SeafloorMapper

User Manual

Version 1.0

2023.10.20

Yiwen Lin

Contact: eeeven0420@gmail.com

Table of Contents

Introduction

About SeafloorMapper

SeafloorMapper is a user-friendly software designed for researchers in the field of bathymetry working with ICESat-2 ATL03 data. It simplifies the process of generating seafloor data by seamlessly integrating automated and manual methods, providing a comprehensive solution for mapping the seafloor.

Target Audience

This user manual is intended for researchers with a specific interest in extracting seafloor data from ICESat-2 ATL03 data.

Getting Started

System Requirements

Before installing SeafloorMapper, please ensure that your system meets the following requirements:

- Operating System: Windows 7/8/10/11
- Internet connection
- 500 MB of RAM
- 1.5 GB of available disk space

Installation

To install SeafloorMapper, follow these steps:

- Download the installation file from https://github.com/ylin152/SeafloorMapper/releases/tag/v1.0.
- Run the installation file and follow the on-screen instructions.

How to use

Prepare your data

The software exclusively processes ICESat-2 ATL03 data, which is organized into tracks. ICESat-2 utilizes both strong and weak beams, resulting in a total of 6 tracks - 3 from strong beams and 3 from weak beams. Each track represents a distinct path of ICESat-2 laser measurements. The received photons (hereafter referred to as “points”) are categorized into noise and signal photons with varying confidence levels, where a higher confidence level indicates a higher probability of a point being a signal photon. For more information, please refer to the documentation available at https://doi.org/10.5067/ATLAS/ATL03.006. While the software can accept ICESat-2 ATL03 data in h5 format from any source, including those directly downloaded from NSIDC, it is recommended to obtain the ICESat-2 ATL03 data from OpenAltimetry (https://openaltimetry.earthdatacloud.nasa.gov/data/icesat2/) with the following instructions:

1. Firstly, on the main page, select any date and zoom into your area of interest. Then, in the top-left corner, click "SELECT A REGION" and drag a rectangle on the map (see Figure 1), and click “View signal photons” in the pop-up box. This brings up any photons available within the region, for the selected date.
2. Then, look for the "ATL03 PHOTON HEIGHTS" section on the profile page (Figure 2). Here, you can draw a rectangle on the plot to zoom in and view the data in more detail. Repeat the zooming process as needed until you find the seafloor structure.
3. Once you've identified the seafloor, click "Download subsetted HDF5 (via NSIDC)" at the bottom of the page (ensure you have an NSIDC account set up). This will initiate the download of a subsetted h5 file, which will be saved as a zipped file on your computer. Unzip the downloaded file to access the data, and you are now ready to start your work!

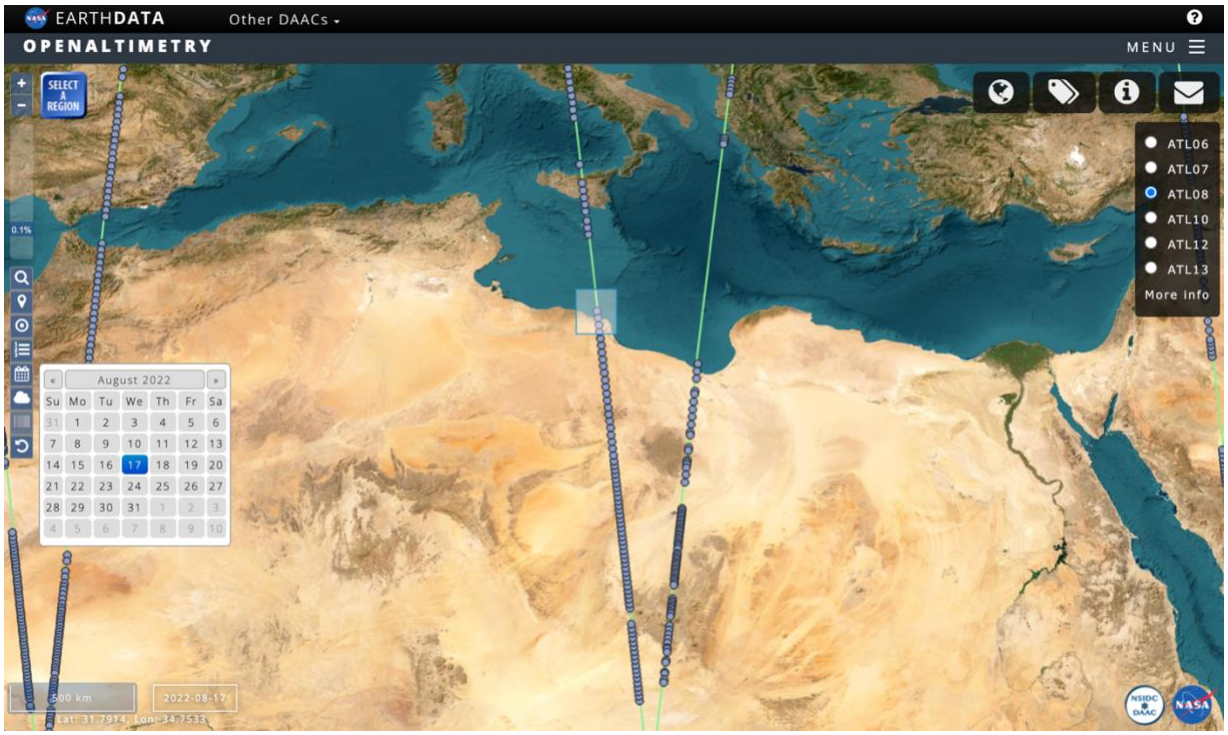

Figure 1: OpenAltimetry web interface.

