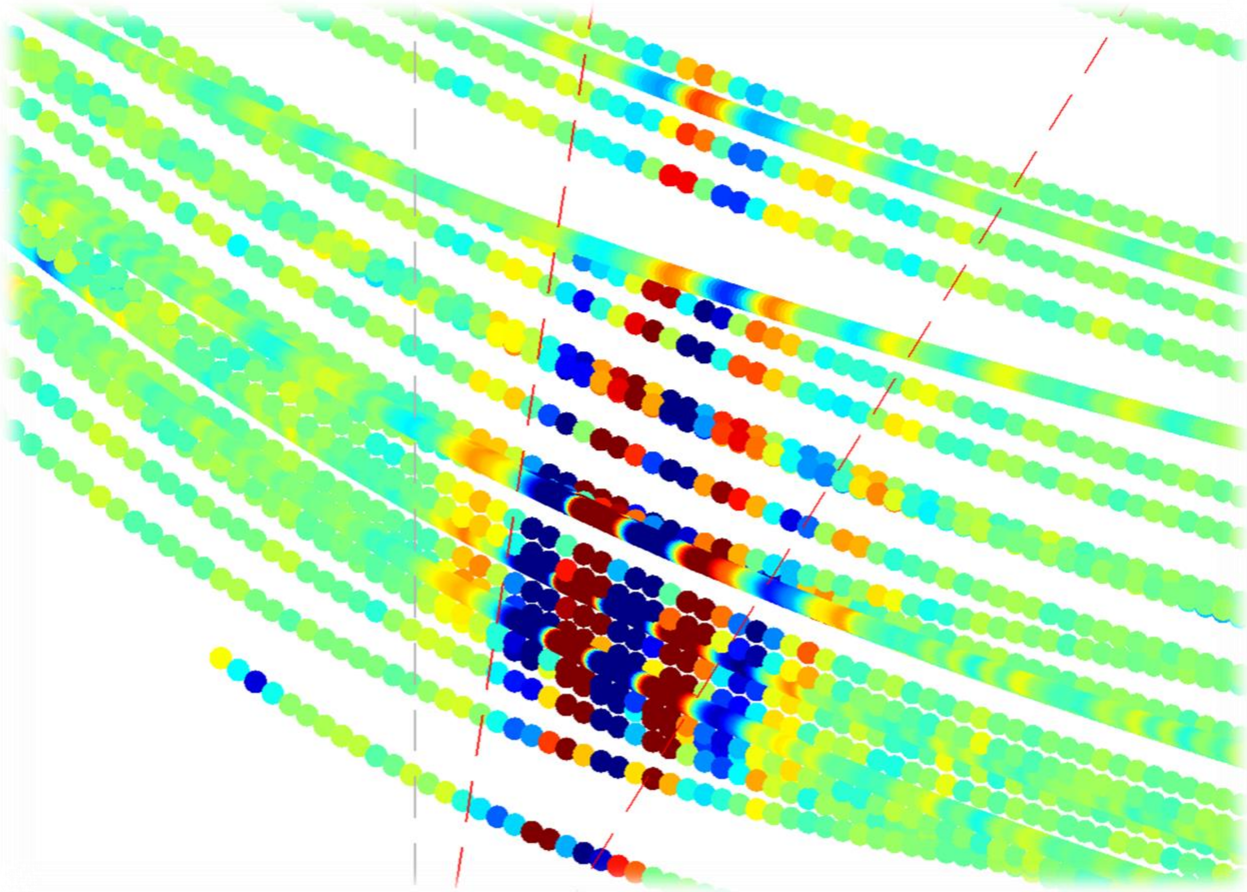
Tutorial Manta-TEC

Author: Fabio Manta

Date: July-2021



INTRODUCTION

This step-by-step tutorial shows how to process RINEX files with our toolkit MANTA-TEC to obtain several parameters useful to study the ionospheric TEC perturbations induced by several type of sources (e.g earthquakes, tsunami, volcanic eruption, etc.). This tool can be also used by scientists which are simply interested in the dynamic of the ionosphere. This toolkit is fully operational on MATLAB and in its *beta* version allows to process both GPS and GLONASS data. To guide the user through the several functions provided by MANTA-TEC, we use an example provided with the package. The example is related to the 2015 magnitude Mw 8.3 Illapel (Chile) earthquake (https://earthquake.usgs.gov/earthquakes/eventpage/us20003k7a/executive).

