

TraPSA (Trajectory-based Potential Source Apportionment)
Software User's Guide

Developer: Chuanlong Zhou

The Center for Air Resources Engineering & Science (CARES)

Clarkson University

Advisor: Thomas M. Holsen; Philip Hopke

Email: zhouc@clarkson.edu; holsen@clarkson.edu;
hopkepk@clarkson.edu

Website: <http://adweb.clarkson.edu/projects/TraPSA/>

TraPSA software is non-commercial and for research and educational purpose only. In order to avoid unpredictable errors, do not install software into 'C:\Program Files\' and do not use a shortcut. Error logs will be generated in the application folder if an error occurs. Please send this error log to zhouc@clarkson.edu for the developer to further improve this software. Thank you very much.

Context

1. Introduction	1
2. Software Structure	3
3. Algorithms.....	4
3.1 Mann-Kendall Test and Sen' s slope	4
3.2 Conditional Probability Function (CPF)	4
3.3 Concentration Field Analysis (CFA).....	5
3.4 Concentration Weighted Trajectory (CWT)	5
3.5 Residence Time Weighted Concentration (RTWC).....	6
3.6 Simplified Quantitative Transport Bias Analysis (SQTBA).....	7
3.7 Potential Source Contribution Function (PSCF).....	8
3.8 Weight Function	9
3.9 Event analysis.....	10
3.11 Grid Calculation	1
4. Data Required	11
4.1 Meteorological data	11
4.2 Map Data	11
4.3 Measurement Data	11
5. Example.....	13
5.1 Case Introduction.....	13
5.2 Software Installation	13
5.3 Before Analysis.....	14
5.4 Monitoring Site Database.....	15
5.5 Generating Trajectories.....	18
5.6 Item to be Analyzed	21
5.7 Basic Analysis.....	22
5.8 Trajectory Ensemble Receptor Models	24
5.9 Event Analysis.....	26
5.10 Emissions Inventory.....	27
5.11 GIS Editing	28
5.12 A Final Result	30
6. Reference.....	31

1. Introduction

The TraPSA (Trajectory-based Potential Source Apportionment) project was developed as a free and open-source toolkit for air pollution potential locations apportionment analysis based on trajectory ensemble receptor models. TraPSA software is a graphical and powerful tool to help the air quality community use massive monitoring and trajectory data for pollutant potential source location apportionment. The main interface of TraPSA is shown in Figure 1.1, which provides monitoring site database management, smart trajectory generating and handling, pollutant temporal pattern analysis, most of existing trajectory ensemble receptor models, and sufficient GIS edit functions.