Figure 2: ICESat-2 ATL03 Elevation Profile for selected area and date.

Main Interface

The main interface of the software (Figure 3) offers two distinct sections for seafloor mapping:

Auto Seafloor Mapping (right column of tools)

In this section, users should use the following tools: Data Preprocessing –> Model Prediction –> Manual Annotation –> Data Output. This workflow is for users who want to use the trained PointNet++ model (https://doi.org/10.1016/j.jag.2023.103512) for automatic classification of seafloor photons, and manually correct model prediction errors before outputting the classified photons.

Manual Seafloor Mapping (left column of tools)

In this section, users have two primary workflow options:

- File Conversion –> Manual Annotation –> Training Data Creation: This workflow is designed for users who want to create training data to serve as input for PointNet++ model training. Please note that the model training process itself is not integrated into this software, but you can access the relevant code to re-train the model, as well as existing training data, on https://github.com/ylin152/SeafloorMapping_PointNet2.git.
- File Conversion –> Manual Annotation –> Data Output: This workflow is for users who want to manually annotate seafloor data without using the PointNet++ model. You can annotate the data and directly generate the output as needed.

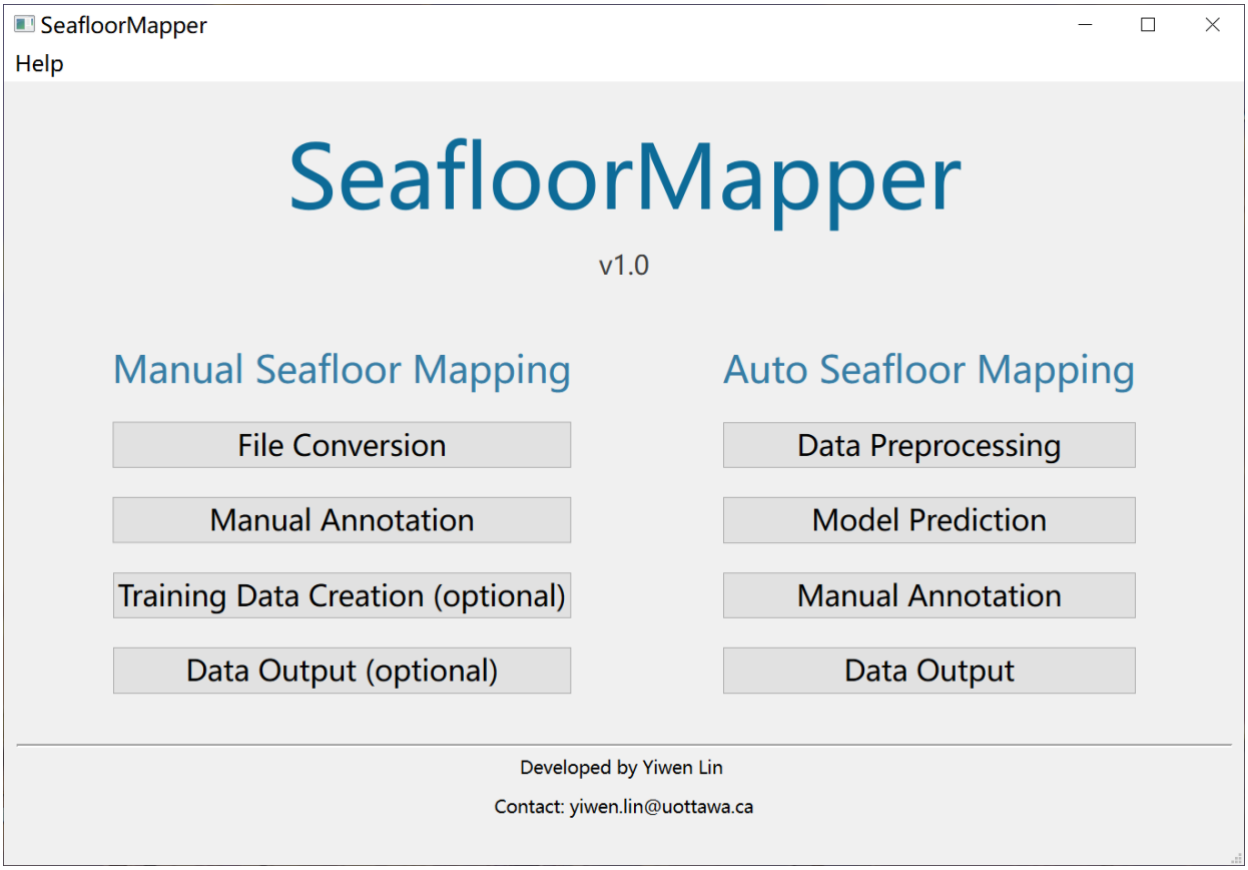

Figure 3: Main interface.