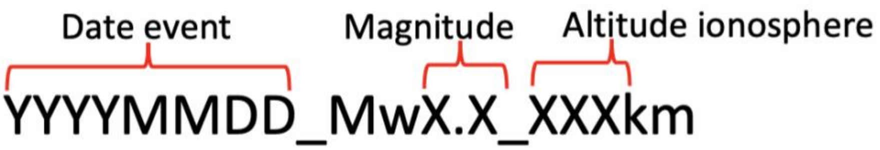
Note that before operating the toolkit it is necessary to install some MATLAB toolboxes as listed in the README.txt.

MATERIALS

We provide 3 RINEX Observations file related to stations AEDA, CMPN and OVLL (**aeda2590.15o**; **cmpn2590.15o; ovll2590.15o**). It is important that the files are located within the folder "input" in a subfolder named as follow:

20150916_Mw8.3_300km

The structure of the subfolder name is important in order to MANTA-TEC to work. The name structure is composed by the date of the event to analyze, the magnitude of the event, and the altitude of the maximum ionization layer of the atmosphere as here depicted:



The input folder also contains other 2 files necessary to corrected for the hardware induced differential delay biases. The IONEX file named **CODG2590.15I** is necessary to correct for the satellite transmitter group delay (TGD), While the SP3 file named **igs18623.sp3** is necessary to correct for the receiver inter-frequency biases (IFB). Now that our input directory is set with all the necessary files, we can move on to process the data.

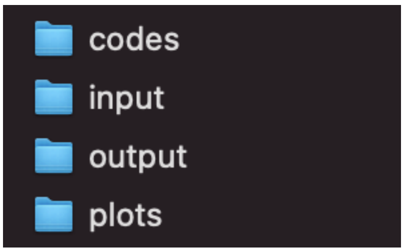
DATA PROCESSING

For the purpose of this example, we provide together with the source codes, a script named "**Test_routine.m**", which after setting some variables related to the event to analyzed, then it calls the function MANTA-TEC, and subsequently produce some useful plots to facilitate the visualization of the results.

As mentioned above, the first step consists of setting some parameter related to the event we would like to analyze, such as coordinates of the epicenter, date and time of the event. At line 28 to line 29, we can set the range of frequency of the Butterworth filter necessary to obtain the filtered TEC. Also, in line 30 it is possible to set the elevation cut-off.

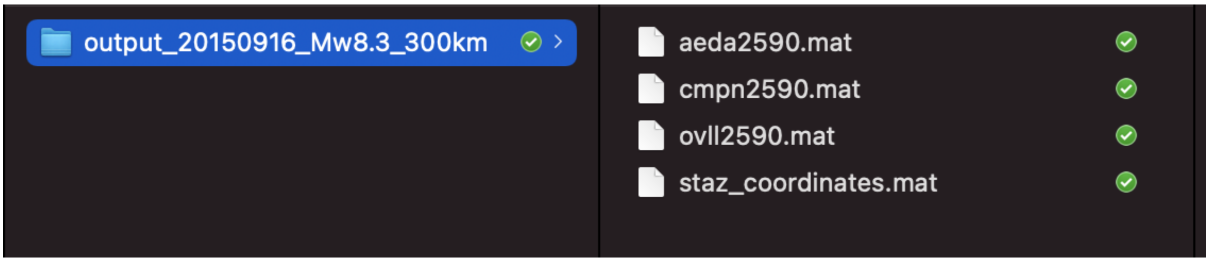
```
23      % Set up Directories and Event info
24
25 -    global staz; % list of GNSS stations
26 -    defaultTimeInterval = 30; %aquisition period in seconds
27      % Frequency range for the BUtterworth filter
28 -    lowFreq = 0.001;% Hz
29 -    highFreq = 0.010;%Hz;
30 -    ele_cutoff = 15; % Elevation cut-off in degrees from the horizon
31
32 -    evntName = 'Mw8.3'; %'sinabung' or 'Mw7.0'
33 -    dateString = '20150916';
34 -    lat_epi = -31.57; %Latitude epicenter;
35 -    lon_epi = -71.654; %Longitude epicenter;
36 -    altitutePPI = 300; %ionosphere altitude in km
37 -    alt_ion = altitutePPI *1000; %ionosphere altitude in meters
38 -    Tmin= 22; Tmax= 24;% Time range to be analyze in hours (UTC time)
39 -    SITE_COORD = []; %coordinates of the GNSS stations
40 -    EventTime = datenum(2015,09,16,22,54,32); % UTC time
41 -    [Year,Month,Day,Hours,Minutes,Seconds]= datevec(EventTime);
42 -    Year = num2str(Year);
43 -    EventTimeHours = Hours+Minutes/60+Seconds/3600;
44      % The following variables are necessary to set the limits of the xaxis in
45      % the plots
46 -    startTime = EventTimeHours-1;% in UTC time
47 -    endTime = 24;%
48
49 -    root_direct = '/Users/fabio/Dropbox/MANTA-TEC_toolkit';
50 -    addpath(root_direct)
```