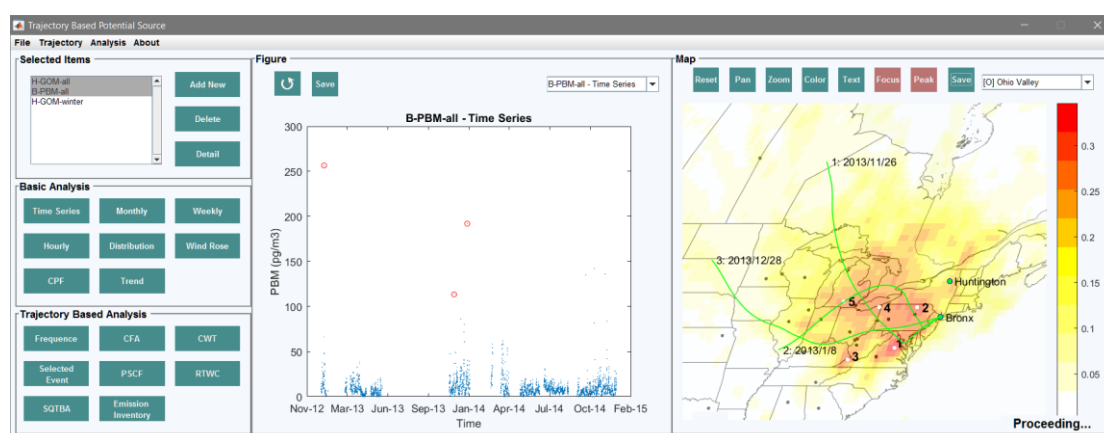


Figure 1.1 TraPSA Main Interface

A database of pollutant monitoring site data can be established in TraPSA. The smart back-trajectory method in TraPSA helps users easily set up, calculate, and import trajectory data. TraPSA includes current popular algorithms for trajectory ensemble receptor models including Conditional Probability Function (CPF), Concentration Field Analysis (CFA), Concentration Weighted Trajectory (CWT), Residence Time Weighted Concentration (RTWC), Potential Source Contribution Function (PSCF), and Simplified Quantitative Transport Bias Analysis (SQTBA).¹⁻⁹ TraPSA provides users sufficient GIS editing functions for mapping air pollutant source apportionment. In addition, GIS data files (ESRI shape file and Geo TIFF file) can be imported or exported by TraPSA allowing further research or editing by GIS software.

TraPSA was developed as an extension package of HYSPLIT_4 software and can be linked to local HYSPLIT_4 installations on the computer (introductions to HYSPLIT are at http://www.arl.noaa.gov/HYSPLIT_info.php). However, TraPSA can also be used without installation of the full HYSPLIT_4 software suite because TraPSA software includes an unregistered version of the HYSPLIT_4 trajectory model executable. However, it is strongly recommended that users register and install the most current version of HYSPLIT software in order to generate the most accurate trajectories.

TraPSA was programmed in MATLAB 2015b. It is only executable on Windows OS platforms and a 1920 X 1080 screen resolution is recommended. MATLAB runtime is required to run TraPSA so that the full version of MATLAB is not necessary. MATLAB runtime will automatically be downloaded if it hasn't already been installed on the computer.

2. Software Structure

The structure of TraPSA is shown in Figure 2.1. Pollutant measurement data, meteorological data and ESRI shapefile of related geographic region (grey blocks) are required. An air pollutant monitoring site database should be established with site information and pollutant measurement data. Trajectory endpoints will be generated based on pollutant measurement date using meteorological data and HYSPLIT model. As an optional data-input, trajectory endpoints can be imported into the database if the files were previously generated by HYSPLIT_4.

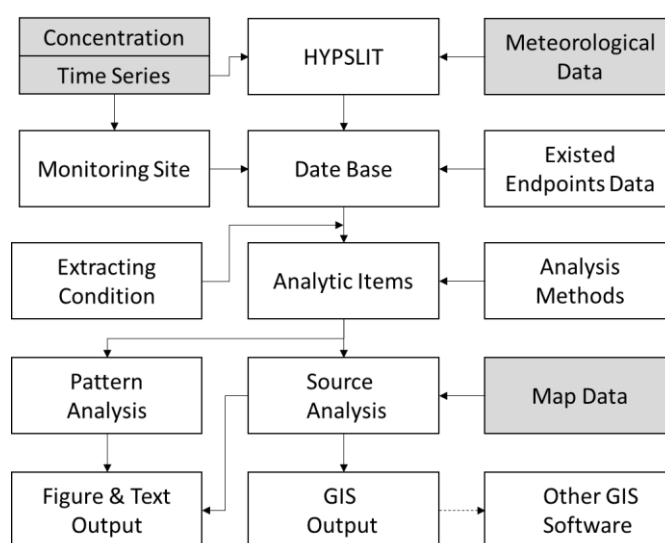


Figure 2.1 Software Structure

After program execution, pollutant data and corresponding trajectory endpoints can be extracted from the database and used for analysis. TraPSA contains modules to compute pollutant pattern analysis, trend analysis, and trajectory ensemble receptor models (including CPF, CFA, CWT, RTWC, PSCF, and SQTBA) for potential source location identification. An ESRI shapefile of the geographic region is required to correctly display the results of the trajectory ensemble receptor models. Any figures generated by TPBS can be exported including their numerical data. Trajectory endpoints and the map raster can be exported in GIS data formats (ESRI shape file and Geo TIFF file) if further analysis and/or editing is necessary in other GIS software.