Auto Seafloor Mapping

For automated seafloor mapping, please follow these step-by-step instructions:

1. Click “Data Preprocessing” and select root data folder (Figure 4). Ensure that there is at least one h5 file in the root folder. Click “Preprocess” – a window will pop up to let you know that the process has started, you can click “OK” to remove this window. Once finished, two new folders, named “csv_data” folder and “input_data”, will have been generated wsithin the root folder, and you will see another pop-up window telling you that the process has been executed successfully. Click “OK” in this window to close it, and also close the Preprocessing window to continue.

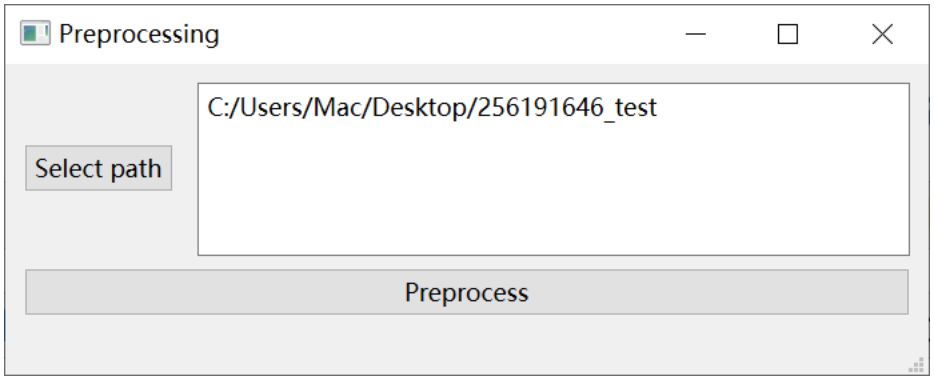

Figure 4: Data Preprocessing Window.

2. Click “Model Prediction” and select the root folder (Figure 5). The default threshold and the number of predictions for model prediction are 0.5 and 10, respectively. The threshold determines the probability at which points are classified as belonging to the seafloor or non-seafloor. The number of predictions signifies how many times you want the model to make predictions on the data, where a higher value improves the robustness and accuracy of the results, but also extends the prediction time. You can test with different parameters, and with the help of the following step (“Manual Annotation”), to find the parameters that best suit your needs. If you want to output the files without any seafloor points to a separate folder, click “Subfolders”. Otherwise, keep that option unchecked. After configuring the parameters and the output option, click “Predict”. As for the

preprocessing window, a pop-up window will tell you that “the operation has started”, and you can click “OK” to remove this window. When the prediction is complete another window will appear to tell you that is the executed successfully, you can also click “OK” to remove this window. You will then find a folder named “model_predictions” created in the root folder. In the “model_predictions” folder, you will see a sub-folder labeled “pred_xx_yy”, where xx and yy are the “threshold” and “number of predictions” parameters you set for the model prediction. **Important note:** The model prediction process can take a long time to complete for a large granule, depending on your CPU speed and processing parameters. To estimate the completion time, you can follow along as new csv files are created in the model_predictions/pred_xx_yy/ folder. The process will be complete when the number of csv files in that folder equals the number of txt files in the input_data folder. When prediction is done, you can close the Model Prediction window.

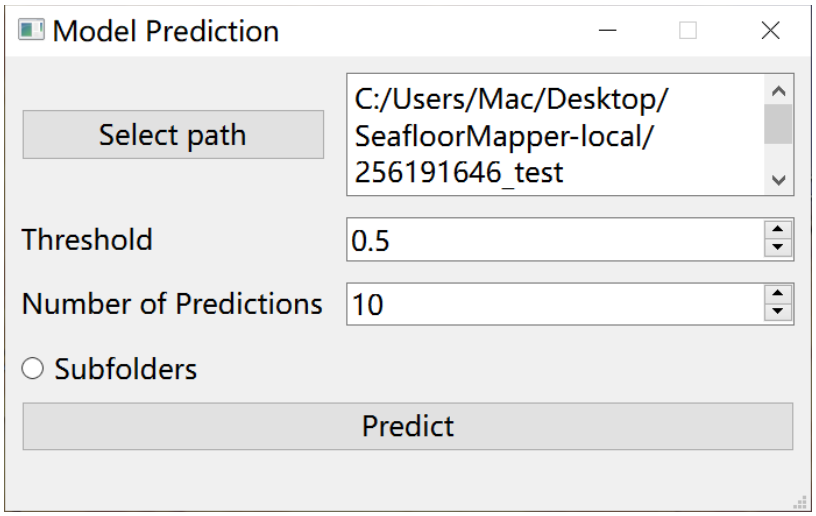

Figure 5: Model Prediction window.