It is very important to set the variable `root_direct` where all the subfolders are located. The general structure of the directory is presented as follow:

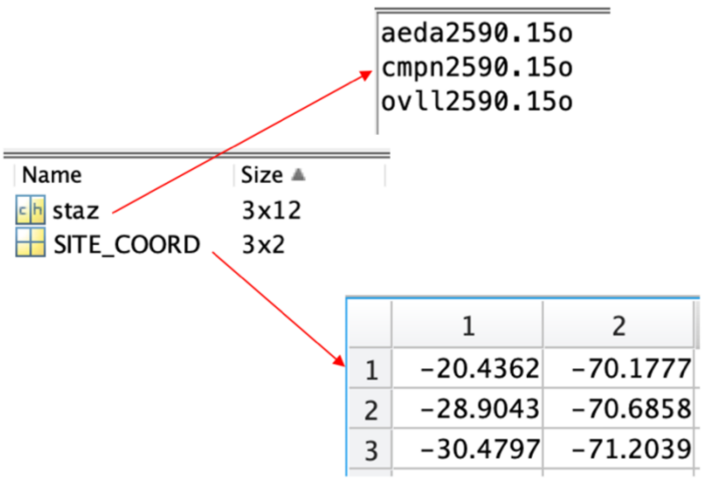


The folder named `codes` contains all the scripts necessary to run **`MANTA_TEC.m`**. The `input` folder, as mentioned above, is the folder where we need to store the RINEX files and the files necessary to corrected for the TEC for the different instrumental biases (IONEX and SP3). The folder `output` contains the processed data (see NEXT SECTION), and finally the folder `plots` contains some key images that can be useful to represent the results. Note that as result of the execution of the first block of the script **`Test_routine.m`**, subfolders with the same name of the event will be automatically created within the folders `output` and `plots` (e.g `output/output_20150916_Mw8.3_300km` and `plots/plots_20150916_Mw8.3_300km`).

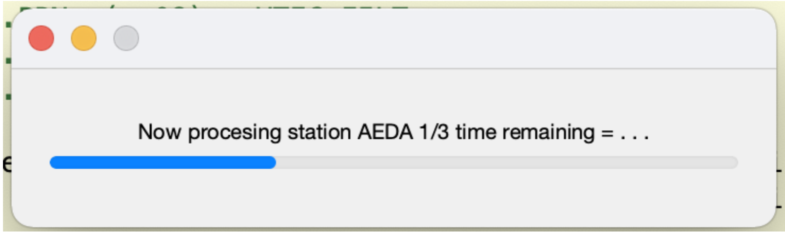
After all the parameters are set, and the directories are ready, the next block calls the main function **`MANTA_TEC.m`** which will return in the `output` directory (`output/output_20150916_Mw8.3_300km`) the processed results. The results are presented with one file `.mat` for each station analysed. For example, the results of the 2015 Illapel earthquake will look like this:



Together with the results files the `output` directory contains also another file named `staz_coordinates.mat` which contains the list of the stations analyzed (`staz`) with their coordinates (`SITE_COORD`).



