Seismic Refraction Data Processing Using Refrapy

Ann Mariam Thomas (EarthScope Consortium), adapted from the PickWin guide by A. Parsekian (University of Wyoming)

This guide is intended for users of the Geometrics Geode seismograph system. The following steps show users how to open raw data files and generate images and tabular data, which students can then interpret using the Excel-based workflow provided with GETSI module [Measuring Depth to Bedrock Using Seismic Refraction](mailto:https://serc.carleton.edu/iguana/teaching_materials/seismic/unit2.html) Unit 2.

This guide is similar to A. Parsekian's Seismic Refraction Data Extraction guide but uses Refrapick as an open-source alternative to the commercial software PickWin. Refrapick is one of the two main programs of [Refrapy](https://github.com/viictorjs/Refrapy?tab=readme-ov-file) (Geudes et al., 2022), a Python-based tool for refraction analysis. While it requires a few command-line scripts for installation and launching, Refrapick is a graphical user interface (GUI) similar in appearance and functionality to PickWin.

Refrapick can be installed on MacOS, Windows, and Linux operating systems. Familiarity with running command-line arguments is recommended for this guide. Although Refrapick can be used in Linux systems, our guide was tailored and tested only on Windows and MacOS machines. Linux users can refer to the following webpage for Anaconda installation: <https://docs.anaconda.com/free/anaconda/install/linux/>.

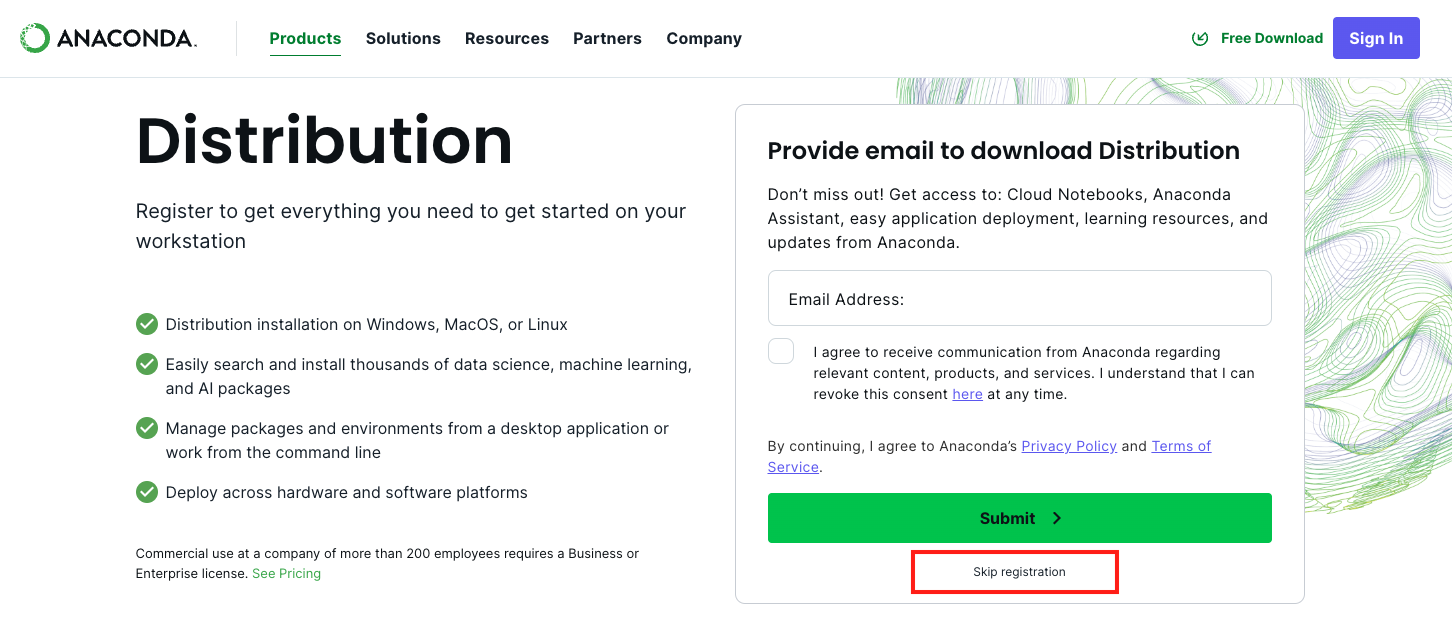
# Have raw data available

During data acquisition, you should have chosen the directory where raw data would be saved. Be sure to download all files related to your data collection from the field computer onto your local computer before returning the field computer. We recommend that the data are saved in the SEG2 format.

# Install Refrapy

* 1. Anaconda is an open-source platform that is used to install Python and the required packages of Refrapy. Install Anaconda by navigating to the webpage: <https://www.anaconda.com/download>. Click “Skip registration” (see Fig. 1) to avoid an email submission and click the green “Download” button (see Fig. 2) to download the appropriate installer for your operating system. This may take 5+ minutes, depending on your system and network.

NOTE: Skip steps a and b if you already have Anaconda or Miniconda installed. Miniconda users, you may need to install a few extra packages (e.g. [scipy](https://scipy.org/install/)) that are preinstalled with Anaconda.



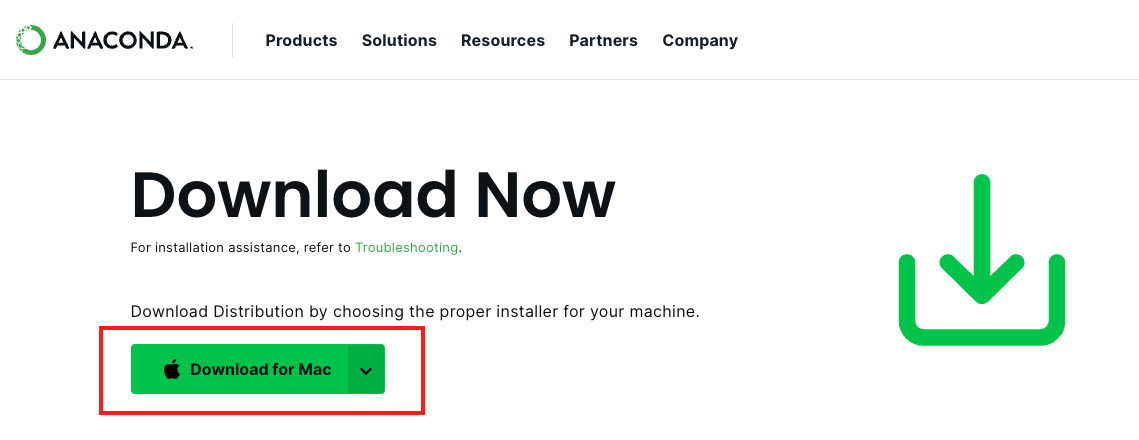
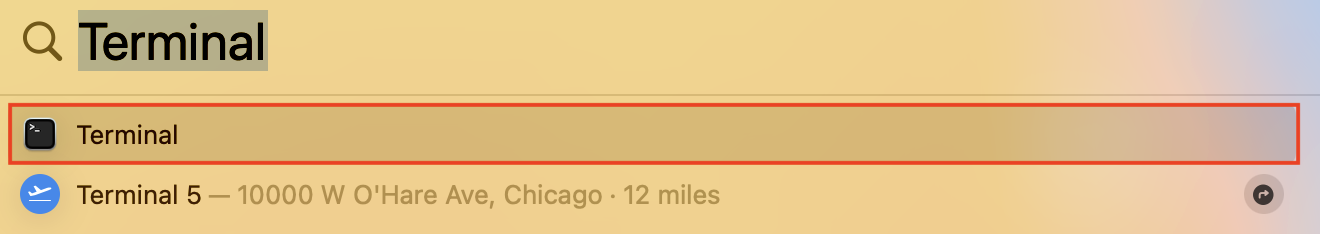


Figure Downloading the Anaconda Installer

* 1. Navigate to your Downloads folder and double-click on the .pkg (MacOS) or .exe (Windows) file that you just downloaded. Follow the instructions on the installer window that appears. We recommend that you keep the default settings.

