TSAnalyzer: a GNSS Time Series Analysis Software

# Introduction

Time series analysis is important in any areas such as finical, astronomy, geophysics and so on. In the geophysics, time series are analyzed to extract the signals caused by geophysics phenomenon. This software is designed to analyze the GNSS coordinates time series and estimated the signals, such as linear, periodical signals. Also, some useful tools are provided for analysis convenience.

# Requirements and Installation

## Requirements

* Python 2.7
* PyQt4
* Matplotlib
* Numpy
* Pandas

## Installation

**TSAnalyzer** is based on Python 2.7, GUI is designed by PyQt4. For scientific use, I recommend WinPython on the windows, which is the leading open data science platform powered by Python and compiles scientific packages already, such as Numpy and Scipy.

Once you have installed WinPython, the requirements already have also been installed.

For Linux users, I would recommend Anaconda which has scientific packages and some requirements except for PyQt4. To install PyQt4, the following commands can be used.

Ubunut: sudo apt-get install python-qt4

Centos: sudo yum install PyQt4

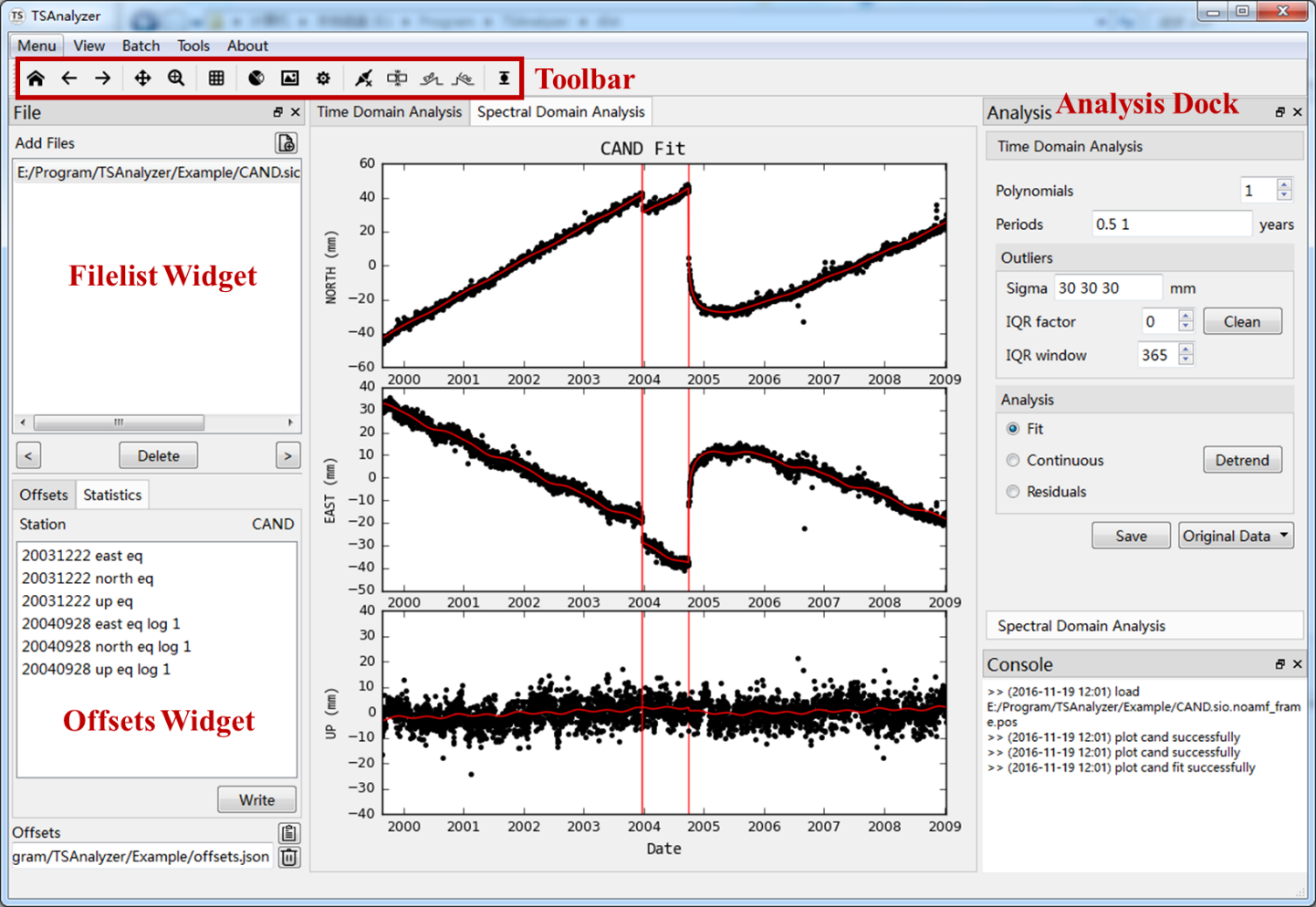
In the root directory, type ‘python main.py’ in the command prompt and it goes!

Also, we provide for win64 executable software.

# GUI Instructions

## Time Domain Analysis Instructions

Figure 1 GUI Time Analysis



* **Tool Bar** (from left to right)

Home plot view, pervious plot view, next plot view, pan the plot, zoom the plot, show grid line in the plot, subplot configuration, save the plot, plot settings (lines, markers, colors). These are same in spectral domain analysis. But the following tools are designed for GNSS time series analysis. Equipment break, earthquake break, earthquake exponential relaxation break, and earthquake logarithmic break. The last one is error-bar toggle button.