3. Click “Manual Annotation” to open the annotation window (Figure 6). Open the “pred_xx_yy” folder (File -> Open folder). All files within the folder will be listed on the right side. All the filenames will appear similar in the window, but you can hover the mouse over each one to see the full filename. Double-click each file in turn to open it. The toolbar at the top enables restoring to the original view,

zooming in/out along the x axis, zooming in/out along the y axis, and drawing on the profile. The seafloor points are displayed in orange. Points predicted as non-seafloor, if belonging to medium confidence, are shown in green, while those with high confidence levels are displayed in blue. Annotate the seafloor points that were missed during the model prediction by using the “pencil” tool at the top-right to draw irregular polygons. Use the “Add” button to identify photons within a polygon as belonging to the seafloor, or conversely use the “Remove” button to identify them as ***not*** belonging to the seafloor. Use “Undo” or “Redo” to reverse the most recent action. You can also click the “Map” tab to view the geographic distribution of the points (Figure 7), which can serve as a reference for annotation. Before saving your annotation, you have to set an export folder by clicking “Set export folder” (Figure 8). By default, it will create a folder called “pred_xx_yy_annoated” under the root folder. Lastly, click “Export file” to save your annotation. Note that every file must be saved to the export folder, even if no seafloor points has been found. When you’re done with the manual annotation, close the annotation window.

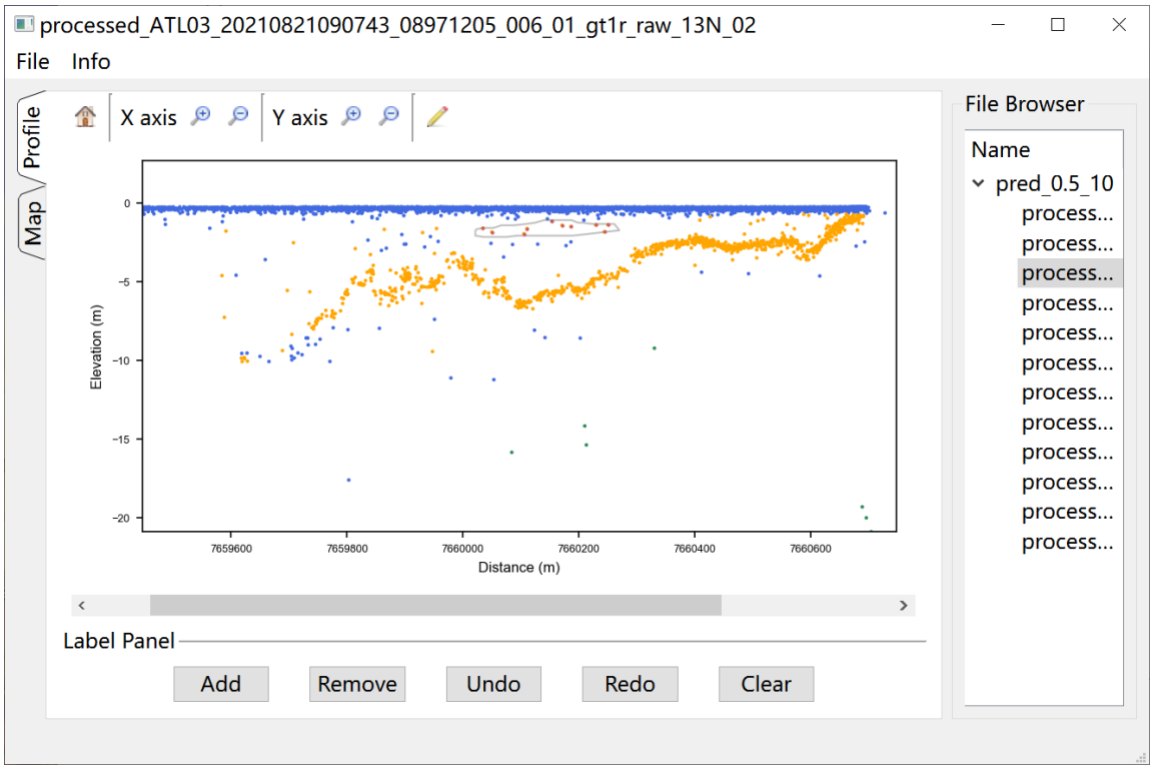

Figure 6: Manual Annotation Window.