For a step-by-step guide through the installer prompts, you can refer to: <https://docs.anaconda.com/free/anaconda/install/mac-os/> (for MacOS users) or <https://docs.anaconda.com/free/anaconda/install/windows/> (for Windows Users).

* 1. Verify your Ana/Mini-conda installation by searching for one of the following command-line interfaces (CLI) using your search bar: “Terminal” for MacOS users or “Anaconda Prompt” for Windows users (Fig. 2).



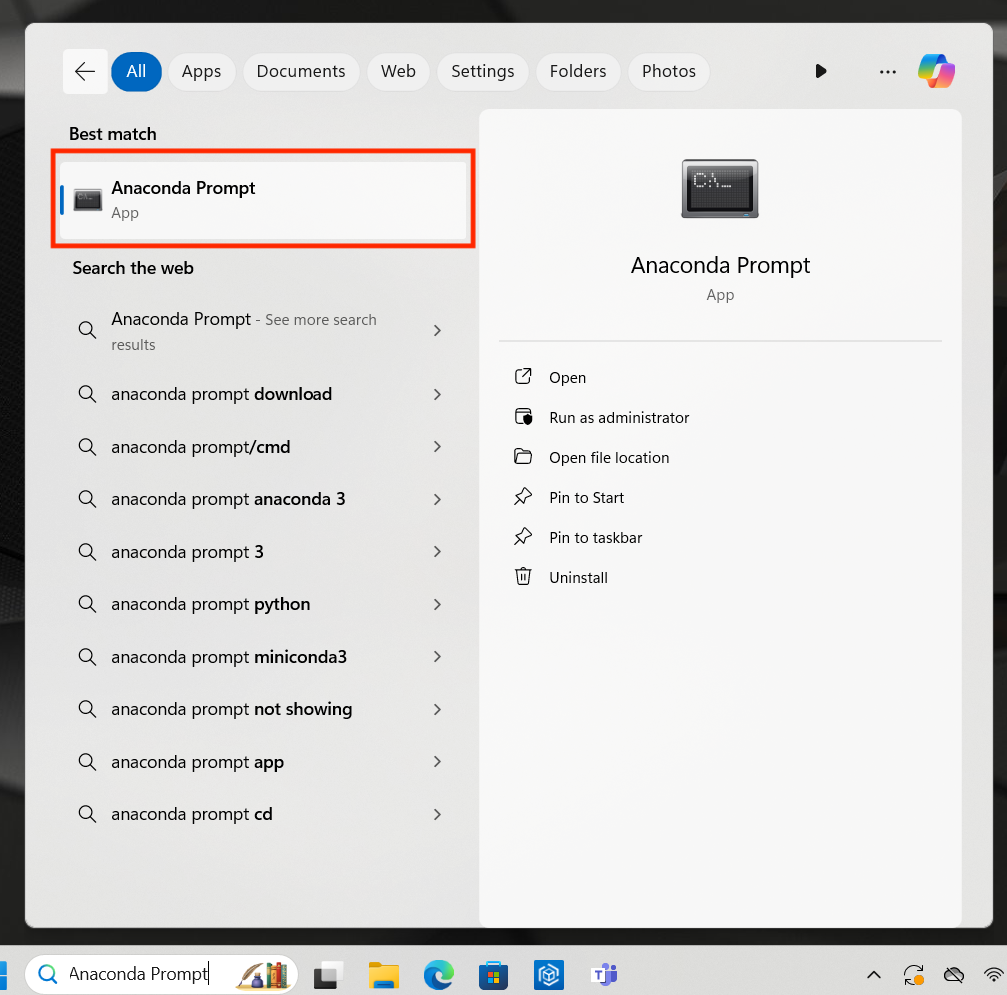


Figure 2 Opening Terminal (MacOS) or Anaconda Prompt (Windows)

* 1. Open your CLI and enter the following command-line prompt: **conda list**

If you see a list of packages and their versions, then you have installed Anaconda successfully :)

* 1. Create a [virtual environment](https://conda.io/projects/conda/en/latest/user-guide/concepts/environments.html) named: “refrapy” with Python 3.8 by running the following prompt in your CLI: **conda create -n refrapy python=3.8**

Enter “y” if asked to “proceed (y,n)?” for this and any of the following command-line arguments.

* 1. Activate your environment by running the prompt: **conda activate refrapy**
  2. Install ObsPy (Beyreuther et al., 2010), a Python package for seismic analysis, by running: **conda install conda-forge::obspy**
  3. Install pmw, a toolkit to build Refrapy’s graphical user interface, by running:

**pip install pmw**

* 1. [Optional step] Install [pyGIMLi](https://www.pygimli.org/index.html) (Rücker et al. 2017), a package used to perform inversion, by running: **conda install -c gimli -c conda-forge pygimli**

NOTE: pyGIMLi is not required for this guide and running Refrapick. It is only needed for Refrainv, an alternative to the commercial software Plotrefa, that can be used to apply time-term and travel-time tomography inversion.

* 1. Navigate to the GitHub page of Refrapy: <https://github.com/viictorjs/Refrapy>. Download the repository by selecting the green “<> Code” button and clicking “Download ZIP:

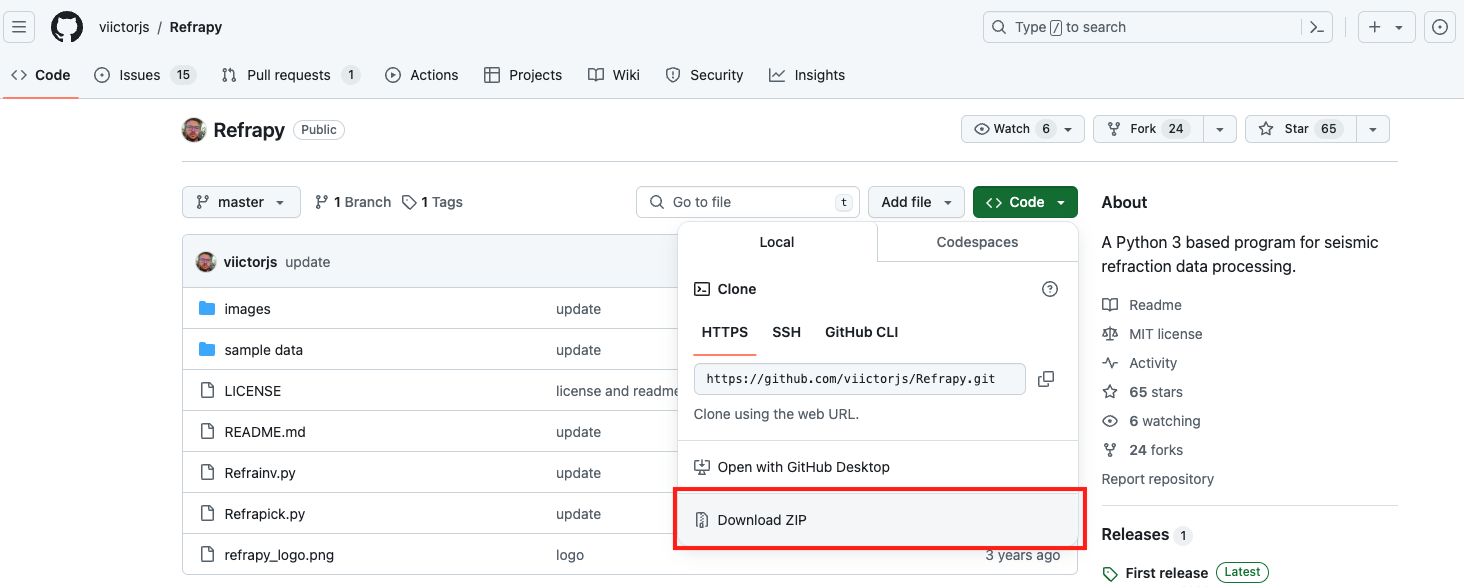


Figure Downloading the Refrapy repository

* 1. Unzip the file you just downloaded and move the “Refrapy-master” folder to your desired directory.