* **Files**

Add files button: ctrl + o

Read previous file: ctrl + q

Read next file: ctrl + w

Delete selected file: ctrl + d

Read file: double click the select file

* Offsets and Statistics

Edit offset item: double click

Write button: write offsets into file

Open offsets button: open offsets records file

* Analysis Area

Polynomials: Integer, 0-10, 0 means search for best orders

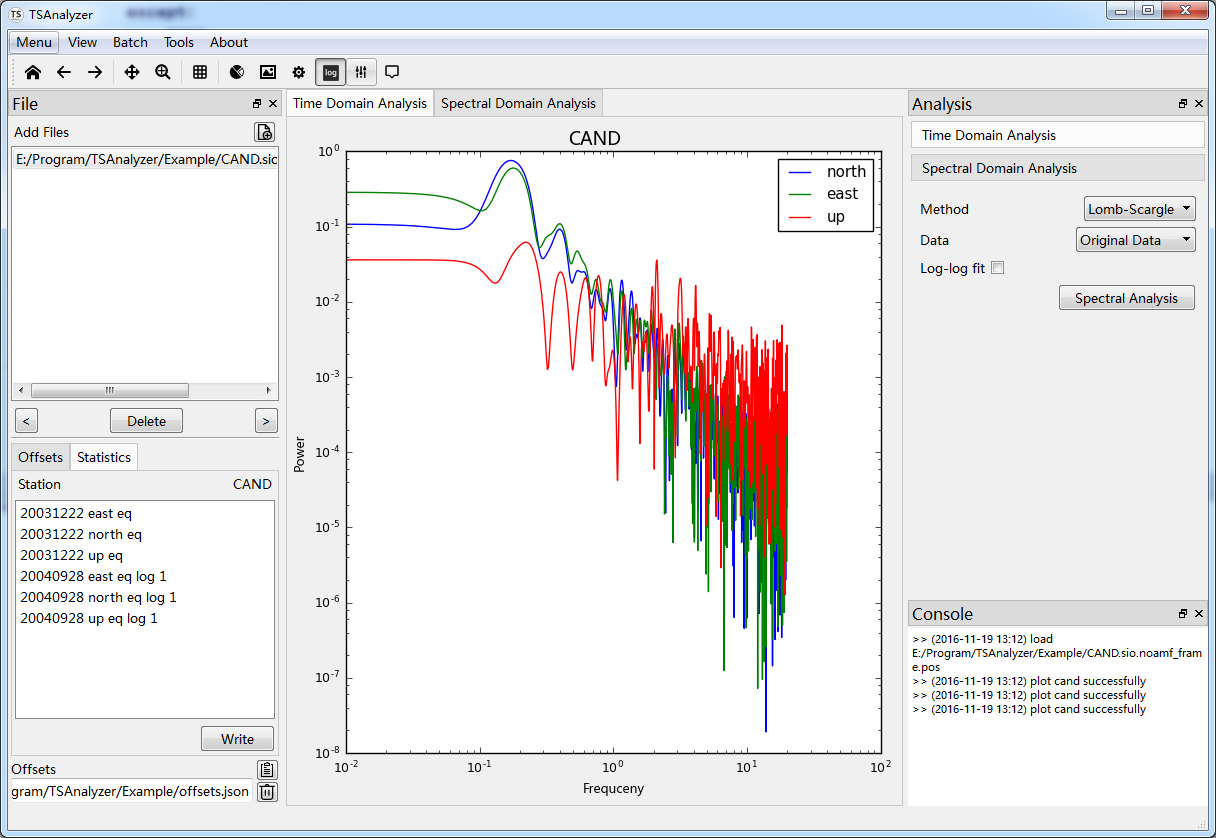
Periods: Number, space between each other

Outliers sigmas: filter the large sigma, number, space between each component.

IQR Factor and IQR window number required.

## Spectrum Domain Analysis Instruction

Figure 2 Spectrum Domain Analysis

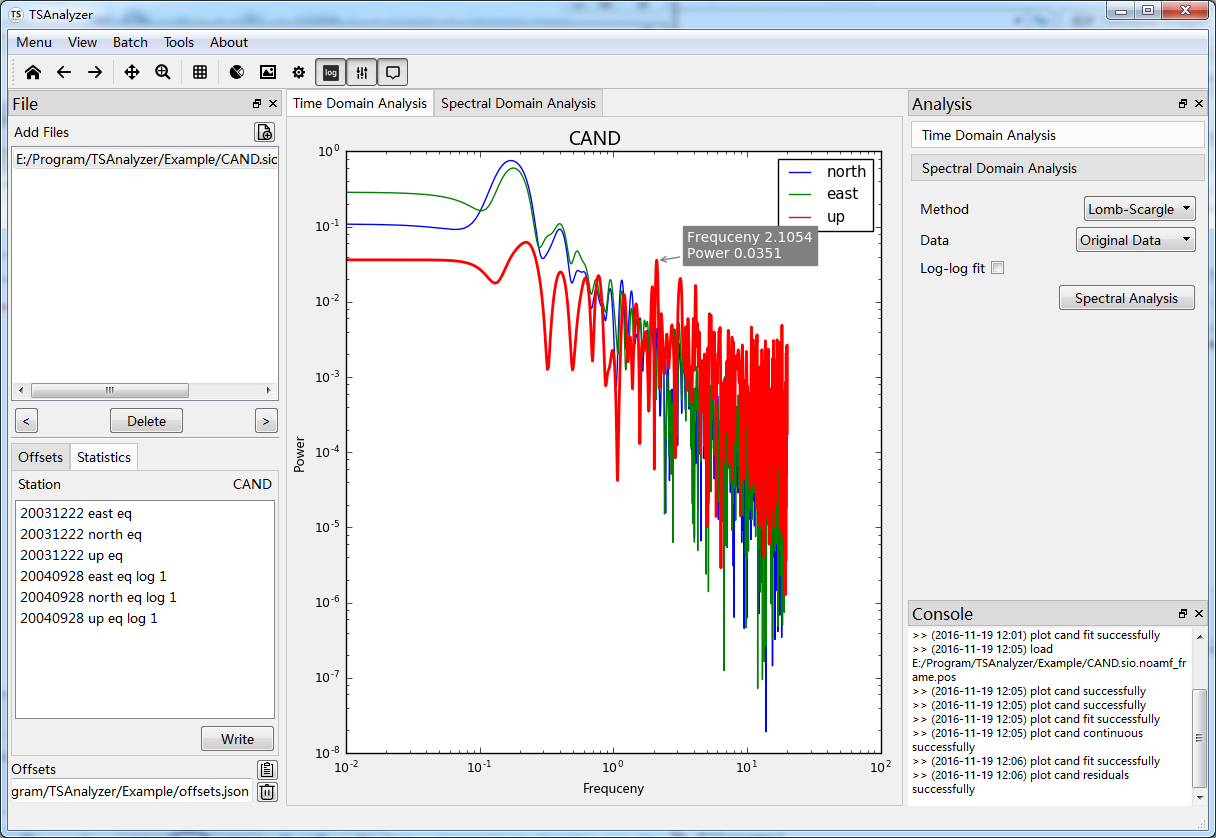


The Tool bar has 3 tools different from time domain analysis toolbar’s.

Log tool: toggle the figure between linear and log-log axis.

Highlight and Annotation tools provide interaction for view detail information better.

Figure 3 Highlight and Annotation tool functions

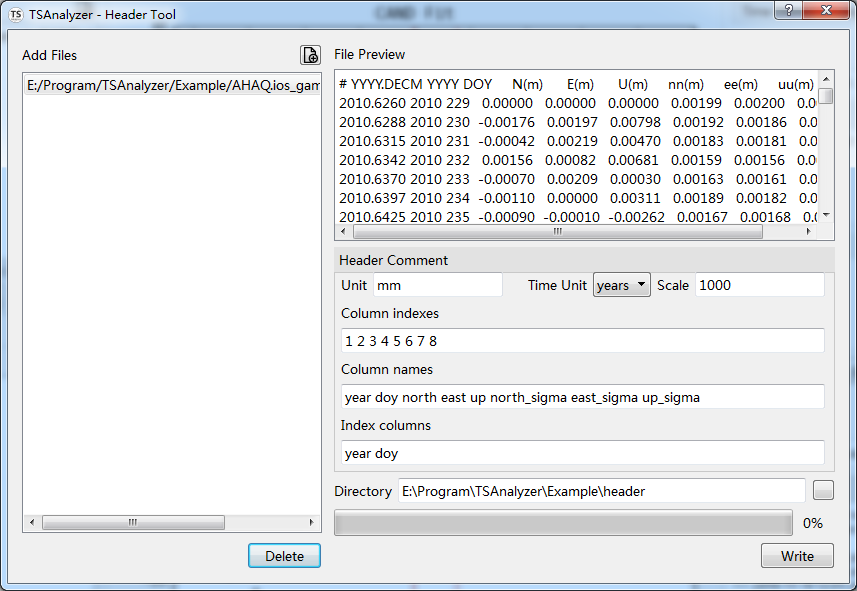


On the Analysis area, you can select methods and data for spectral analysis. **TSAnalyzer** supports lomb-scargle method only at present.