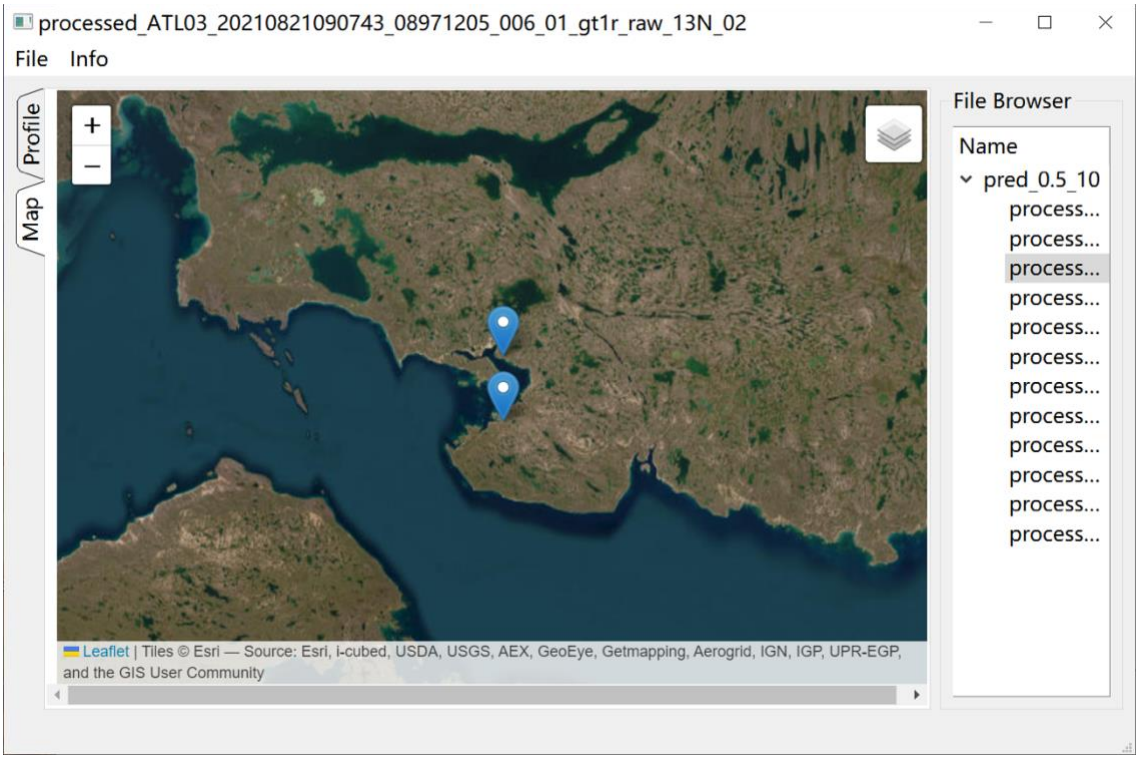

Figure 7: Annotation Window - Map view.

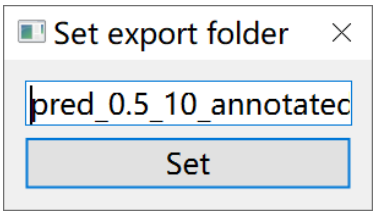

Figure 8: Set annotation export folder.

4. Click “Data Output” (Figure 9). Select the “pred_xx_yy_annoated” folder as the input. You can choose to apply refraction correction to the data, and select whether to output all points or only seafloor points, and whether to output separate tracks or a single combined track. Refraction correction is performed through: 1) for each seafloor point, identify the closet water points, where the projected y distance to that point is lower than 100 meter, 2) calculate the mean elevation of those identified water points, 3) apply the approximate correction equation $Z' \approx Z + 0.25416D$, where, D represents the uncorrected depth between the seafloor point and the water surface, Z is the uncorrected elevation of the seafloor point, and Z' is the corrected elevation of the seafloor point (please refer to doi:10.3390/rs11141634 for details). Set the output format as csv or txt. The default output path is “pred_xx _yy_annotated_output”, but you can specify the path yourself. Once all parameters are set, click “Output” – windows will pop up to let you know that the operation has started, and ended, click “OK” to close them. The output data will be stored in the user-specified folder, or in the default output folder “pred_xx_yy_annotated_output” under the “model_predictions” folder. Close the output data window when you are done.