Now, you should have all the tools needed to run Refrapick. If you have completed Section 1 once, you can move directly to Section 2 any time you need to launch Refrapick.

# Launch Refrapick

1. If you haven’t already done so, open your CLI (Terminal for MacOS users or Anaconda Prompt for Windows users) and activate your “refrapy” conda environmentby running: **conda activate refrapy**
2. Navigate to your “Refrapy-master” folder by running the following command and replacing *path/to/Refrapy* with the appropriate directory path: **cd *path/to/Refrapy***
3. Launch Refrapick by running: **python Refrapick.py**

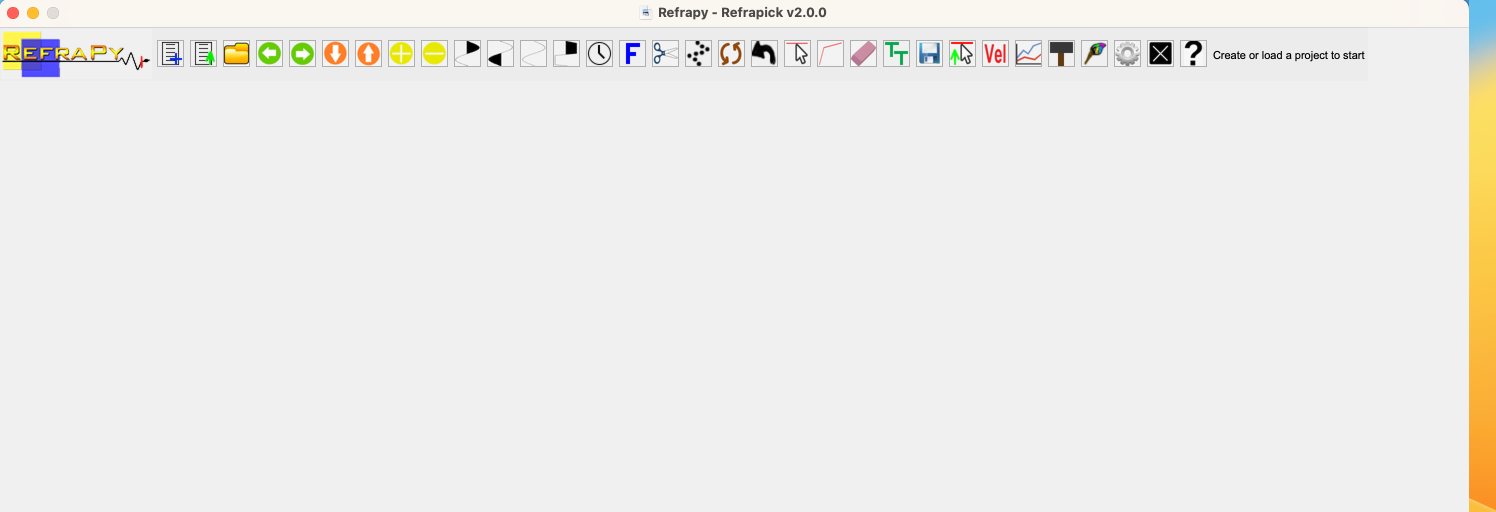
If you see a window that looks like Figure 4, then you have successfully launched Refrapick!

Figure 4 Refrapick GUI

NOTE: If you only see a squished view of the Refrapick GUI on your Mac, double click on the Title bar (the bar containing the title, minimize, and close buttons) to expand your window to the full view.

# Load your measured data

1. Create a new project path by clicking on the leftmost icon in menu bar of Refrapick

Figure Create a new project path

1. Choose the desired directory where you want to create a new project folder that includes empty folders to store waveform data, first-arrival picks, and more. Name your project folder and click OK once it has been successfully created.

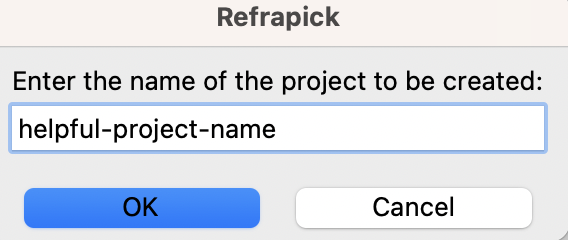


Figure 6 Name your project

1. Using your preferred file management application (Finder or File Explorer), copy your field data (.dat files preferably stored in a SEG2 format) to the “data” folder of your new project folder.

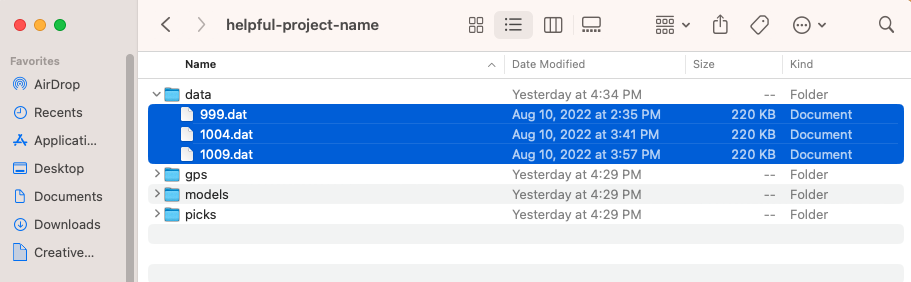


Figure 7 Move waveform files to data folder

1. In the Refrapick GUI, click on the folder icon to open waveform file(s).



Figure 8 Open waveform file(s)

1. You will be automatically promoted to your “data” folder. Select one of your .dat files and click Open.

NOTE: You can also load multiple .dat files and flip through them using the green arrows in the Refrapick menu bar.

After successfully loading your data, your window should look something like Figure 9.