Log-log fit can plot a log-log fit line on the plot and display the slope.

## Header Tool

Figure 4 Header Tool



Head tool is in the menu tools. By using this tool, you can set your files head comments and can be understood by **TSAnalyzer**. Suppose that you have many other format time series files, you can use this tool to add comment and avoid doing any conversion. The hands-on tutorial will show a detail introduction.

# Analysis Instructions, a Hands-on Tutorial

## Header Tool

**TSAnalyzer** supports *pos, tseries, cats neu* formats currently. For more flexible formats, Header Tool is provided for add comment header to original time series files.

From Figure 4, there some parameters should be input.

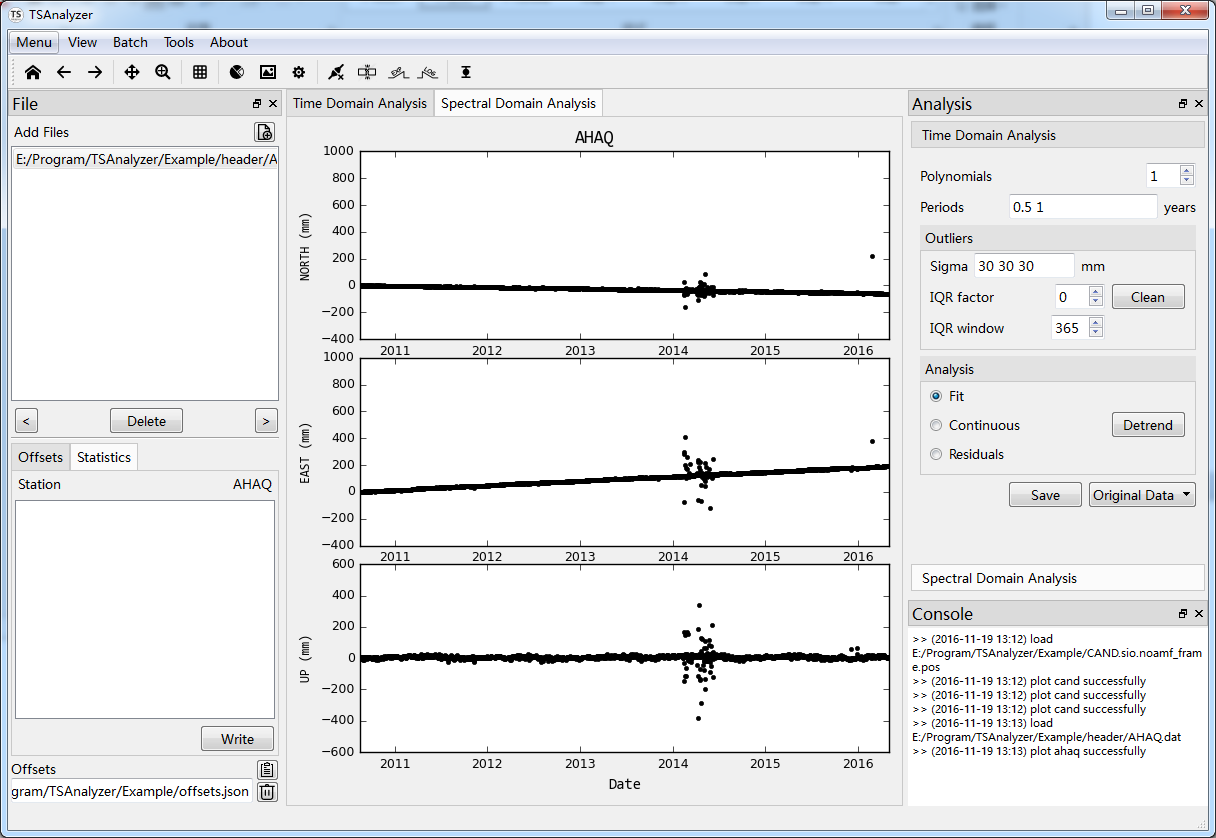
* **Unit** mm abbreviation for millimeter, although the original file’s unit is meter, but we use **scale** factor 1000 to convert meter to millimeter.
* Time unit years and days supports at present.
* Column indexes and column names, starting from 0 in this example, we use columns year (1), doy (2), north (3), east (4), up (5), north\_sigma (6), east\_sigma (7) and up\_sigma (8). Index columns supports the following key words:

year, month, day, hour, minute, seconds, doy, ymd, hms, mjd

* doy, day of year
* ymd, year month day, for example 2010101
* hms, hour minute seconds, for example 120000
* mjd, Modified Julian Date
* The directory is used to save new files after adding header comments.

After adding header comments, we could use TSAnalyzer to load these new files (Figure 5).