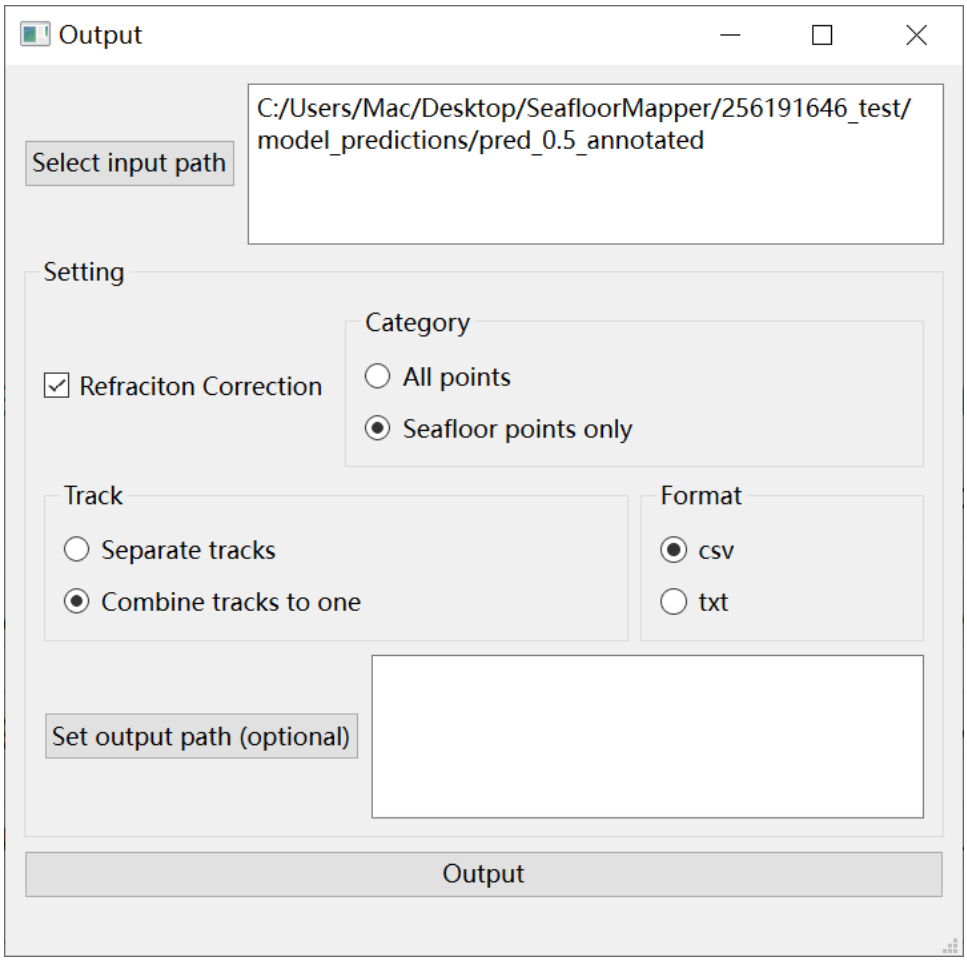

Figure 9: Output Data Window.

The final output file includes the following columns: “x”, “y”, “elev”, "lon”, “lat”, “class”, “prob”, “pred” and “w_elev”. The "x" and "y" columns represent the x and y distances projected from longitude and latitude coordinates to the local UTM zone, which is automatically calculated by the software and appended to the filename. The "elev" column indicates the elevation of the points relative to the ellipsoid, while the "lon" and "lat" indicate the longitude and latitude. The "class" field denotes the original classification by the preprocessing algorithm, where 5 corresponds to water surface, and 3 indicates other points. The "signal_conf_ph" column is the confidence level of the points being signal photons, where 1 denotes medium confidence level and 2 denotes high confidence level. The "prob" column denotes the probability, ranging from 0 to 1, of the points belonging to the seafloor class as predicted by the model. The "pred" column indicates the predicted class assigned by the model based on the “prob” column, with 0

representing non-seafloor and 1 representing seafloor. Lastly, the "w_elev" column represents the mean elevation of the closest water points to a given point if its "pred" column is 1 (see section on refraction correction on page 12 for information on how the closest water points are found). In addition, selecting the option “combine tracks to one” will create a new 'track' column, indicating the track ID to which each point belongs.

Training Data Creation

To create training data for the PointNet++ model, the steps are as follows:

1. Click “File Conversion” and select the root data folder (Figure 10). Be sure to check the “split data for training data creation” option. Then, click “Convert”. This action results in the creation of two folders: “csv_data” and “split_data”. Windows will pop up to tell you that the operation has started, and that it has successfully executed. Click “OK” to close these windows, and close the File conversion window.

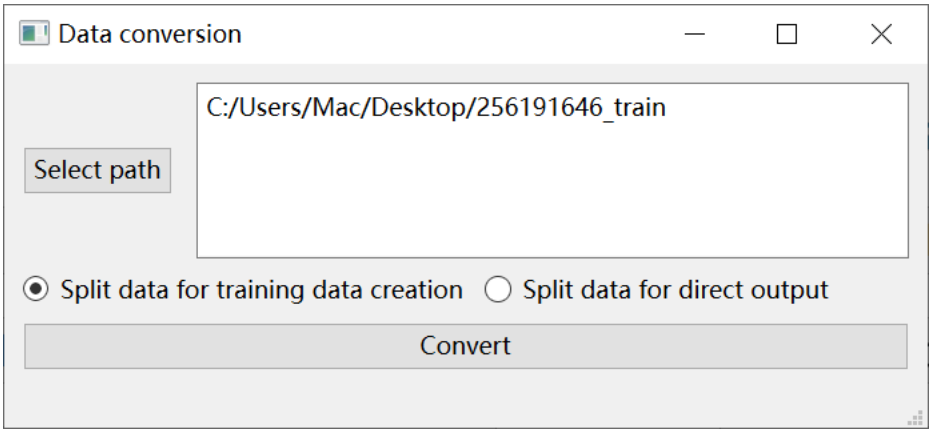

Figure 10: File Conversion Window.