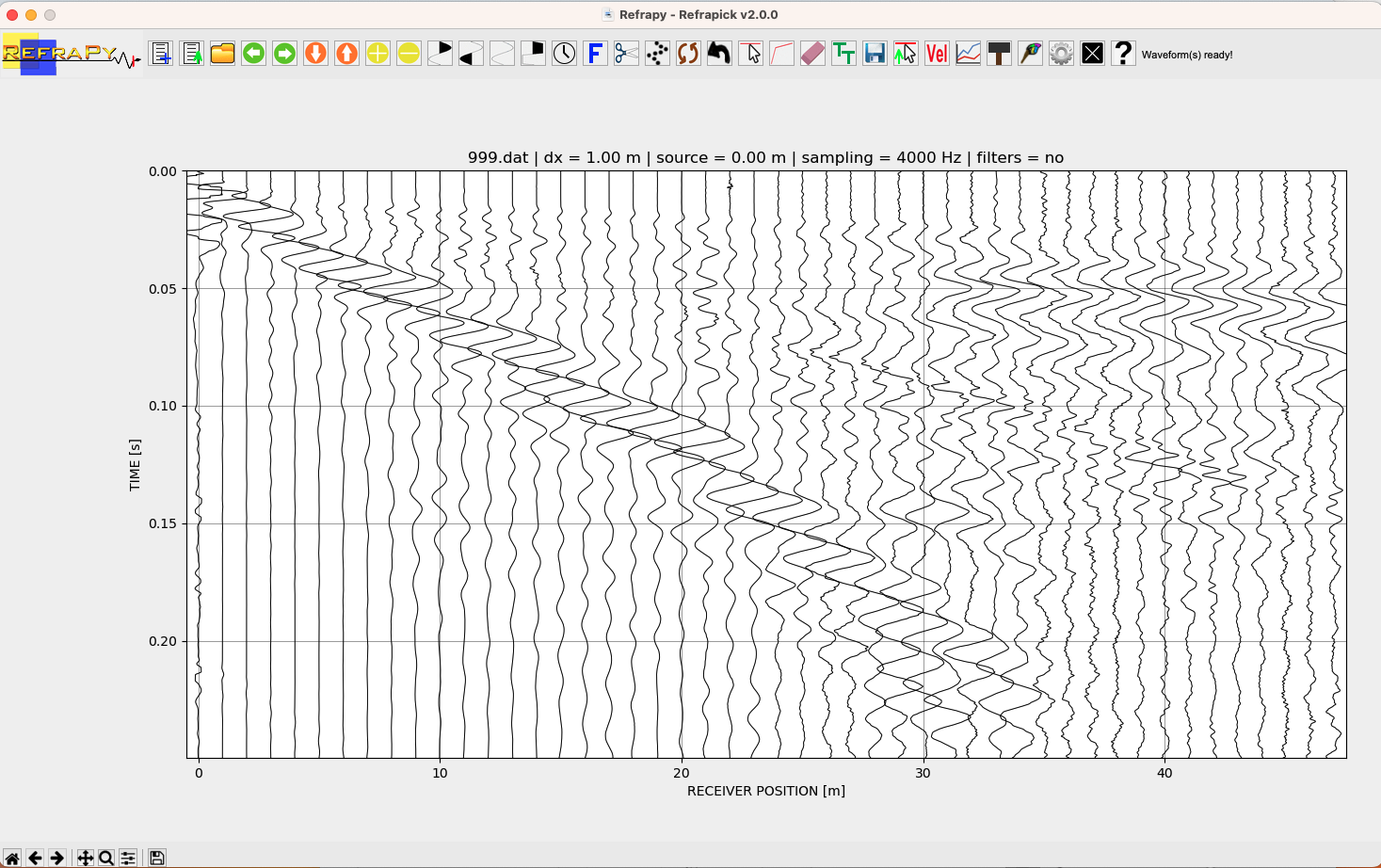


Figure 9 Loaded waveform data

# Adjust data display settings

1. Now, let us improve our display so that we can easily see the first arrival at each geophone trace. Click the two icons boxed in red below. The left boxed icon will shade the positive amplitudes of each trace black. The right boxed icon will clip the traces so that large amplitudes in a trace do not obscure the view of nearby traces.

Refrapick menu bar with two icons boxed in red

Figure 10 Shade and clip traces

Now, your window should look something like display in Figure 11.

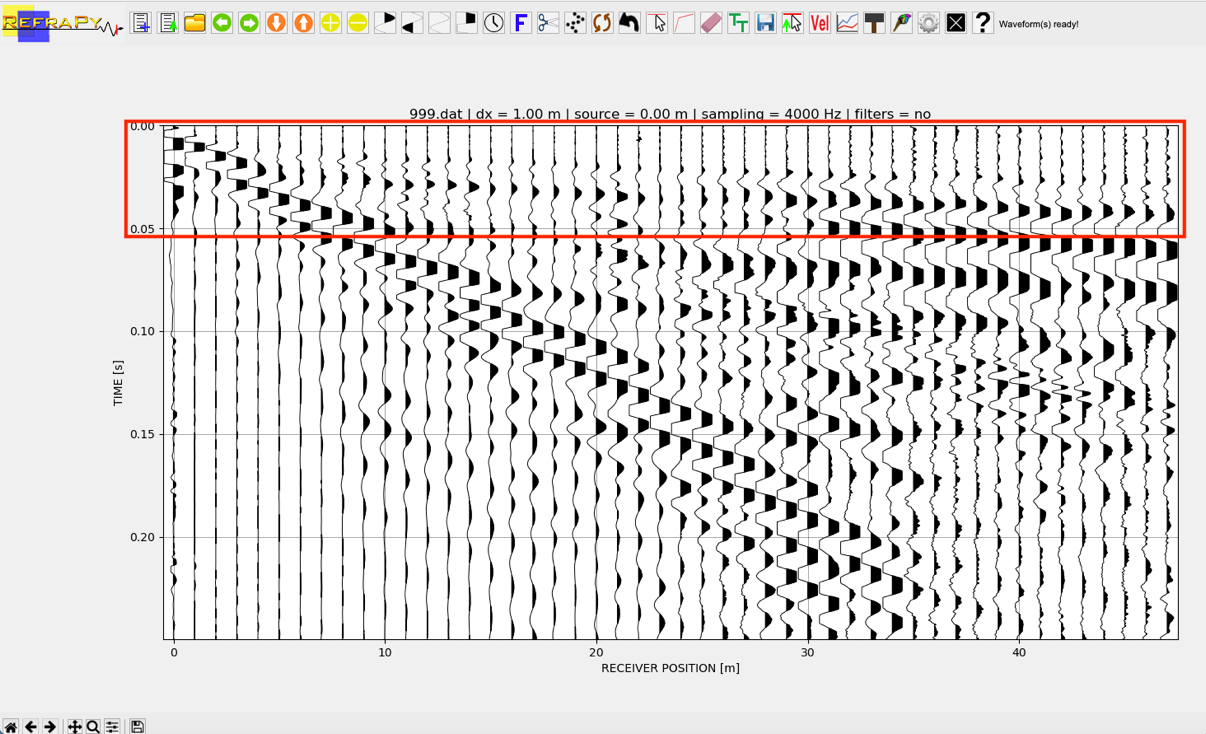


Figure 11 Waveform data with desired signals outlined by a red box

1. We can see the whole data set now but the signals we want to analyze are just in the area boxed in red in Figure 11, within the first 50 msec of each time series. Let us zoom into this area and amplify the arrivals by using the following icons

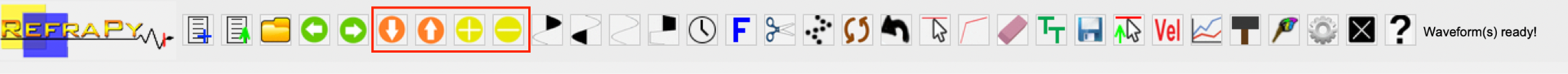


Figure 12 Adjust axis limits and gain

**Orange arrows:** Decrease/increase time axis limits

**Yellow plus/minus signs**: Increase/decrease scale gain

Decrease the time axis limits and adjust the gain until you have a display that looks like Figure 13, with the refraction first arrivals clearly following a linear trend dipping to the right (this will vary depending on the shot point of the dataset you chose. In this case, the shot point is at geophone 1).