Figure 5 TSAnalyzer AHAQ After Adding Header

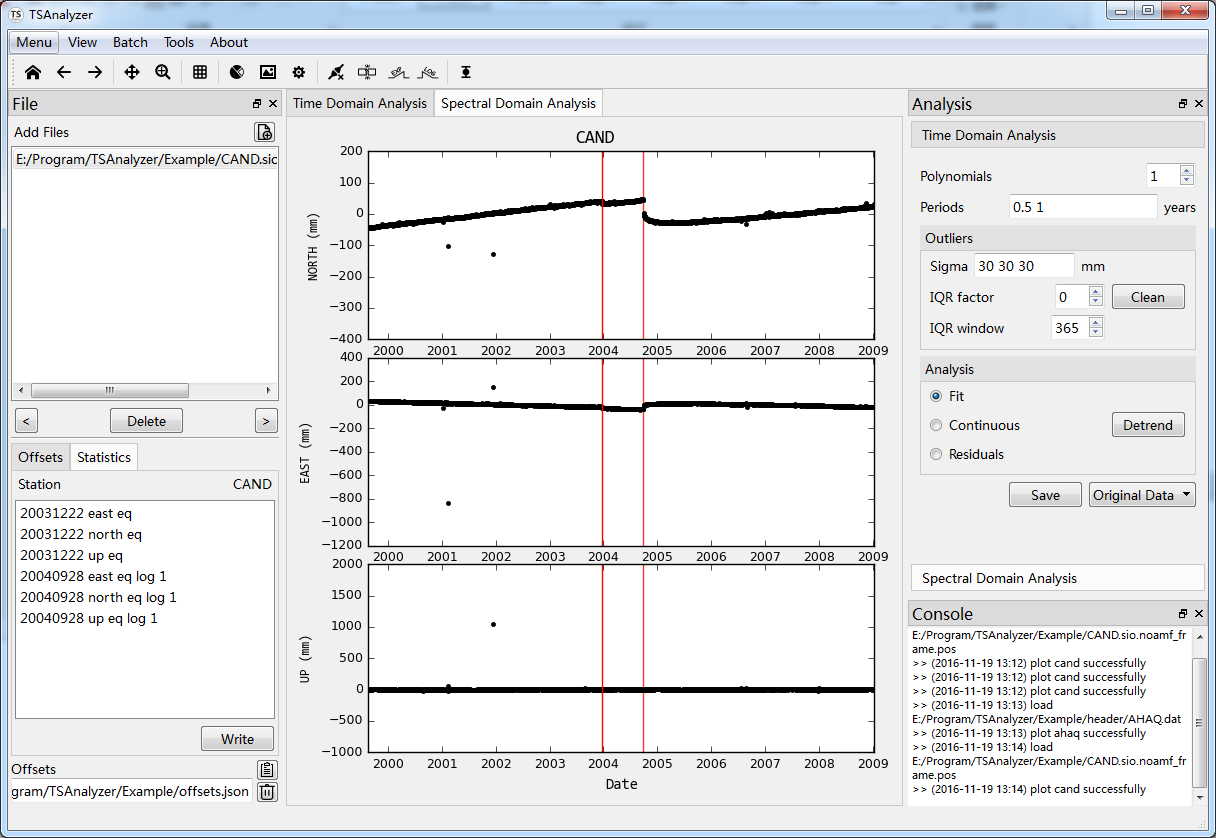


## Time Series Analysis Step by step

1. Read time series file

**Double click** the filename to load the file. **TSAnalyzer** will plot the time series files components separately.

Figure 6 Load CAND pos

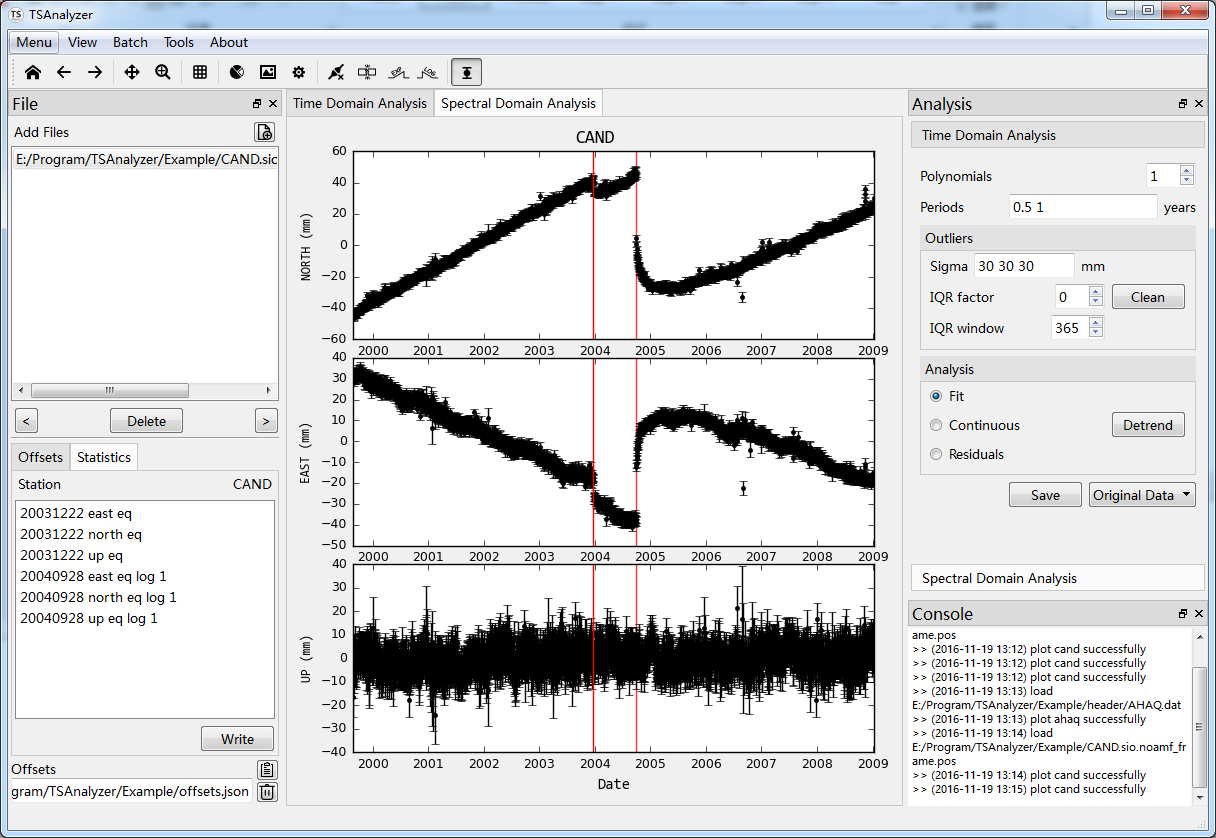


1. Pick the offsets

In the toolbar, there are four kinds of offsets, from left to right: equipment break, earthquake break, earthquake exponential relaxation break, and earthquake logarithmic break.

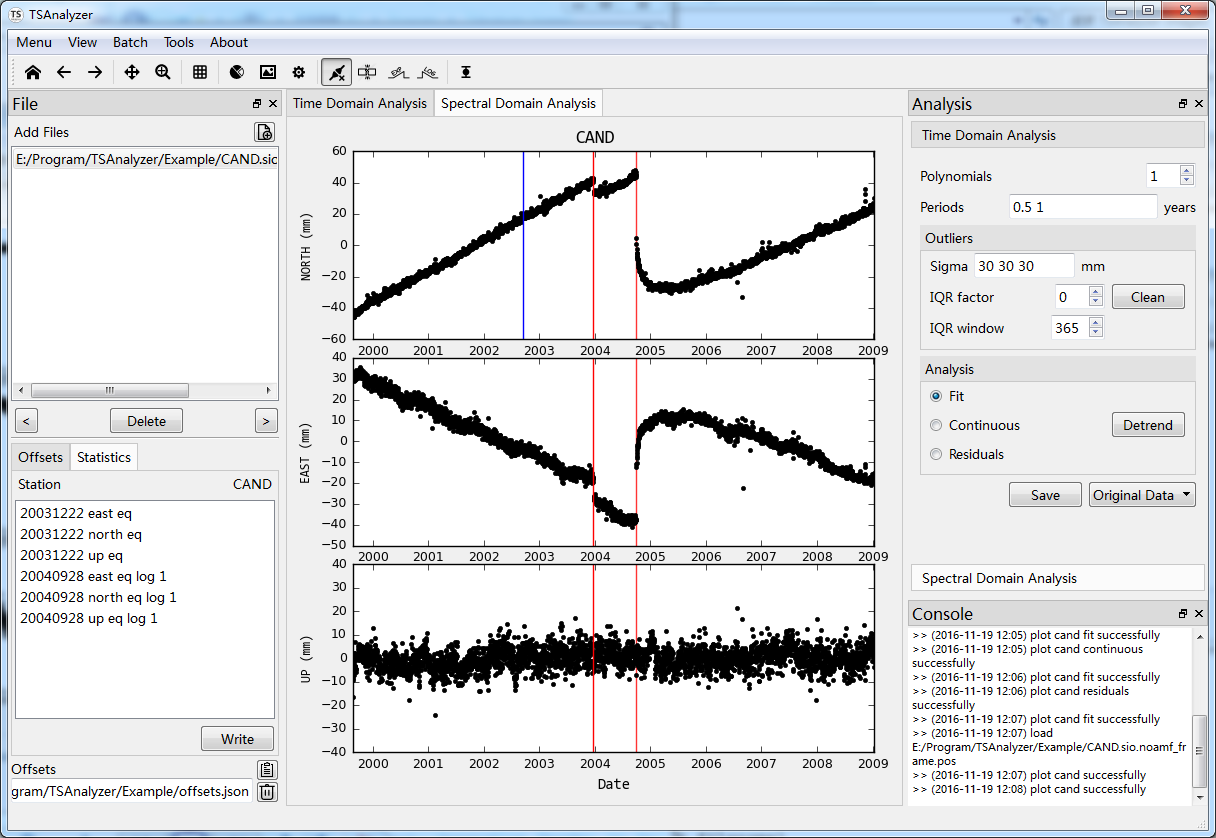
It is not obvious to see offsets because of the outliers’ existence. The outliers have very large uncertainties. Hence we could set sigma criterion to filter theme. In the outlier group, set the IQR factor to 0, and do the remove outliers.