3. Algorithms

3.1 Mann-Kendall Test and Sen's slope

The Mann-Kendall test is a nonparametric test for monotonic trends. Kendall's Tau coefficient permits a comparison of the strength of monotonic correlation between two data series. When a monotonic trend is demonstrated by the Mann-Kendall's Tau (calculated using MATLAB built-in function), Sen's slope, a nonparametric linear regression models is used for estimating the slope of the trend, which is robust to outliers and non-normality in residuals¹⁰. This method has been widely used in climate and hydrological trend analysis^{11, 12}. Sen's slope is defined as the median of the slopes $(y_i - y_j)/(x_i - x_j)$ determined by all pairs of sample points in a set of two-dimensional points (x_i, y_i) .

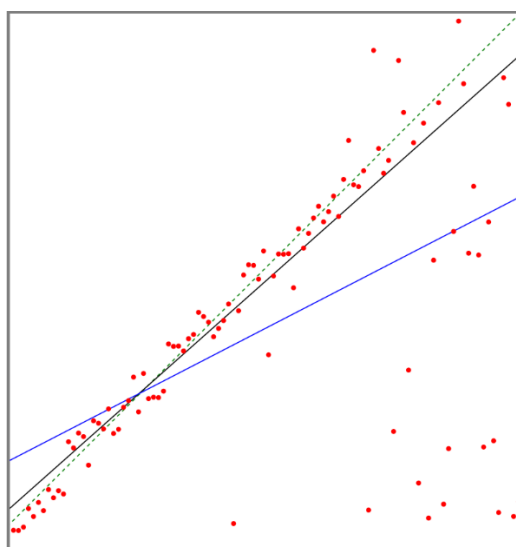


Figure 3.1 The Theil–Sen estimator of a set of sample points with outliers (black line) compared to the non-robust simple linear regression line for the same set (blue). The dashed green line represents the ground truth from which the samples were generated.¹³

3.2 Conditional Probability Function (CPF)

CPF determines the probability of a wind direction to be associated with specific pollutant levels and can be useful for determining local source directions¹⁴. A criterion value C (which is adjustable in TraPSA) representing high concentration events is arbitrarily set (usually 75%~90% of the highest concentration). The CPF value for different wind directions is calculated as:

$$CPF_i = \frac{P_i}{F_i}$$

4

CPF_i : CPF value of direction i

P_i : the wind frequency of the direction i in the high concentration events

F_i : the wind frequency of direction i

Note that the wind direction has a large uncertainty for low speeds (typically [5 km/h] although the parameter can be manually set). Therefore, these data should be discarded. Also note that the wind speed and direction, which are generated by HYSPLIT4, will probably change with the trajectory starting height used.

3.3 Concentration Field Analysis (CFA)

The CFA model determines air pollutant source locations by combining concentration measurements with back trajectories^{9, 15}. For this method, the whole geographic region is divided into an array of grid cells defined by the cell indices i and j . Back-trajectories (presented by l) will be generated starting at concentration measurement time at receptor site. Trajectory endpoints in grid cells (presented by τ) will be counted after back-trajectory calculations. If a trajectory endpoint lies in the grid cell, the trajectory is assumed to collect and transport the material emitted in this cell along the trajectory to the receptor site. The CFA values then can be calculated as shown below so that grid cell with larger CFA value implies the higher contribution of pollutant to the receptor site:

$$CFA_{i,j} = \frac{\sum_{l=1}^L \log(c_l) \tau_{i,j,l}}{\sum_{l=1}^L \tau_{i,j,l}}$$

$CFA_{i,j}$: CFA value of grid i, j (latitude, longitude)

c_l : receptor concentration measured at the starting time of back trajectory l

L : total back trajectory line number

$\tau_{i,j,l}$: endpoint number of back trajectory l in grid i, j

3.4 Concentration Weighted Trajectory (CWT)

The concentration field values of the grids that a trajectory passes through with a concentration close to 0 will significantly reduce their CFA due to the logarithmic calculations used in CFA. The CWT model is a modification of CFA using a linear calculation, which is more robust to low pollutant measurements. The CWT values are calculated as follows^{7, 16}:

$$CWT_{i,j} = \frac{\sum_{l=1}^L c_l \tau_{i,j,l}}{\sum_{l=1}^L \tau_{i,j,l}}$$

$CWT_{i,j}$: CFA value of grid i, j (latitude, longitude)