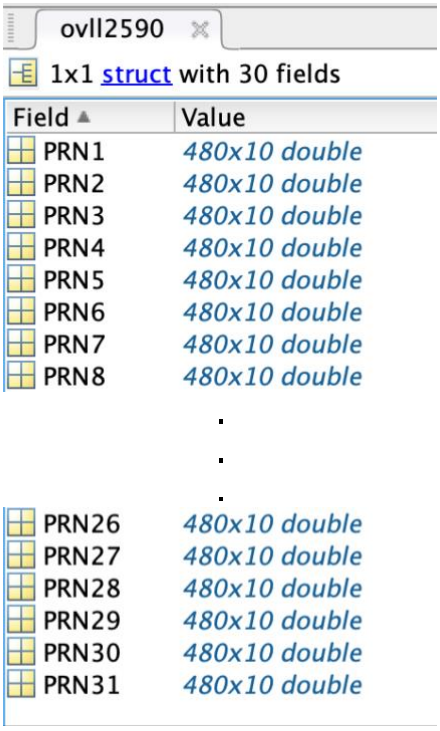
While MANTA_TEC is running a progress bar will pop-up in the screen indicating the name of the station currently processed, the number of the station over the total number, and the hypothetical total processing time.



If the processing is concluded without any issues, we should be able to see in the `output` folder the same number of files corresponding to the RINEX Observations files stored in the `input` folder. Like for RINEX Observations files, also the name of the output is constituted by the name of the station plus the Day Of the Year (DOY) and extension `.mat` (e.g. `ovll2590.mat; cmpn2590.mat; aeda2590.mat`).

STRUCTURE OF THE OUTPUT FILES

Each output file is a MATLAB structure array which contains a number of fields *n*, equal to the number of GPS satellites available at the time. The name of each field is constituted by the acronym PRN (pseudo-random noise) followed by the number of the corresponding satellite. As we can see from the figure below, each field contains an array with a number of rows equal to the number of points in the time series, and a number of columns which corresponds to the number of parameters calculated (in this version of the code is set to be 10). For instance, for the 2015 Illapel earthquake, considering a sampling rate of 15 seconds and a 2-hour window (22:00 to 24:00 UTC), we will obtain array with 480 rows and 10 columns.



According to this structure, if we want to access to the results related to the station OVLL with respect to the GPS satellite number 2, we have to type in the command window the following command:

```
>> ovll2590.PRN2
```

As mention above for each station-satellite pairs **`MANTA_TEC.m`** calculate a total of 10 parameters represented by 10 column vectors.

```
nameDOY.PRNxx(:,01)  = Time
nameDOY.PRNxx(:,02)  = LAT_IPP
nameDOY.PRNxx(:,03) = LON_IPP
nameDOY.PRNxx(:,04) = ELE
nameDOY.PRNxx(:,05) = AZIM
nameDOY.PRNxx(:,06) = EFM
nameDOY.PRNxx(:,07) = STEC
nameDOY.PRNxx(:,08) = VTEC_FILT
nameDOY.PRNxx(:,09) = STEC_FILT
nameDOY.PRNxx(:,10) = DISTANCE
```

As shown in the figure the parameters vectors are:

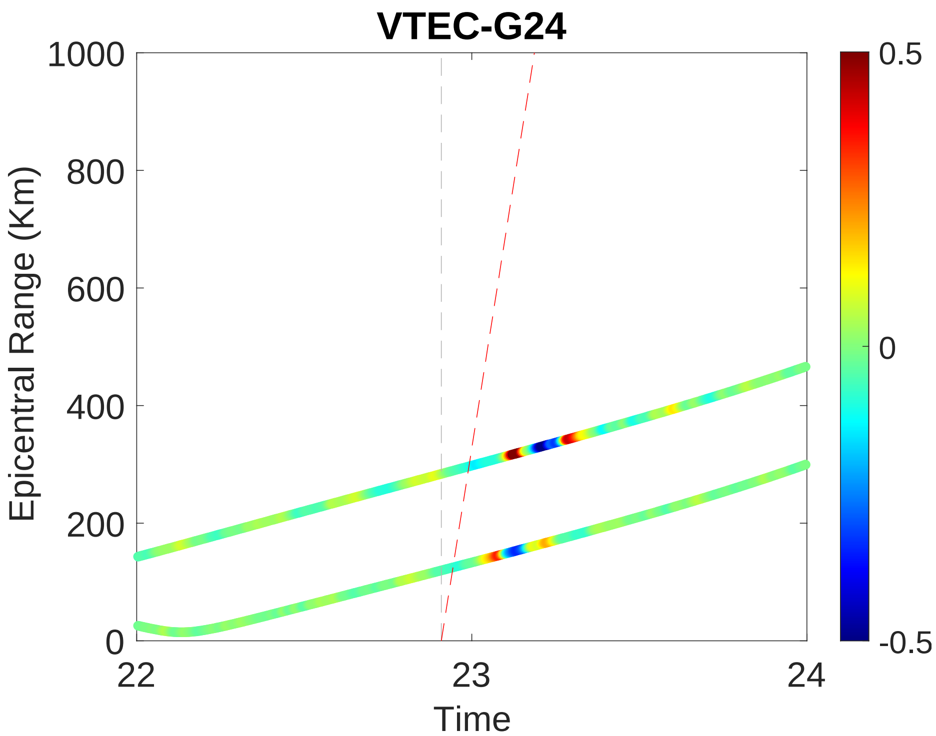
1) The **Time** vector, which contains the time represented in decimal hour of the day in UTC time;
2) The **LAT_IPP** vector, which contains the time dependent latitude in degrees of the Ionospheric Pierce Point (IPP);
3) The **LON_IPP** vector, which contains the time dependent longitude in degrees of the Ionospheric Pierce Point (IPP);
4) The **ELE** vector, which contains the time dependent elevation angle, in degrees, of the IPP with respect to the GNSS station;
5) The **AZIM** vector, which contains the time depended azimuthal angles in degrees of the IPP with respect to the GNSS station;
6) The **EFM** vector,
7) The **STEC** vector, which contains the time dependent absolute Slant TEC (STEC) before filtering;
8) The **VTEC_FILT** vector, which contains the time dependent filtered Vertical TEC (VTEC);
9) The **STEC_FILT** vector, which contains the time dependent filtered STEC;
10) The **DISTANCE** vector, which contains the time dependent distance of the IPP from the epicenter measured in km.