Figure 7 Filter large sigma outliers



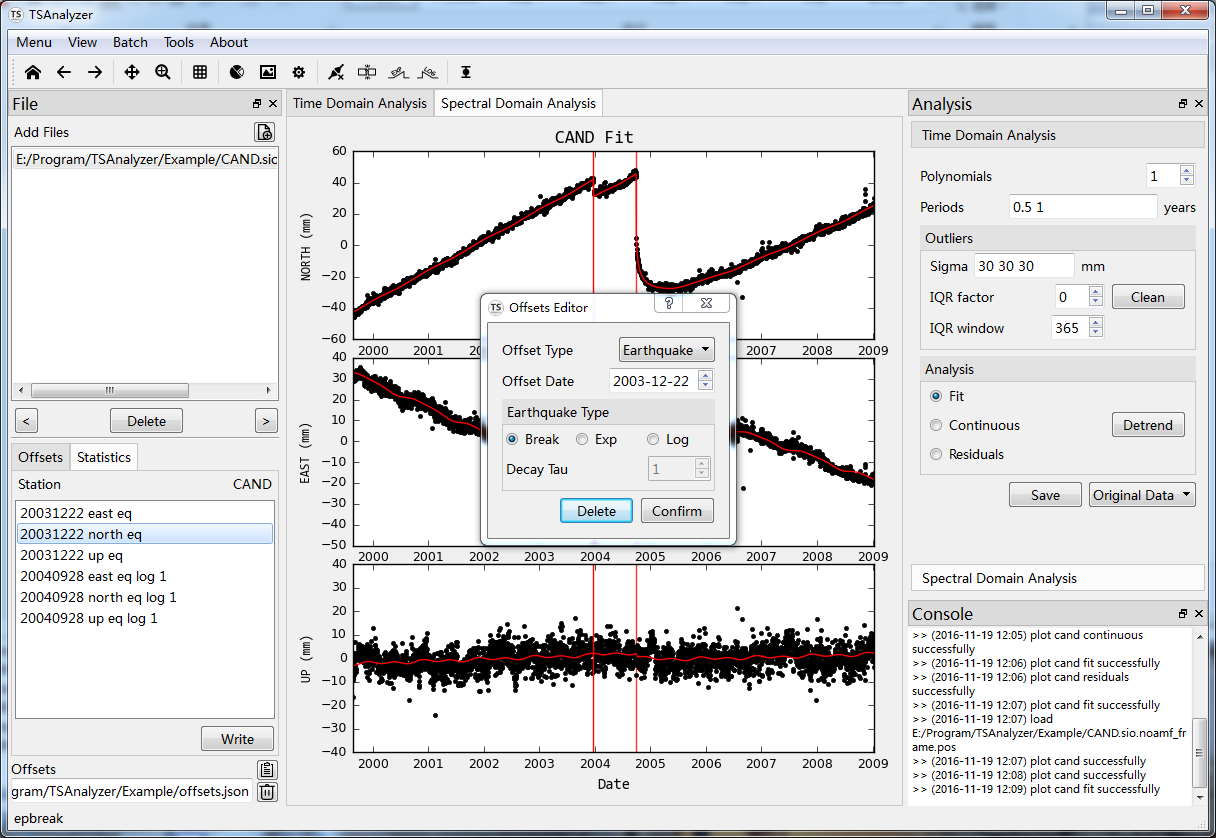
From the Figure 7, we could see there existed two offsets. We use break and log break to pick them out. Zoom and pan has been provided on the toolbar to pick accurately.

Figure 8 Pick offsets



In the Figure 8, we could see the red rectangle, it records these offsets. You can double click these items to edit accurately from the popup dialog (Figure 10). After you have pick these offsets, click the write button on the left bottom corner to save these offsets for later use.

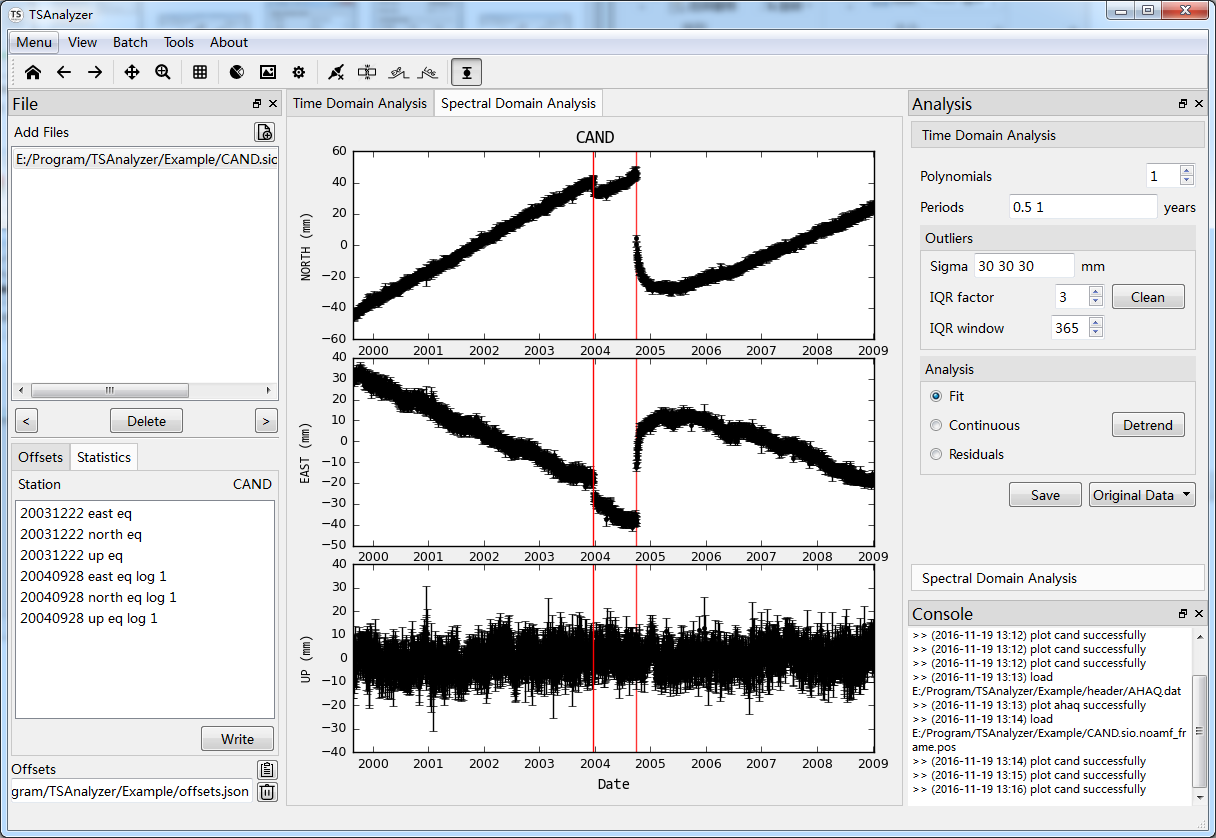
Figure 9 Popup Offsets Dialog



1. Remove the outliers

In the step 2, we adopt sigma criterion to remove some gross errors. There still some outliers displayed on the plot. So IQR method can be used. IQR factor and IQR window can be input according to your data. Here we select factor 3 and window 365 (year).