c_l : concentration at the starting time of back trajectory l

L : total number of back trajectories

$\tau_{i,j,l}$: endpoint number of back trajectory l in grid i, j

3.5 Residence Time Weighted Concentration (RTWC)

In order to distinguish a large source from a moderate source¹⁷, the RTWC model can be used to calculate a grid concentration field^{5, 8}. The rationale for the redistribution used in the RTWC approach is that no major pollutant sources are located along a “clean” trajectory (i.e. one with a very low concentration at the receptor site). The “polluted” trajectory (i.e. one with a high concentration at the receptor site) must have been influenced by sources along its path through which no “clean” trajectories pass. Also RTWC can be applied into multiple sites since a trajectory going to one monitoring site cannot typically go to another and thus, it helps clean up trailing effects.

The initial field can be calculated by the CWT or CFA model and the redistributed concentrations for every trajectory can be calculated:

$$C_{l,t} = c_l \times \frac{X_{l,t}}{\frac{1}{N_l} \sum_{t=1}^{N_l} X_{l,t}} = c_l \times \frac{X_{l,t}}{\bar{X}_l}$$

$C_{l,t}$: redistributed concentration for time t in back trajectory l

c_l : receptor concentration of back trajectory l

N_l : total hour of back trajectory l

$X_{l,t}$: $X_{l,t} = CF_{i,j}$ if back trajectory l hit grid i, j at time t

Then concentration field is iterated until the average difference between the concentration fields of two successive iterations is below a criterion value (typically 0.5%):

$$CF_{i,j}^k = \frac{\sum_{l=1}^L \sum_{t=1}^{N_l} c_{l,t}^{k-1} \tau_{i,j,l,t}}{\sum_{l=1}^L \sum_{t=1}^{N_l} \tau_{i,j,l,t}}$$

CF_{ij}^k : CF value of grid i, j (latitude, longitude) at iteration time k

L: total number of back trajectories

τ i, j, l, t: endpoint number of back trajectory l at time t in grid i, j

3.6 Simplified Quantitative Transport Bias Analysis (SQTBA)

The uncertainties in a trajectory pathway increase with the increasing trajectory length^{18, 19}. These uncertainties are considered in the SQTBA model, a simplification version of QTBA⁶. SQTBA assumes a normal distribution caused by atmospheric dispersion is approximated about the trajectory centerline with a standard deviation that increases linearly with time in the upwind direction. Thus the transition probability density function can be expressed as¹⁷:

$$Q(x, t|x', t') = \frac{1}{2\pi\sigma_x(t')\sigma_y(t')} \exp \left[-\frac{1}{2} \left\{ \left(\frac{X - x'(t')}{\sigma_x(t')} \right)^2 + \left(\frac{Y - y'(t')}{\sigma_y(t')} \right)^2 \right\} \right]$$

$Q(x, t|x', t')$: the probability for the air parcel at position x' at time t' to arrive at position x at time t

X, Y: coordinate of the grid center

x', y' : center line of trajectory

σ : standard deviation of trajectories at two directions which are assumed to grow with time $\sigma_x(t') = \sigma_y(t') = at'$ (a is a dispersion speed, equals to 5.4 km/h)

The potential mass transfer potential field for a given trajectory l, arriving at time t,

$\bar{T}_l(x|x')$, is integrated over back trajectory time τ :

$$\bar{T}_l(x|x') = \frac{\int_{t-\tau}^t Q(x, t|x', t') dt'}{\int_{t-\tau}^t dt'}$$