2. Click “Manual Annotation” to open the annotation window. Open the “split_data” folder (File -> Open folder). All files within the folder will be listed on the right side of the window. Double-click each file to annotate them all in turn.
    a. Before you start annotating the data points, set the folder to save the annotated files to by clicking “Set export folder” (Figure 14). By default, it will create a folder called “split_data_annoated” under the root folder.
    b. Annotate the points. See the instructions for manual annotation starting on page 9.
    c. Lastly, click “Save file” to save your annotation. Noted that every file must be saved to the export folder, even if no seafloor points has been found.

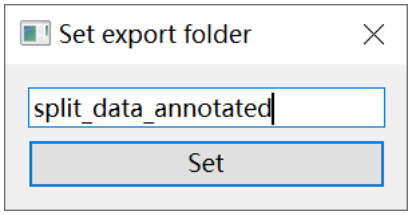

Figure 11: Set the annotation export folder.

3. Click “Training data creation (Figure 12, optional)”. Select the “split_data_annoated” as the input. Click “Create from predicted data” option if you want to generate the training dataset based on predicted data from “Auto Seafloor Mapping”. Otherwise, keep that option unchecked. This step will create a new folder named “input_data” within the root folder.

The final output file is in txt format and includes the following columns: “x”, “y”, “elev”, "lon”, “lat”, “class”, “signal_conf”, and “annot”. Note that this output file does not include header names for these columns in order to align with the input requirements of the PointNet++ model. See the description on page 13-14 for what the data in these columns mean. The "annot " column records the annotation made by the user, with 0 representing non-seafloor points and 1 representing seafloor points.

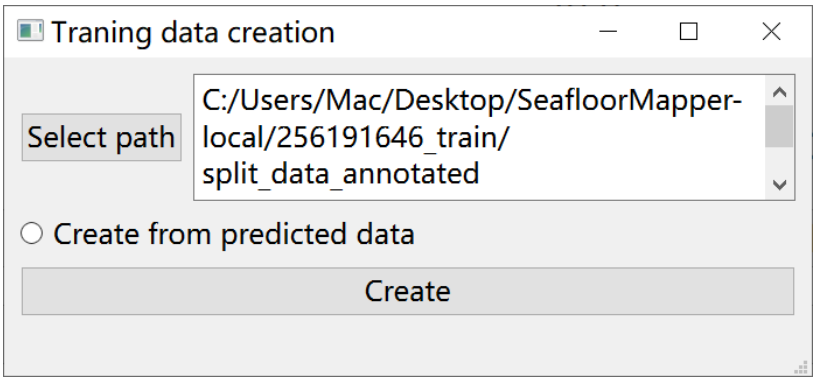

Figure 12: Training Data Creation Window.

Manual Seafloor Mapping

For manual seafloor mapping, please follow these steps:

1. Click “File Conversion” and select the root data folder (Figure 13). Make sure to check the “split data for direct output” option. Then, click “Convert”. This results in two folders: “csv_data” and “split_data”.

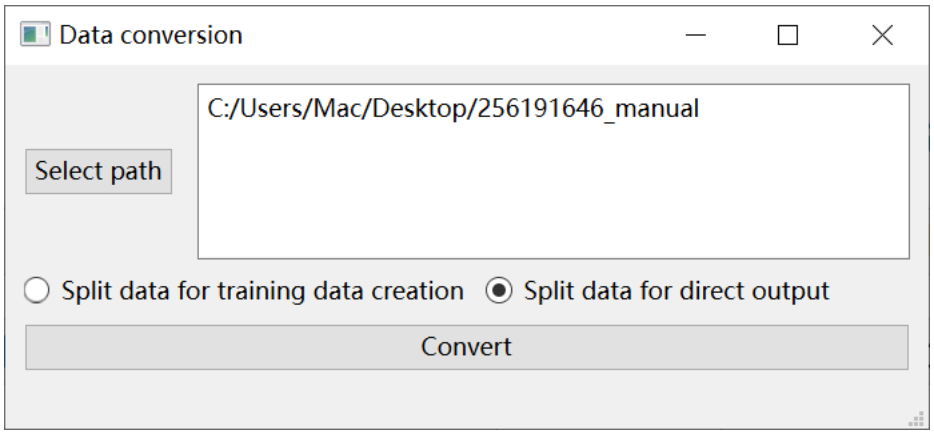

Figure 13: File Conversion Window.

2. Click “Manual Annotation” to open the annotation window. Open the “split_data” folder (File -> Open folder). All files within the folder will be listed on the right side of the window. Double-click each file to annotate them all in turn.
    a. Before you start annotating the data points, set the folder to save the annotated files to by clicking “Set export folder” (Figure 14). By default, it will create a folder called “split_data_annoated” under the root folder.
    b. Annotate the points. See the instructions for manual annotation starting on page 9.