Figure 10 IQR Method

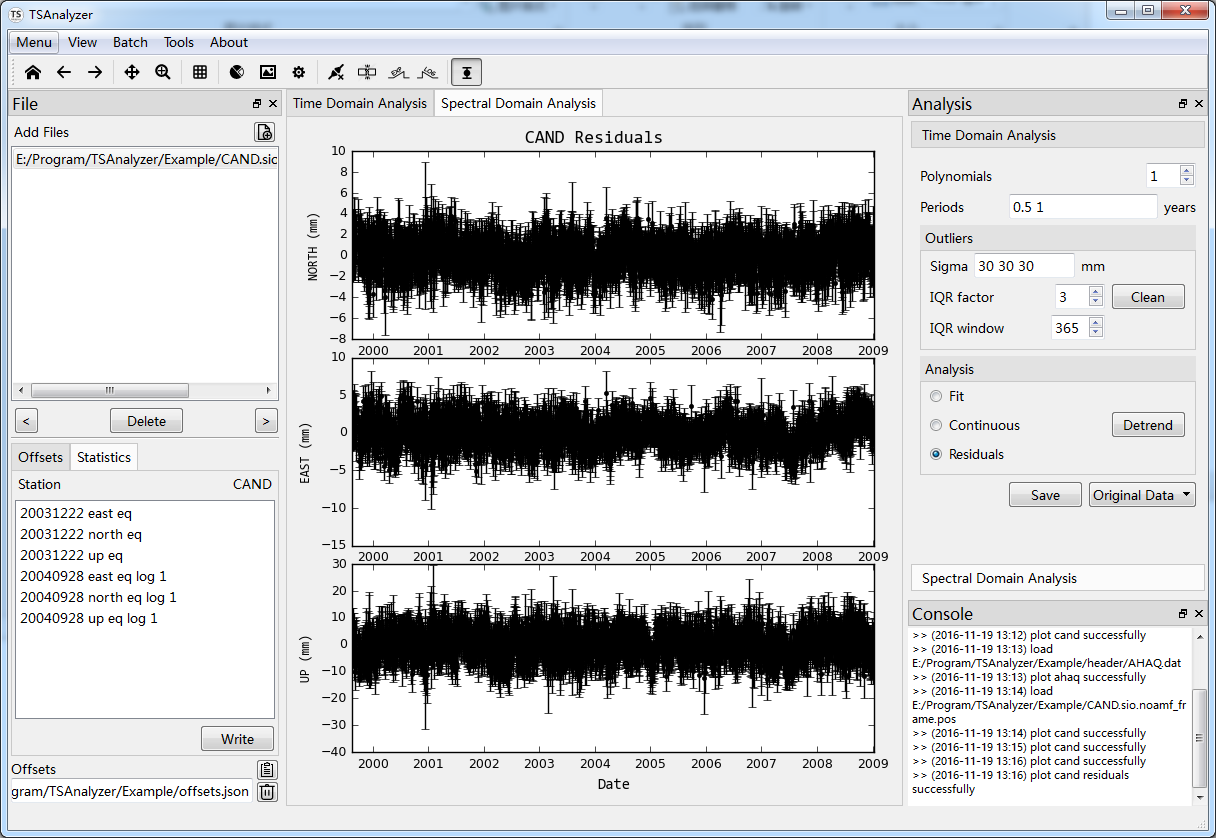


1. Time Series Analysis

In the step 3, IQR method adopt the paraments polynomials and periods. Polynomial order can be set from 0 to 10, 0 stands for searching the best polynomial order. Period not fixed to one, space between each other, and the unit in this data is year, this means we input semi-annual and annual periods.

If the “Fit” is checked, it plots the fit line on the clean plot, “Continuous” stands for removing any offsets, “Residuals” means plotting residual time series. Click the detrend button to do analysis.

Figure 11 Detrend



The result will be showed in the popup dialog. If you are boring with popup dialog, in the Menu View to uncheck the show detrend log button and the result will be should in the console area.

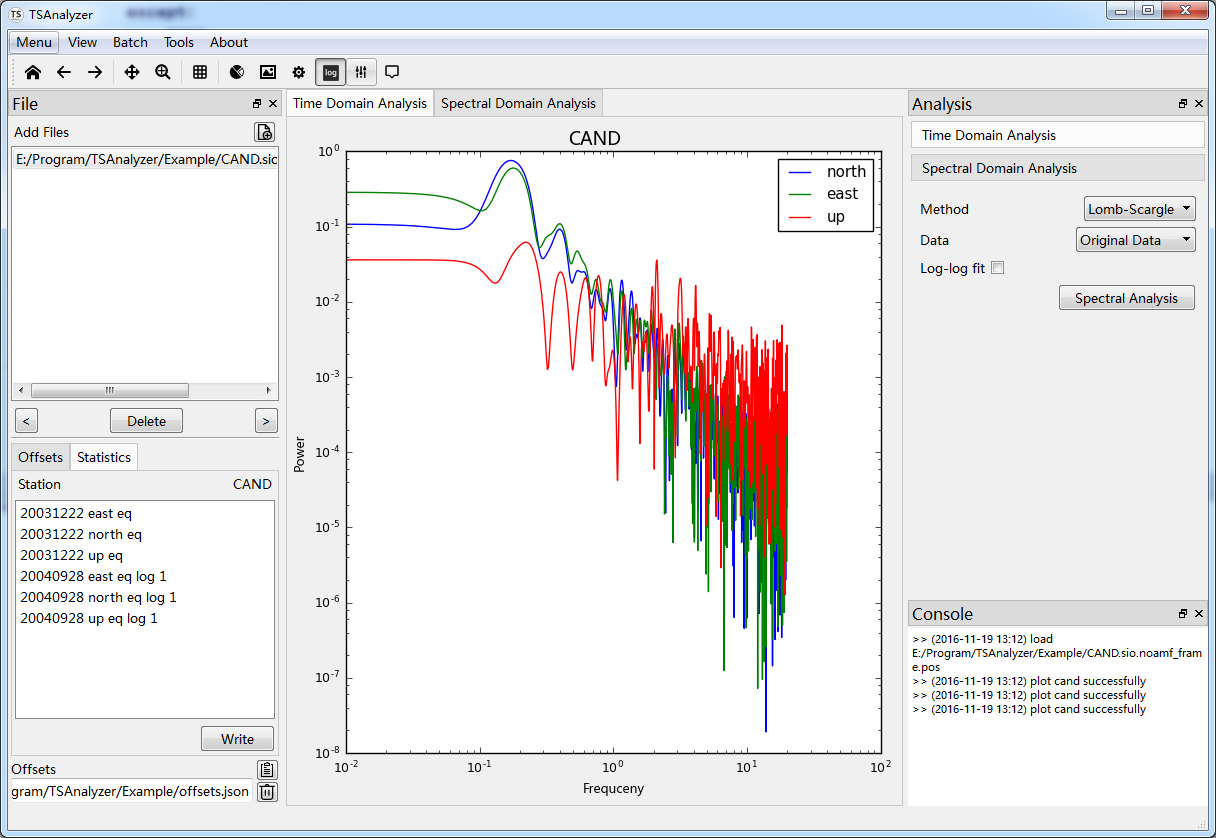
1. Save the Data

The original data means the input data, if you do some outliers action, then the data is the clean data. The fit data is the red fit line data and the residuals is the difference the original data between fit data.