The concentration-weighted mass transfer potential field is calculated as:

$$\tilde{T}(x|x') = \sum_{l=1}^L \bar{T}_l(x|x') c_l$$

c_l : receptor concentration of back trajectory l

L: total number of back trajectories

Then final SQTBA field is obtained as follow:

$$SQTBA(x|x') = \frac{\tilde{T}(x|x')}{\sum_{l=1}^L \bar{T}_l(x|x')}$$

3.7 Potential Source Contribution Function (PSCF)

The PSCF model calculates conditional probabilities (if the value is close to 1 means the trajectories cross the grid cell effectively transport the emitted contaminant to the receptor site) to identify the source contribution of each grid to receptor site ^{1, 20, 21}. The contribution function is determined by a criterion value C (the parameter is adjustable in TraPSA), which is arbitrarily set (usually 50%~90% of the highest concentration). The contribution value of the endpoint on a back trajectory with receptor concentration larger than C is 1, otherwise the contribution value is 0:

$$p_l = \begin{cases} 1 & c_l \geq C \\ 0 & c_l < C \end{cases}$$

p_l : result of conditional value for trajectory l

c_l : receptor concentration of back trajectory l

The PSCF values are calculated for every grid as follow:

$$PSCF_{i,j} = \frac{\sum_{l=1}^L p_l \tau_{i,j,l}}{\sum_{l=1}^L \tau_{i,j,l}}$$

$PSCF_{i,j}$: PSCF value of grid i, j (latitude, longitude)

L: total number of back trajectories

$\tau_{i,j,l}$: endpoint number of back trajectory l in grid i, j

The PSCF model can't be directly applied into multi-site analysis as a fixed criterion value is used. The multi-site PSCF model is defined as ^{3, 22}:

$$p_{s,l} = \begin{cases} 1 & c_{s,l} \geq C_s \\ 0 & c_{s,l} < C_s \end{cases}$$

C_s : criterion value of site s

$c_{s,l}$: receptor concentration of back trajectory l from site s

$p_{s,l}$: result of conditional value for trajectory l

$$PSCF_{i,j} = \frac{\sum_{s=1}^S \sum_{l=1}^{L_s} p_{s,l} \tau_{i,j,s,l}}{\sum_{s=1}^S \sum_{l=1}^{L_s} \tau_{i,j,s,l}}$$

$PSCF_{i,j}$: PSCF value of grid i, j (latitude, longitude)

S: total site number

L_s : total number of back trajectories from site s

$\tau_{i,j,s,l}$: endpoint number of back trajectory l from site s in grid i, j

3.8 Weight Function

When cells are crossed by small number of trajectories, false source areas maybe identified if some of the trajectories also pass real through actual source areas; this phenomenon is called the “trailing effect”^{5, 7, 23}. A weighting function can be used to avoid the “trailing effect” defined as follows:

$$W_{i,j} = v_{i,j} \times \text{Weigth} \begin{cases} p1 & n_{i,j} < l_1 \times \text{avg} \\ p2 & l_1 \times \text{avg} < n_{i,j} < l_2 \times \text{avg} \\ \dots & \dots \\ pk & l_{k-1} \times \text{avg} < n_{i,j} < l_k \times \text{avg} \end{cases}$$

$W_{i,j}$: weighted value of grid i, j (latitude, longitude)

$v_{i,j}$: value of grid i, j

$n_{i,j}$: endpoint number of grid i, j

k: weight levels

l: weight level range

$n_{i,j}$: endpoint number of grid i, j

avg: average endpoint number of grid cell, avg = total endpoints/ grid number

pk: weight factor of level k

I_k : upper range of level k

3.9 Event analysis

Some special events (highest or lowest) will be important to understand the potential source of pollutants, TraPSA allows users to pick any time period for event analysis. TraPSA also provides a function for automatically determining highest or lowest concentration events. A duration time input is required, then average concentrations in any duration time periods will be calculated and sorted to determine the highest or lowest concentration events.

3.10 Grid Calculation

TraPSA provides a basic grid focalization and modification function, calculated as follow:

$$G_{i,j} = \frac{\sum \forall (g \neq no \ data)}{9}$$

g: grids directly next to grid i, j

TraPSA also provides a basic grid peak identification function. One grid will be identified as a potential peak if its value is larger than any other surrounding grids. Then the potential peaks will be sorted to determine the highest peaks (the displayed peak number is adjustable in TraPSA).