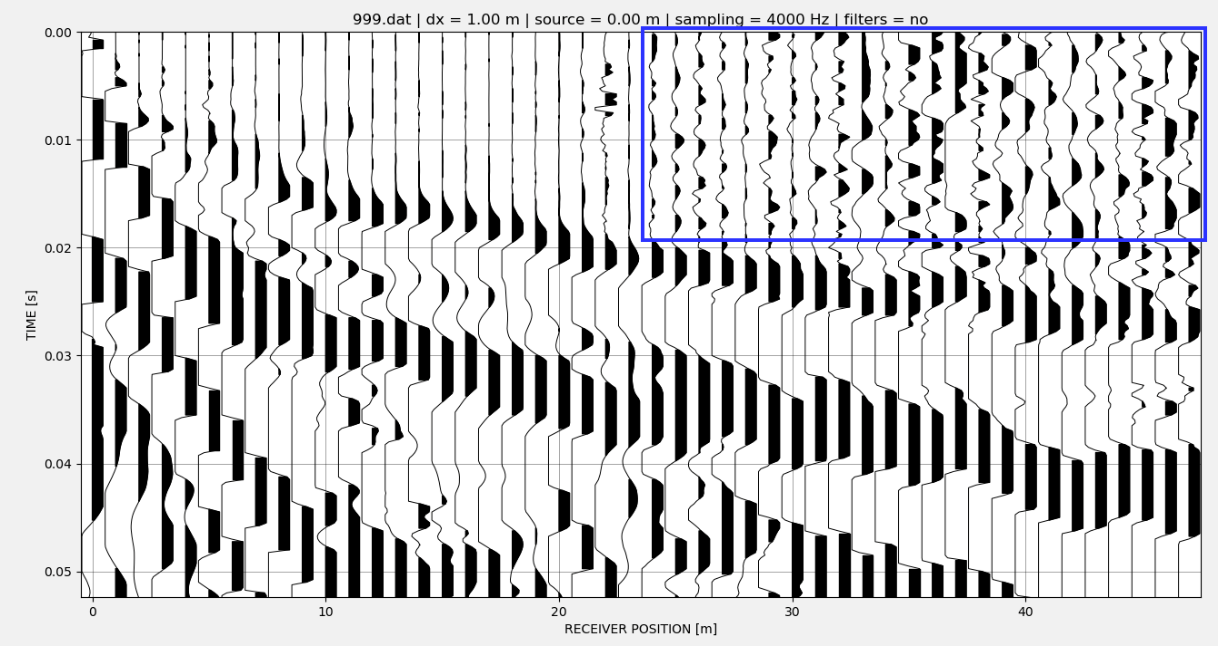


Figure 13 Zoomed-in view of data with ambient noise outlined in a blue box.

Inside the blue outlined area in Figure 13, you can see that the ambient noise is amplified. We don’t need to address this because it won’t impact our analysis too much. Just remember to ignore the “random” noise in this area of the seismogram.

# Analyze

1. At this point, you can take a screenshot of your display and print the seismogram to provide to students for manual picking (picking = digitizing the travel time-offset data points)
2. Alternatively, you can use the Pick mode to pick your first arrivals by selecting following icon.



Figure 14 Enable Pick mode

TEACHING TIP: You can do this for the students, or you could ask them to use Refrapy themselves and follow this guide. Since Refrapy installation and launching may take some time and troubleshooting, we recommend Refrapy for graduate students or upper-level undergraduates with some programming experience. There is also an opportunity, outside the scope of this guide, for graduate students to use Refrainv and engage with inversion methods to create a more realistic subsurface velocity model.

Once you have enabled the Pick mode, simply use the mouse to click on the location where you see the first change in waveform amplitude along each trace. If you want to move your pick, click on a new location and the pick will automatically change. You can connect a line through each pick by selecting the following boxed icon.



Figure Connect first arrivals with a red line

Figure 16 shows an example of the picked arrivals.

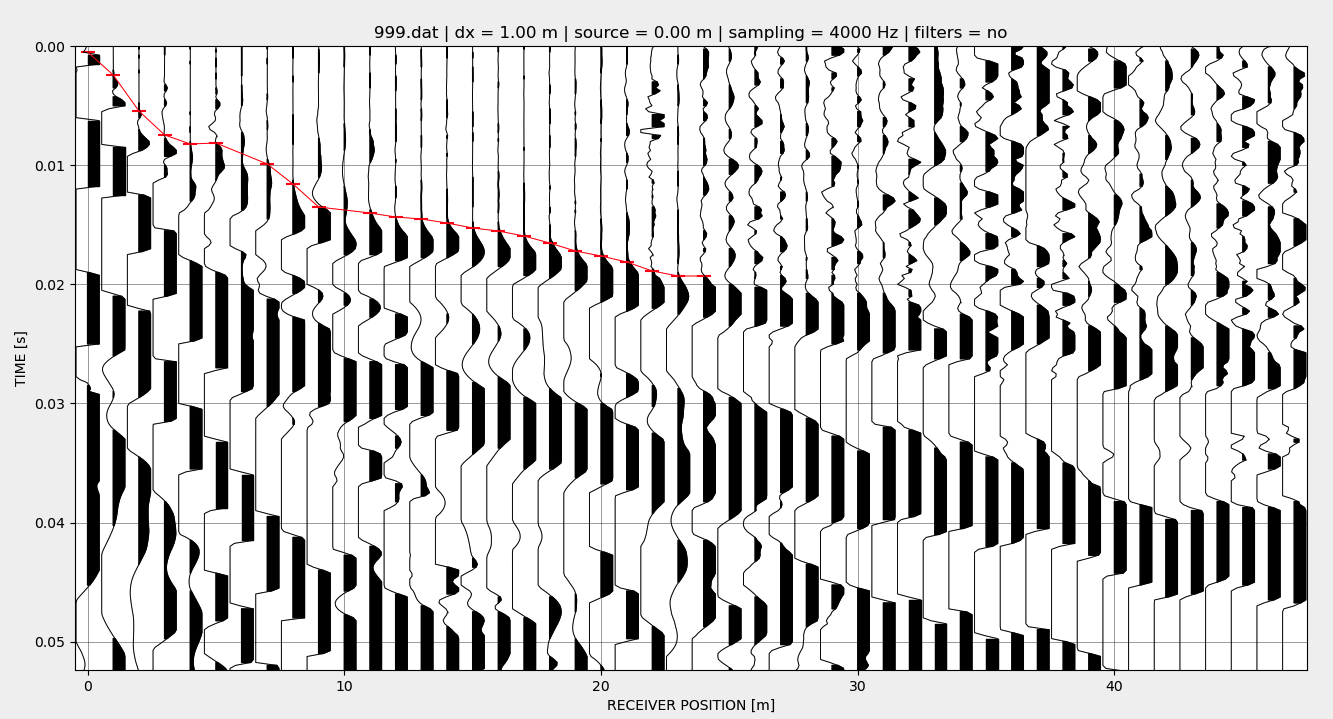


Figure 16 Waveform data with picked arrivals

1. CLASS DISCUSSION TOPIC: There is a lot of judgment in the role of seismic interpretation. Sometimes the waveforms will not be totally clean, but you can get a sense for which is the right place to pick based on the adjacent traces. You can see that we did not pick arrivals on some traces. On the right side of the image where the noise starts to encroach on the data, we stopped picking because it was too hard to distinguish between signal and noise. This is fine for our purposes where we just want to calculate a 2-layer velocity model.

[Optional step] To save these picks, select the following boxed icon (you can say No to loading a topography file) and save the file with your preferred name.

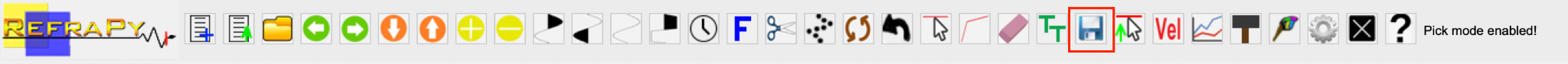


Figure 17 Save your picks

1. One way to make it more clear which parts of the seismogram to pick is to draw lines on the image that might represent refractions from different layers. If possible, it is best to have each refraction arrival line defined by at least four traces. Here are examples of two ways that this dataset might be interpreted.