## Spectral Domain Analysis

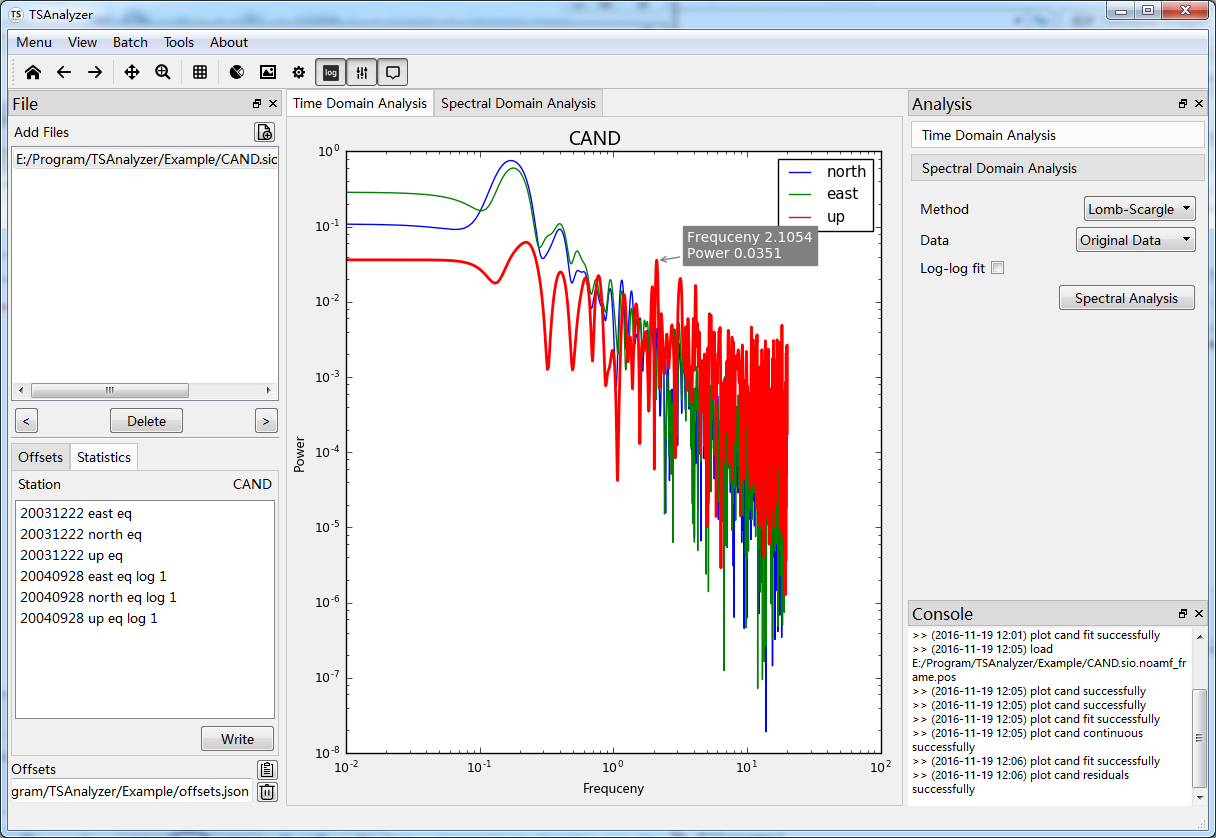
Spectral domain analysis is useful for time series analysis. For GNSS coordinates time series, it is common for gap and discontinuities’ existence. So Lomb-Scagle method is used very often.

Figure 12 Spectral Analysis



Click the spectral analysis button, the result will be plotted. In the toolbar, there are some tools for better visualization. Log-log axis, highlight lines and annotation tools.

Figure 13 Spectral Tools



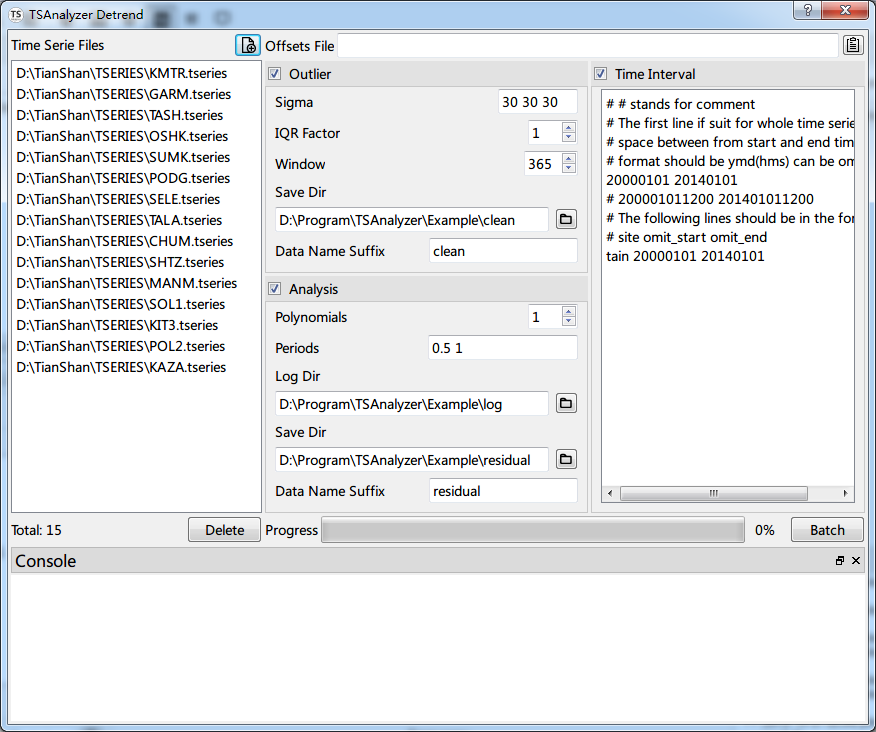
## Batch Operation

Because of the growing number of GNSS stations, many organizations will process the hundreds or thousands of GNSS time series. **TSAnalyzer** provide tools for batch analysis, including detrend and plot.

In the menu batch, there are detrend and plot batch menus.

### Detrend Batch

Figure 14 Analysis Batch



The batch dialog has a similar operation as the previous time domain analysis’. The parameters are the same, except for some directories and time interval (the right panel).

* Outlier **Save Did**, in which the data will be saved after remove the outliers
* Outlier **Data Name Suffix**, the name of the clean data saved, taking KMTR.tseries for example, the clean data will be saved in D:\Program\TSanalyzer\Example\clean\KMTR**\_clean**.neu.
* Analysis **Log Dir**, analysis log mentioned in time domain analysis will be saved in this directory.
* Analysis **Save Dir**, the original clean data after removing the polynomials, periods and offsets from will be saved in this directory.
* Analysis **Data Name suffix**, like the outlier data name suffix
* **Time interval**, designed for individual stations flexibly.

