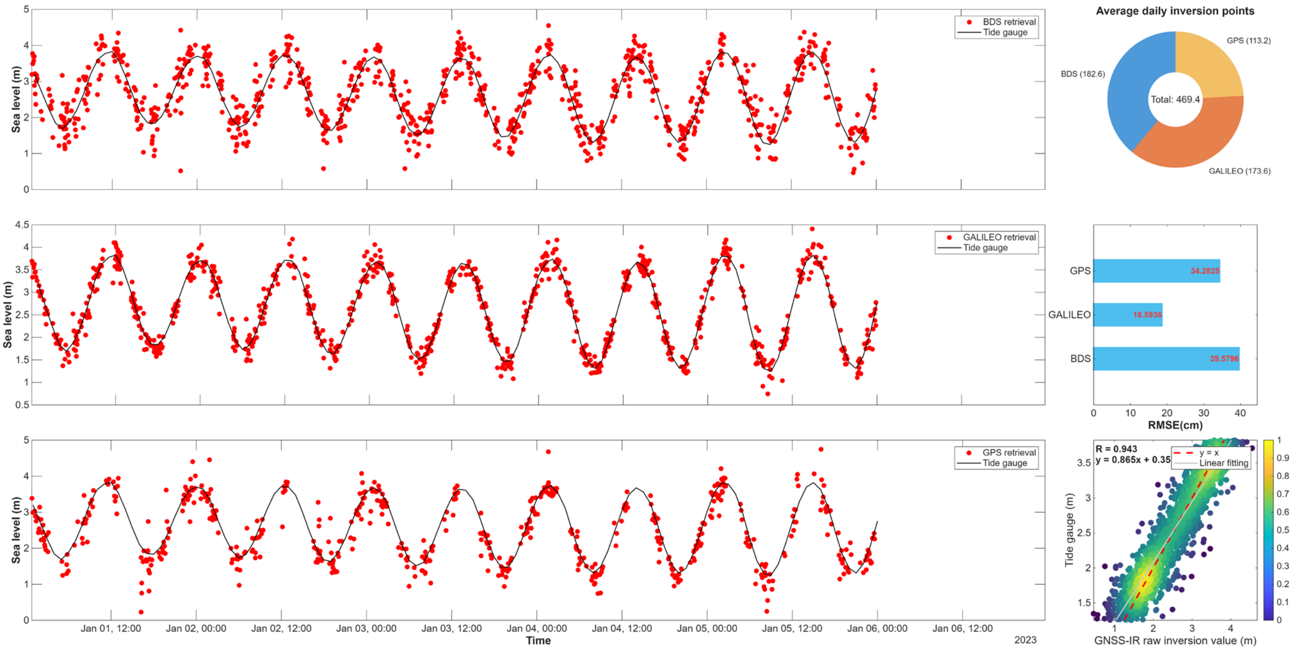

Figure 7-12. Triple-frequency carrier phase combination results, daily average number of inversion points for each system, accuracy, and correlation scatter plot.

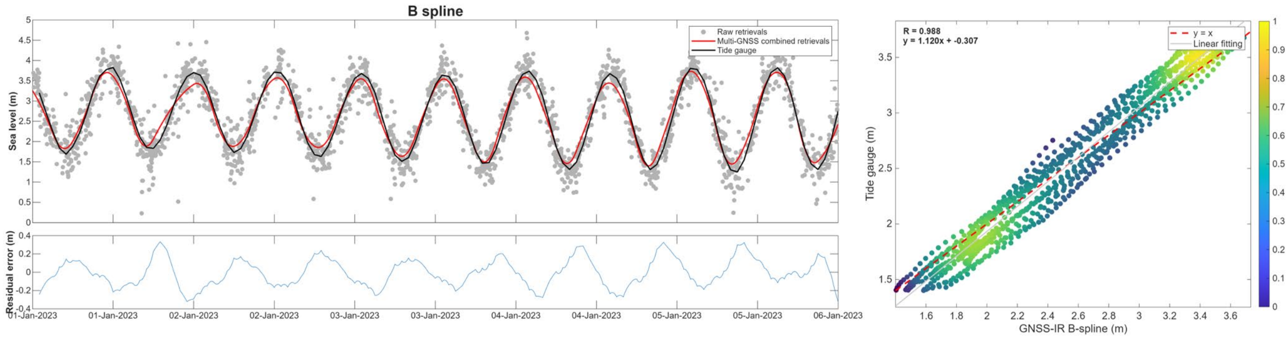

Figure 7-13. Triple-frequency carrier phase combination results, B-spline estimation results, residuals, and correlation scatter plot.

8 Contact

If you have any questions about this software or have suggestions for improvement, please feel free to contact us.

Fei Yang: yangfei@cumtb.edu.cn

Ruixian Hao: 2110200210@student.cumtb.edu.cn

Vitamin_N@outlook.com