*Three-layer structure*

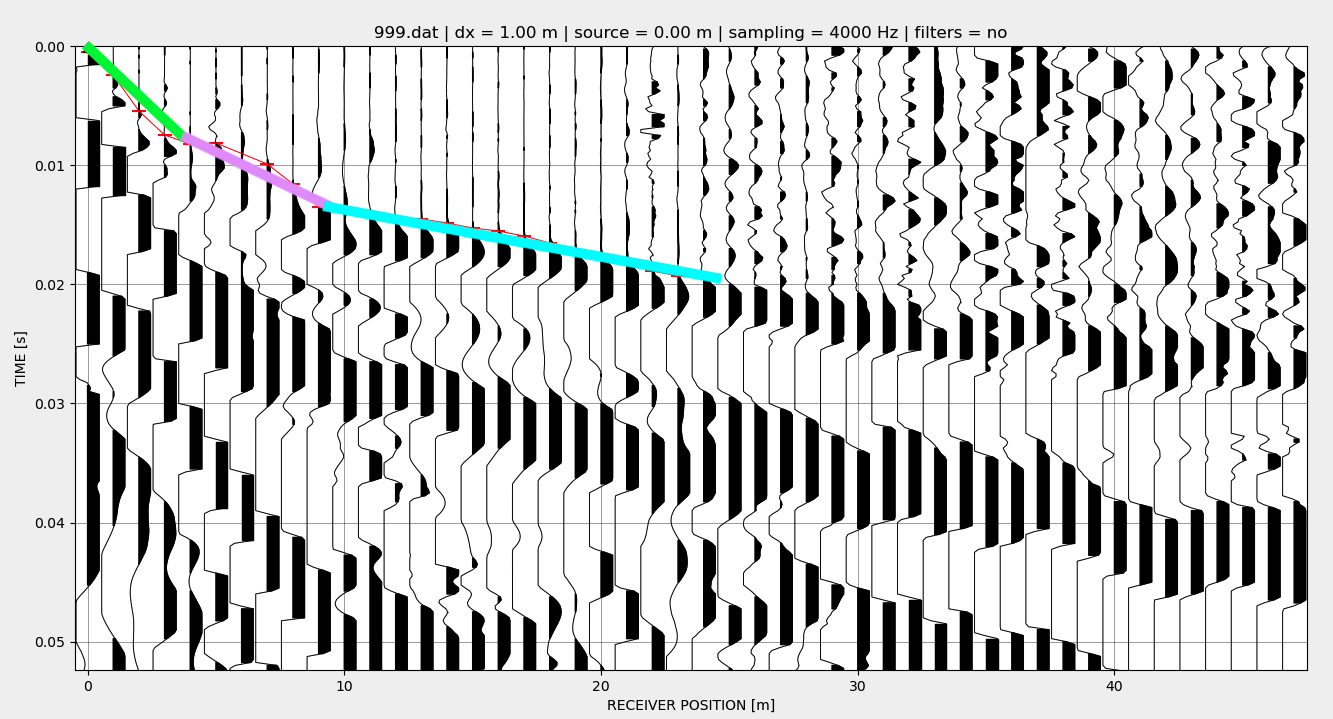


Figure 18 Waveform data with annotated lines corresponding to three velocity layers

*Two Layer Structure:*

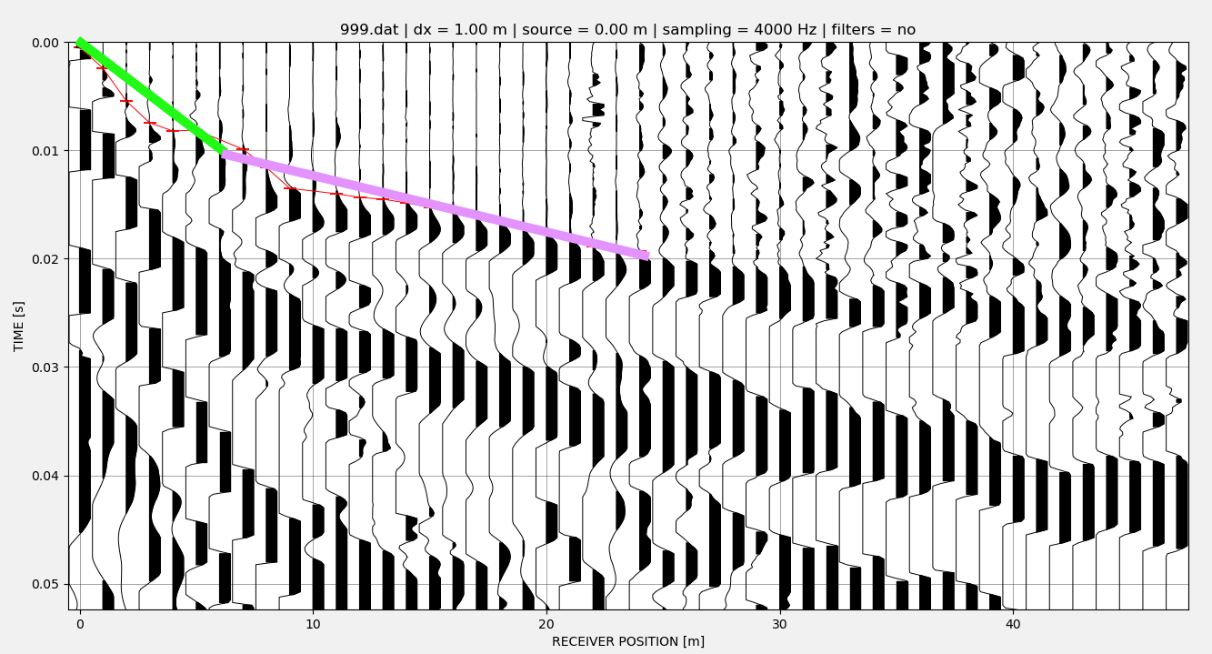


Figure 19 Waveform data with annotated lines corresponding to two velocity layers

When interpreting the data, we need to have a conceptual model in mind for what structure we will try to recover. In some cases, there may be outside data (e.g., boreholes) that could directly tell us if a 2-layer or 3-layer case is most appropriate. Otherwise, we need to make this judgment based on the data alone. Generally, we want to choose the simplest possible structure that appears to be appropriate for the data. Specifically, we want to identify the locations of sharp changes in the slope of the first arrivals that might indicate refraction to progressively faster velocity layers as we go deeper. In this example data, you can see that the change in slope as you move toward larger distances is gradual—sharp changes are not easy to identify. This is not uncommon and often occurs when the geology transitions gradually between layers. Nonetheless, for our analysis approach to work, we must define layers. As you can see by the thick lines overlaid on the seismogram for the 3-layer case, this seems to do a pretty good job of falling along the first arrival threshold. In contrast, the thick lines drawn on the 2-layer case don’t seem to fall right on the threshold of first arrivals quite as well. However, 2 layers is simpler than 3 layers, so that might make the simpler option more desirable. (the thick lines shown are purely a visual approximation – you may choose to shift the locations of these lines however you want for your dataset).

1. Example: Shot in the middle of the line



Figure 20 Waveform data with short location at 25 m

In this raw seismogram, you can see that the shot/source location is at 25 m because this is where the first arrivals are closest to 0 ms travel time. Since we have geophones on both sides of the shot, we get two refraction arrivals that we can analyze. In this image you can see that the change in slope is much sharper than in the previous demonstration dataset and this confirms that it would be most appropriate to use a 2-layer model.