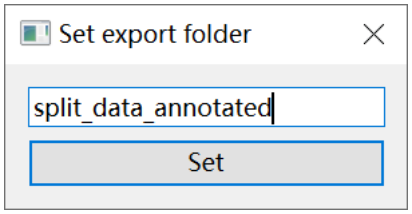

Figure 14: Set annotation export folder.

c. Lastly, click “Save file” to save your annotation. Noted that every file must be saved to the export folder, even if no seafloor points has been found.

3. Click “Data Output (Figure 15, optional)”.
    a. Select the “split_data_annoated” folder as the input path.
    b. Choose whether you want to apply refraction correction to the seafloor points.
    c. Select whether to output all points or seafloor points only.
    d. Select whether to output the data in separate tracks, or in one combined track.
    e. Set the output format as csv or txt.
4. The default output path is “split_data_annotated_output”, but you can specify the path yourself. Once all parameters are configured, click “Output”, and click ok in the window that says the operation has started.
5. The output data will be stored in the default output folder “split_data_annotated_output” or in the user-specified folder.

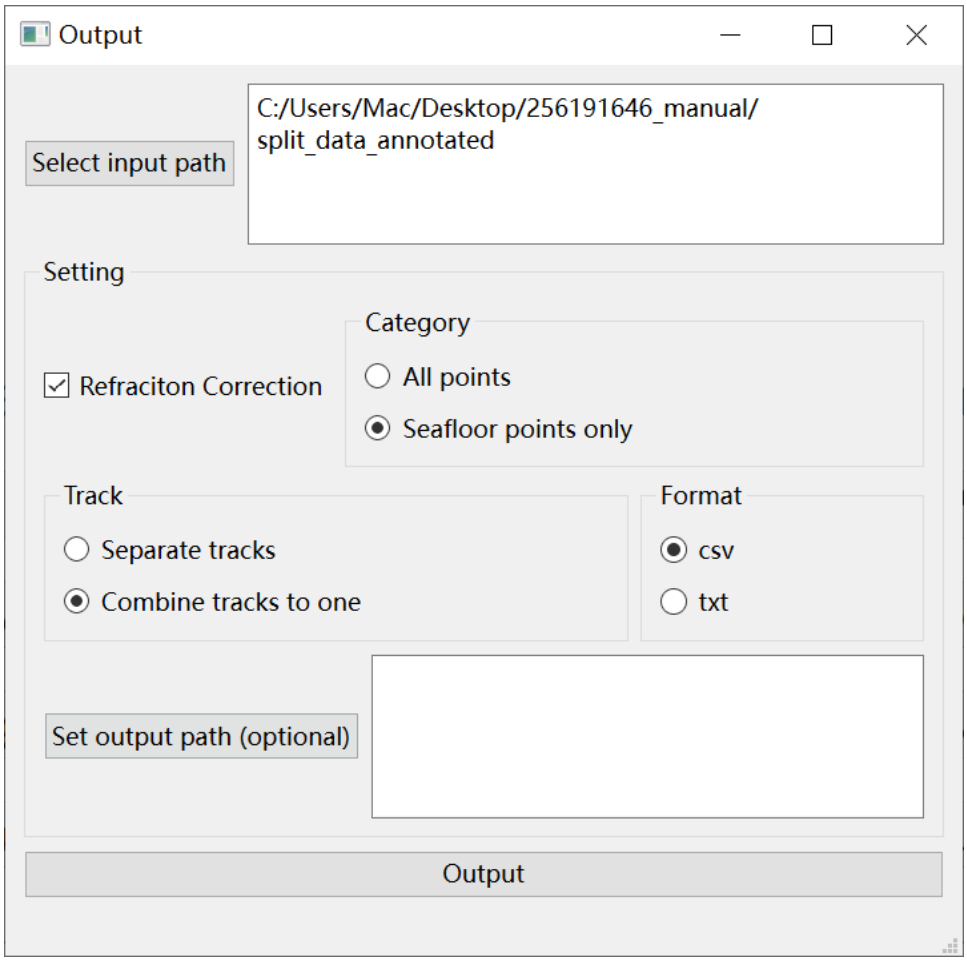

Figure 15: Output Data Window.

The final output file includes the following columns: “x”, “y”, “elev”, "lon”, “lat”, “class”, “signal_conf”, “annot” and “w_elev”. See the description on page 13-14 for what the data in these columns mean. The " annot " column records the annotation made by users, with 0 representing non-seafloor and 1 representing seafloor. When following this workflow, the "w_elev" column represents the mean elevation of the closest water points to a given point if its "annot" column is 1. In addition, selecting the option “combine tracks to one” will create a new 'track' column, indicating the track ID to which each point belongs.

Open-Source Collaboration

SeafloorMapper is an open-source software, which means its source code is freely available to the public. We encourage users to actively engage with the software's development. Here's how you can get involved:

Reporting Issues

If you encounter any problems, or have suggestions for improvements, please consider reporting them. Your feedback is invaluable to us. You can report issues on our GitHub repository: https://github.com/ylin152/SeafloorMapper.git.

Contributing to the Code

We welcome you to contribute to the software's development. You can fork our repository, make improvements, and submit your contributions for review. Together, we can make SeafloorMapper even better.