At this point MANTA-TEC have provided the user with all the information necessary to carry on a complete study of ionospheric perturbations generated by a source. In the next section, we provide some examples of useful *data visualization.*

DATA VISUALIZATION

Data visualization gives us a clear idea of what the information means by giving it visual context through maps or graphs. This makes the data more natural for the human mind to comprehend and therefore makes it easier to identify trends, patterns, and outliers within large data sets.

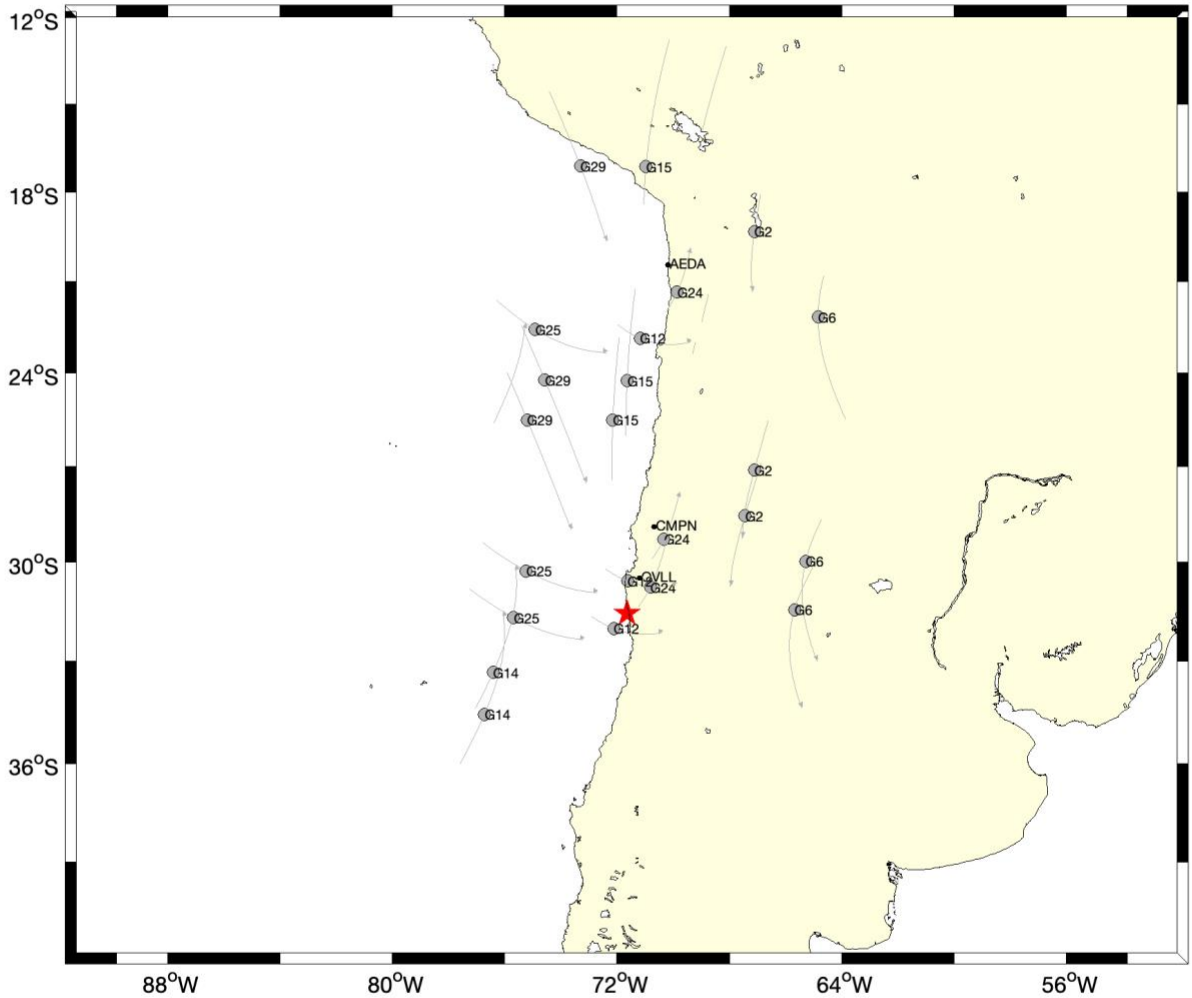
For these reasons, the last but not least step of the exercise named **`Test_routine.m`** provide a few examples of plots that can be useful to represents the **`MANTA_TEC.m`** results. The *hodochrones* are probally the most common in ionospheric monitoring. The block starting in line 102 and finishing in line 157, loops through all the stations and the satellites available producing one figure per satellites. By setting the variable `save_fig = 'y'`, we can decide to save the plots in the directory name `plots`. As we can see from line 120, as default the script is set to plot the vector number 8 which contain the values of VTEC.