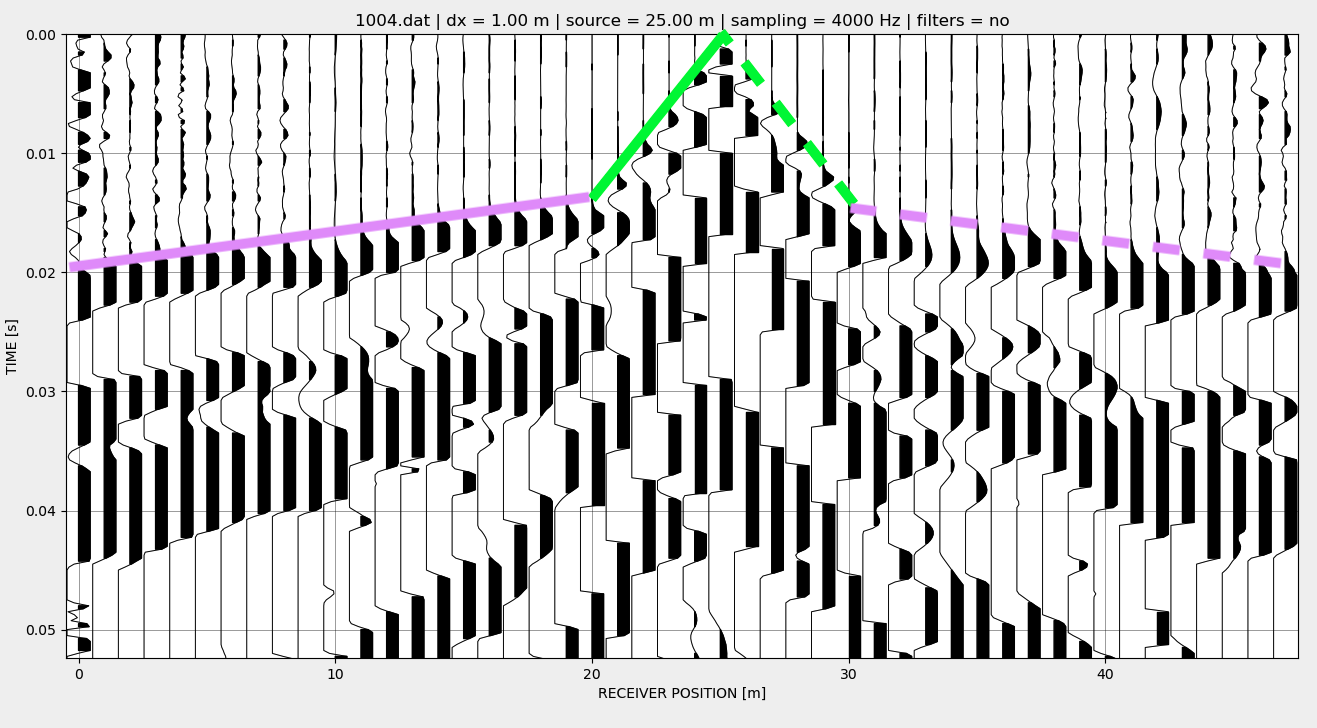


Figure 21 Waveform data with annotated lines corresponding to two velocity layers

Here is the same raw seismogram as Figure 20, but with the two refraction arrivals highlighted with thick lines. In a case like this, the solid line refractions should be analyzed separately from the dashed line refractions, but both may be used to estimate velocity and layer thickness.

1. Example: Shot on the far end of the line

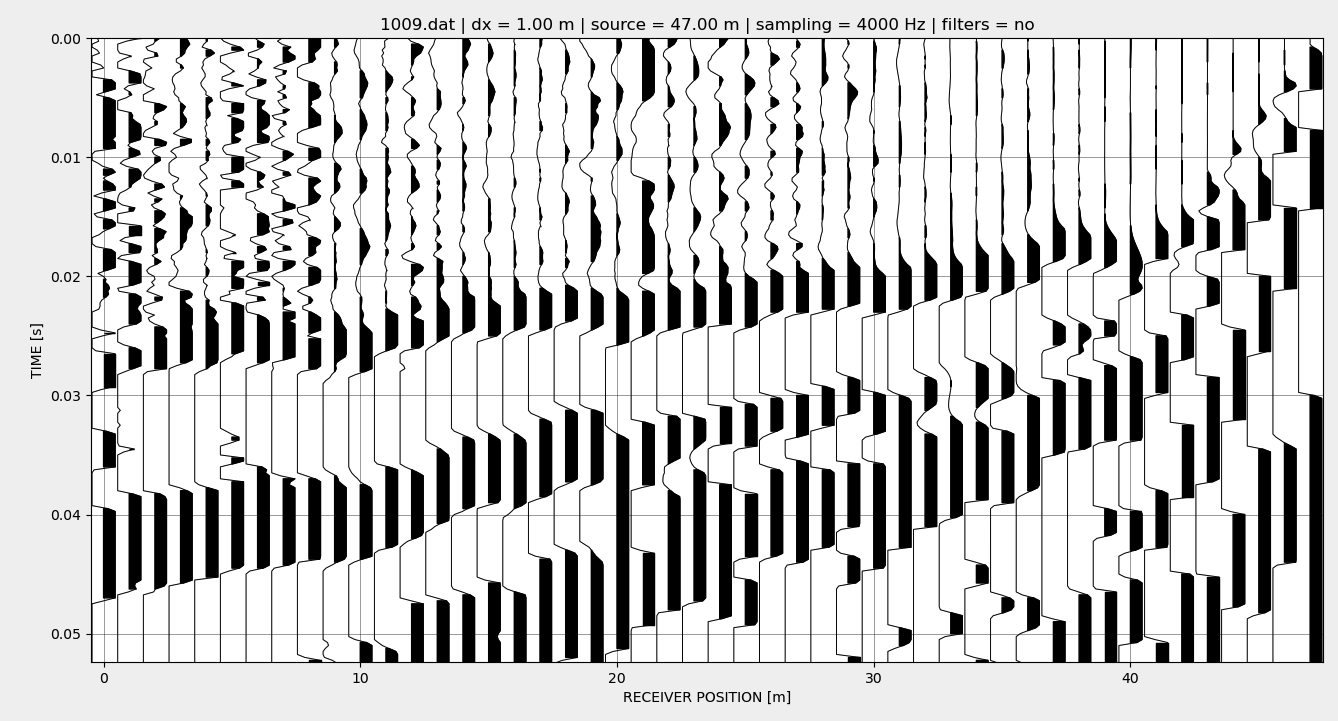


Figure 22 Waveform data with a shot location at 47 m

Here is an example where the shot point is at the far end of the line from the seismograph. This looks just like the example used above, except in this image the change in slope of the first arrivals is much sharper.