Figure 15 Time Interval Comments

The example show detail comments.

# # stands for comment

# The first line if suit for whole time series file

# space between from start and end time

# format should be ymd(hms) can be omitted

20000101 20140101

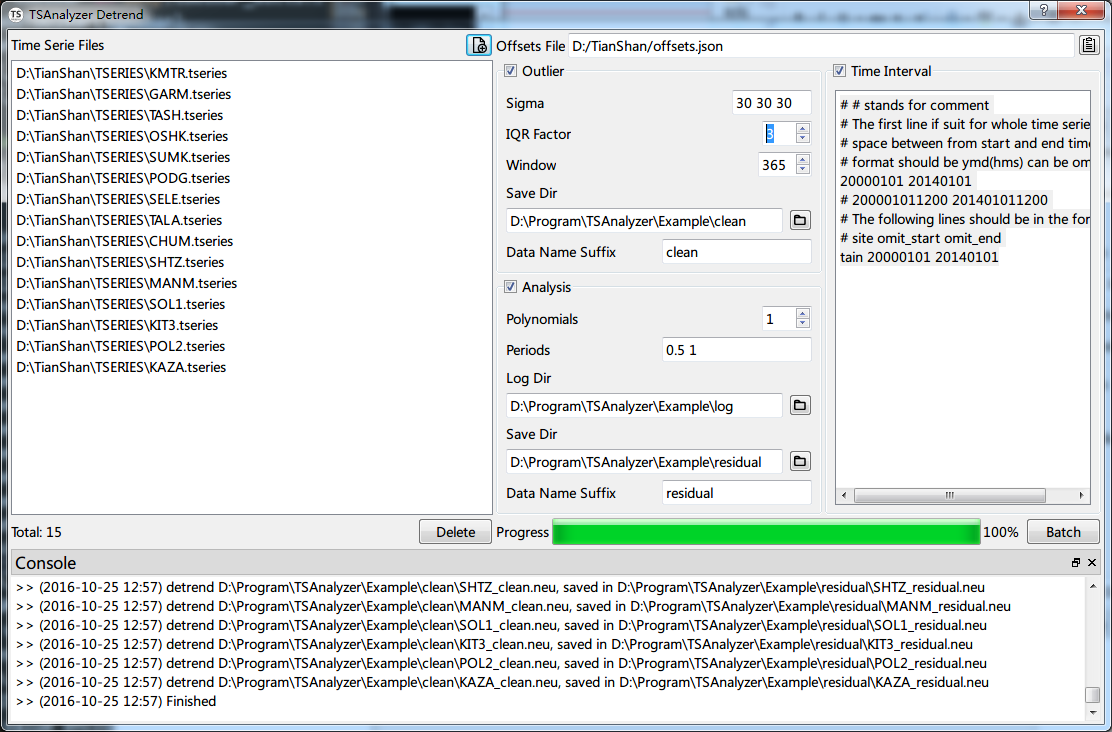
# 200001011200 201401011200

# The following lines should be in the format

# site omit\_start omit\_end

tain 20000101 20140101

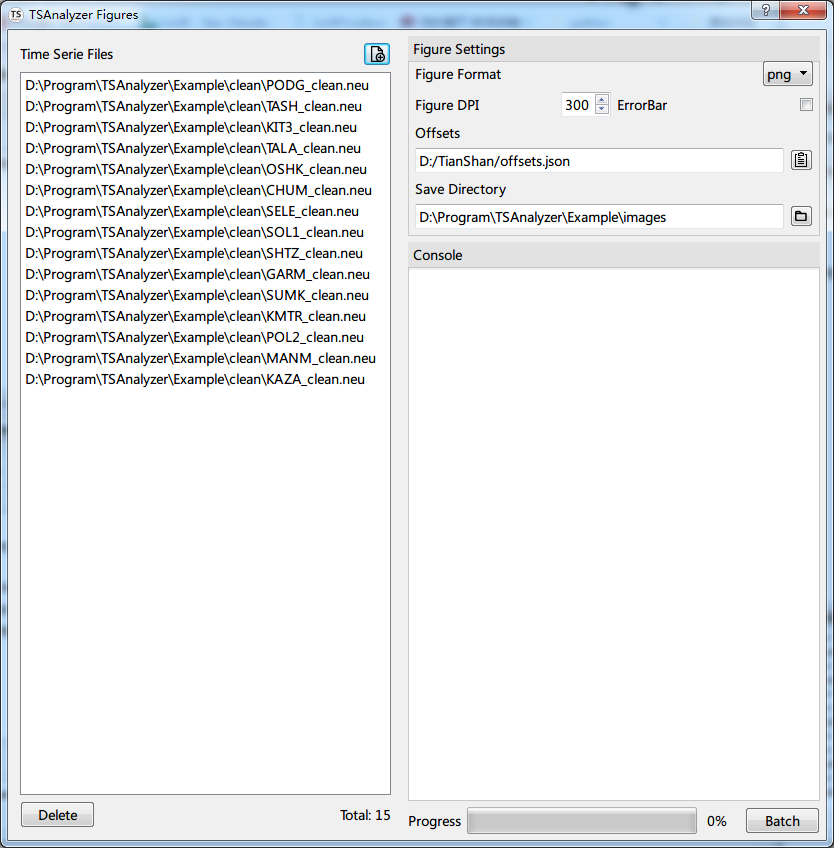
Figure 16 Batch Detrend End

After setting the parameters mentioned, click batch button and it goes!

### Figure Batch

Also for plot figures easily, **TSAnalyzer** provides figure batch tool (menu - batch - figure). It has limited customs style at present, including figure format (png, jpg, eps, pdf) figure dpi, and error bar plotting or not. If you set the offsets file, it will plot the offsets events on the figure. Click the batch button and it goes!

Figure 17 Figure Batch



# Mathematic Model

## Least Square Analysis

In the equation, is polynomials*,* is the velocity, and are harmonic components, usually annual and semi-annual. The is a step function, and can be used to explain the sudden change event caused the equipments or co-sesimics. and terms are stand for logarithmic or exponential function to be fit after an Earthquake. These term are the accumulated post-seismic motion, is the co-seismic offset, and is the decay time.

This model can be used to estimated polynomials (max order is 10, if zero is selected, it can search for best order), harmonics (users can define the period) as well as step function (including sudden change, post-seismic log or relaxation) at specific times (users pick interactively or input offsets file), detail information is displayed in the picture.

## Outliers

### Big sigma criterion

This criterion is according the sigma (user input, accepted one or three numbers) to mask out the data whose sigma is larger than sigma criterion to get the cleaner data. One number input means that three sigma criterion for all components are the same. If you want to specific for individual component, please use space between numbers.

Apart from this, we also use 3 times standard deviations criterion here. Equation is:

f sigma is not set, this criterion will not be adopted.

### IQR

By using the least square model, one could get the residuals , then the software use the window (user input) to get the percentile of 75%, median and 25%.

By using the above model, the outliers can be found and removed. If iqr fatcor is 0, this criterion will not be used.

## Spectrum Analysis

Spectral analysis method is based on Lomb Scargle, the source code comes from Astrop. For datail information, the instructions can be found on <http://docs.astropy.org/en/stable/stats/lombscargle.html>.

We adopted the astropy’s lombscargle code here.