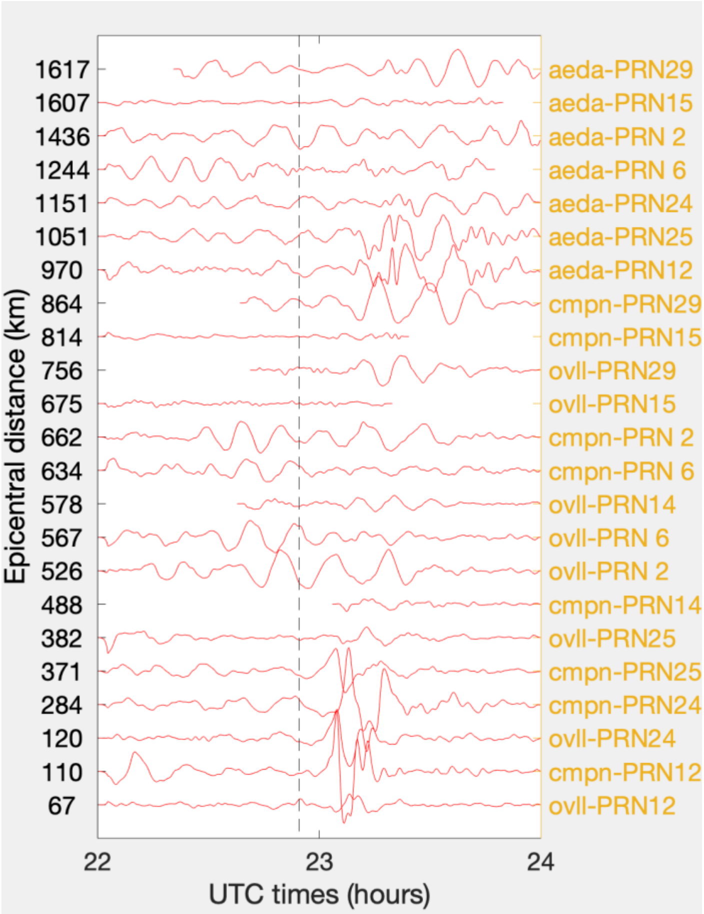
The figure shows an example of hodochrone related to satellite 24. The horizontal axis represents the time in hours of the date (UTC). And the vertical axis represents the IPPs distance from the epicenter. The color scale indicates the VTEC variations. The vertical grey dashed line represents the time of the evet. While the red dashed lines represent the propagation velocity of an acoustic wave propagating from the epicenter at 1000 m/s. **We can notice that some of the hodochrone plots are empty. This is due to the fact that during the selected time window some of the satellites were not operating. This is the case for satellites 1, 3, 4, 5, 7, 8, 9, 10, 11, 13, 16, 18, 19, 20, 21, 22, 23, 26, 27, 28, 30, 31. For this reason, we suggest that during the first run, the user should not save the hodochrones, and after having identify the satellites which are operating to save only them.**