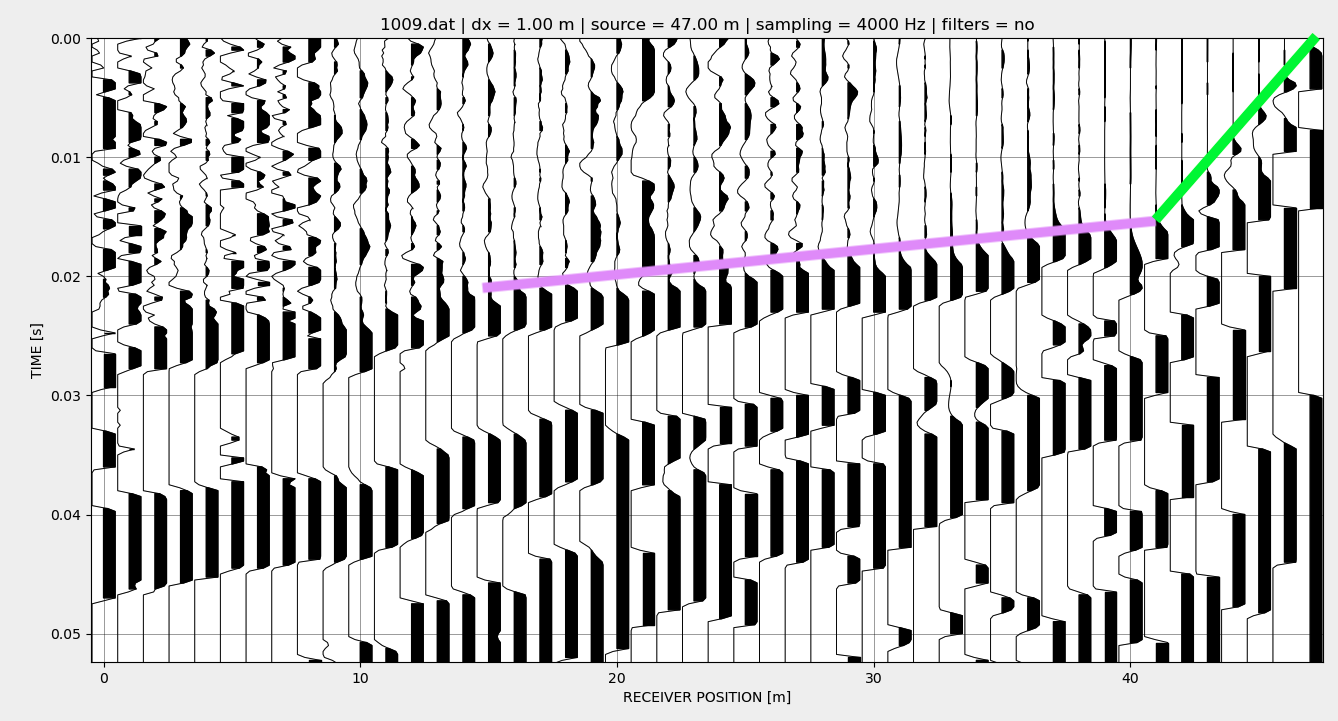
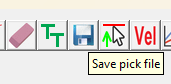


Figure 23 Waveform data with annotated lines corresponding to two velocity layers

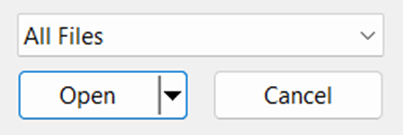
Here are annotated lines illustrating that a 2-layer model is appropriate in this case. This data can be analyzed the same way as the example dataset, but the x-direction values should be reversed so the shot is at 0 and the geophones are increasing positive values.

TEACHING TIP: If you are discussing data quality, errors, and/or uncertainty in your classroom, a good point of discussion is how different people will pick a given dataset and how each set of picks will result in a slightly different set of velocity/thickness results. This can be turned into a synthesis exercise by calculating summary statistics (mean, median, standard deviation) based on the thickness/velocity results from several students or groups within the class.

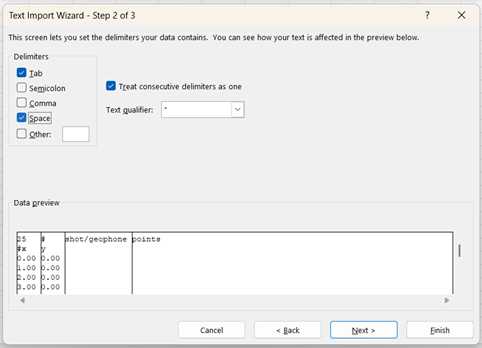
1. **Exporting Data and Working in Excel**
   1. To save the distance and time data for your picks, select the Save pick file button, outline in red on image below. You can select No when asked to load a topography file. The data will be saved as a .sgt file.



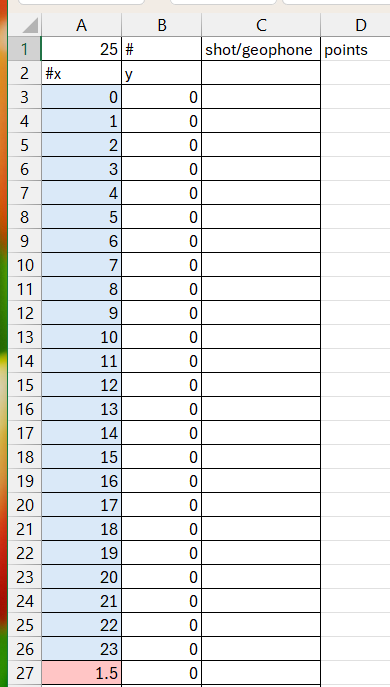
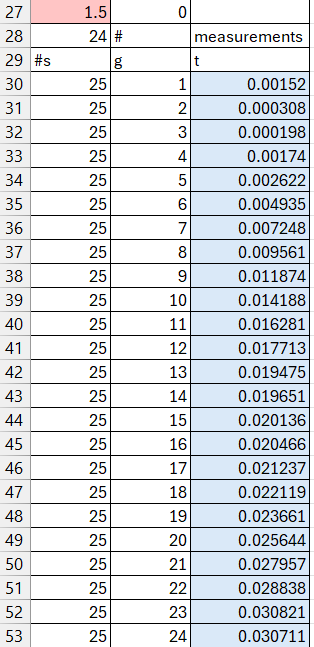
* 1. To open this file type in excel:
     1. Open excel, select File -> Open. In the browser, make sure you choose All Files instead of All Excel Files in order to see the .vs files. Select your file.
     2. In the Text Import Wizard pop up, on page one, make sure Delimited is checked and then hit Next.
     3. On page two, make sure Space is checked, the select Finish.



A screenshot of a computer

AI-generated content may be incorrect.

* 1. Once in Excel, your data should look similar to the example below. The first two lines and the third column can be ignored. The red box, in this example column A row 27, is the position of the shot. The blue boxes contain the positions of the geophones (column A) and the arrival times (column B) for the picks. The top half of the spread sheet (rows 3-26) gives the positions of the geophones, while the bottom half (rows 30-53) labels them by number, so be careful that you are using the correct set of data. This difference will be more obvious if you are using a spacing other than 1 m. This data can be used to replicate the excel exercises in the module or used with your own analysis method.

# Connection to GETSI Curricular Materials

The “2-layer seismic refraction exercise” XLSX workbook provided in the Seismic refraction Unit 2 can be either used directly or adapted to analyze the travel time/offset data extracted from the raw seismograms.

<https://serc.carleton.edu/iguana/teaching_materials/seismic/unit2.html>

# References

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., and Wassermann, J., 2010.

ObsPy: A Python Toolbox for Seismology, Seismological Research Letters, 81(3), 530-533. doi: 10.1785/gssrl.81.3.530.

Guedes, V.J.C.B., Maciel, S.T.R., Rocha, M.P., 2022. Refrapy: A Python program for seismic

refraction data analysis, Computers and Geosciences, 159, 105020. doi: 10.1016/j.cageo.2021.105020.

Rücker, C., Günther, T., Wagner, F.M., 2017. pyGIMLi: An open-source library for modelling and

inversion in geophysics, Computers and Geosciences, 109, 106-123. doi: 10.1016/j.cageo.2017.07.011.