The section of the script, from line 158 to line 181 can also be neglected for this exercise. This part of the code is useful when dealing with large number of stations (~100) and it helps to identify and select a certain group of station which satisfy a defined criterion, for example a distance from the source smaller then 500 km. In this exercise we have only a total of 3 stations and this step is not necessary.

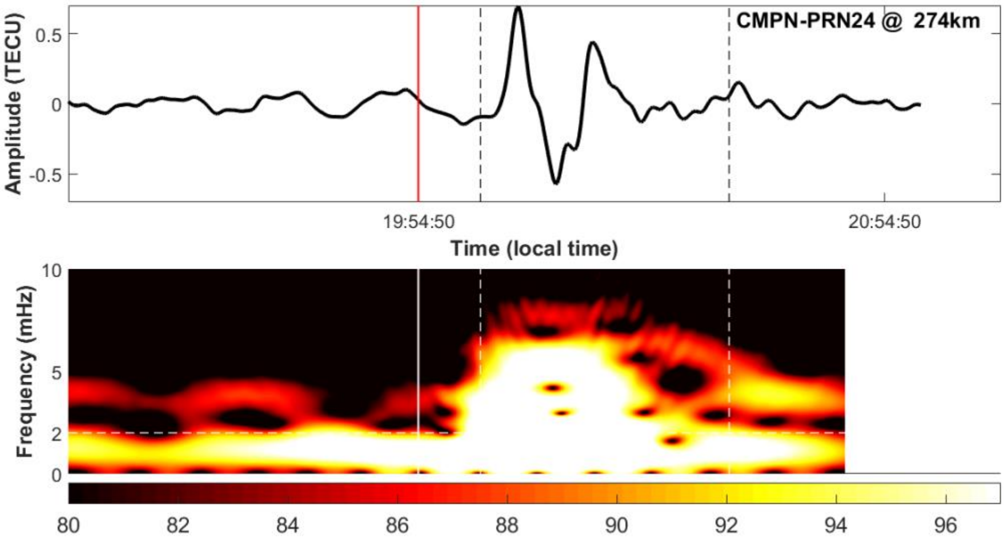
The following section, from line 182 to line 292 allows to produce a trajectory map of related to the IPPs of the available satellites and selected stations. As we can see from the figure below, the trajectory map represents the position of the IPPs during the time of the event. The red star represents the location of the epicenter. The black dots represent the position of the GNSS stations (labelled with their respective names). The grey dots represent the position of the IPPs at the time of the earthquake, while the gray lines they trajectory during a certain time window.



Another important and plot are the time series. This sub-section of script **`Test_routine.m`** which goes from line 293 to line 384 print in a single plot all the available TEC time series organized by distance to epicenter at the time of the earthquake. On the left side we can read the corresponding values of distance in km, while on the right side we find the associated GNSS-satellite name.



Finally, another important data visualization support, is the Power Spectral Density (PSD). The spectrogram of the filtered VTEC (or STEC) can provide us with information about the frequency content of the signal, and further allow us to understand the nature of the of the perturbation. In the figure below, we have an example obtained running the section of the code from line 383. The example shows on top the time series related to the station CMPN with respect to the GPS satellite 24. In the text notation we can find reported the distance to the epicenter at the time of the earthquake (~280km). With a red vertical line, is highlighted the time of the earthquake. The lower panel of the figure shows the frequency content and the color scale represent the intensity of the frequency. A horizontal dashed line marks the Brünt-Vaïsalla frequency that represents the limit between gravity and acoustic domains. This sub-routine save the output plots in a sub-folder of plots name **`spectro`**.


CMPN-PRN24 @ 274km
Amplitude (TECU)
0.5
0
-0.5
19:54:50
20:54:50
Time (local time)
Frequency (mHz)
10
5
2
0
80
82
84
86
88
90
92
94
96