4. Data Required

4.1 Meteorological data

Meteorological data is required by TraPSA for generating HYSPLIT trajectories. Various meteorological data formats are available as long as they are recognizable by HYSPLIT_4, although the HYSPLIT_4 suite contains numerous converters that users can use to convert data into HYSPLIT format. Typically, the Eta Data Assimilation System (EDAS) with the 40 km resolution data and North American Regional Reanalysis (NARR) data are used for North American locations.^{3, 17} The Global Data Assimilation System (GDAS) and the NCEP/NCAR Reanalysis Archive have been used for sites in Europe and Asia.^{2, 24-26} In addition, NOAA ARL maintains an archive of meteorological datasets called HYSPLIT format that can be used to drive the HYSPLIT trajectory model (as well as other HYSPLIT executables).

These meteorological data can be download at these sites:

EDAS: <ftp://arlftp.arlhq.noaa.gov/archives/edas40/>

NARR: <ftp://arlftp.arlhq.noaa.gov/narr/>

GDAS: <ftp://arlftp.arlhq.noaa.gov/archives/gdas1/>

NCEP/NCAR Reanalysis: <ftp://arlftp.arlhq.noaa.gov/archives/reanalysis/>

Other meteorological data: <http://nomads.ncdc.noaa.gov/data.php>

4.2 Map Data

ESRI shapefiles is required to display maps in TraPSA. Two GIS data formats, ESRI shapefile (‘.shp’, ‘.dbf’ and ‘.shx’) and GeoTIFF file (‘.tif’) are available for used by TraPSA. Shapefiles of related regions should be downloaded before analysis. Note that it will take a longer time to display more specified shapefiles. Note that it will take a longer time to display the shapefiles with more geographic information.

Country administrative boundaries shapefile are available at:

<http://www.gadm.org/country>

4.3 Measurement Data

Pollutants measurement data are required for receptor models. The data file format can be ‘.csv’, ‘.xlsx’ or ‘.xls’. The data should include species names, units, collection date and concentration and organized as shown in the following figure. Note that collection date format is flexible as long as it can be identified by TraPSA, but **must include year, month, day and hour information**. TraPSA will handle the date format, missing data and replicate data automatically.

	A	B	C	D
1	Name	PBM	GOM	GEM
2	Unit	pg/m3	pg/m3	ng/m3
3	Time			
4	Data			
3	12/31/2012 23:05	14.66	2.31	1.539
4	1/1/2013 2:05	11.12	2.408	1.594
5	1/1/2013 5:05	8.988	3.71	1.645
6	1/1/2013 8:05	8.754	3.326	1.637
7	1/1/2013 11:05	8.039	2.188	1.551
8	1/1/2013 14:05	8.332	1.448	1.488
9	1/1/2013 17:05	5.688	0.986	1.418
10	1/1/2013 21:05	4.195	0.511	1.371
11	1/2/2013 0:05	5.662	0.422	1.354
12	1/2/2013 3:05	5.033	0.471	1.335
13	1/2/2013 6:05	5.477	0.667	1.366
14	1/2/2013 9:05	5.008	0.81	1.388
15	1/2/2013 16:05	-9	-9	1.739
16	1/2/2013 19:05	8.144	5.922	1.662
17	1/2/2013 22:05	7.986	5.434	1.666
18	1/3/2013 1:05	6.599	4.646	1.627

Figure 4.1 Input Measurement Data format

5. Example

5.1 Case Introduction

Atmospheric mercury data, including gaseous elemental mercury (GEM), gaseous oxidized mercury (GOM) and particulate-bound mercury (PBM), collected as part of the Atmospheric Mercury Network (AMNet) at two monitoring sites, Huntington Forest (44.0, -74.2) and Bronx (40.9, -73.9), and data from 2013 to 2014 were used in this example. The NCEP/NCAR Reanalysis Data Archive was used to obtain the meteorological data for generating trajectory endpoints.

5.2 Software Installation

Install TraPSA software following the install wizard. The wizard will install an installation folder containing TraPSA and MATLAB runtime (MATLAB Runtime will be automatically installed if it has not been previously installed). It is fastest to download the installation package with MATLAB Runtime included. If the installation package without MATLAB Runtime is downloaded and MATLAB Runtime has never been installed on the computer, it will take several minutes to download MATLAB Runtime alone (about 440 MB). Note: to avoid unpredictable errors, do not install TraPSA into "C:\Program Files\" and do not use a shortcut.

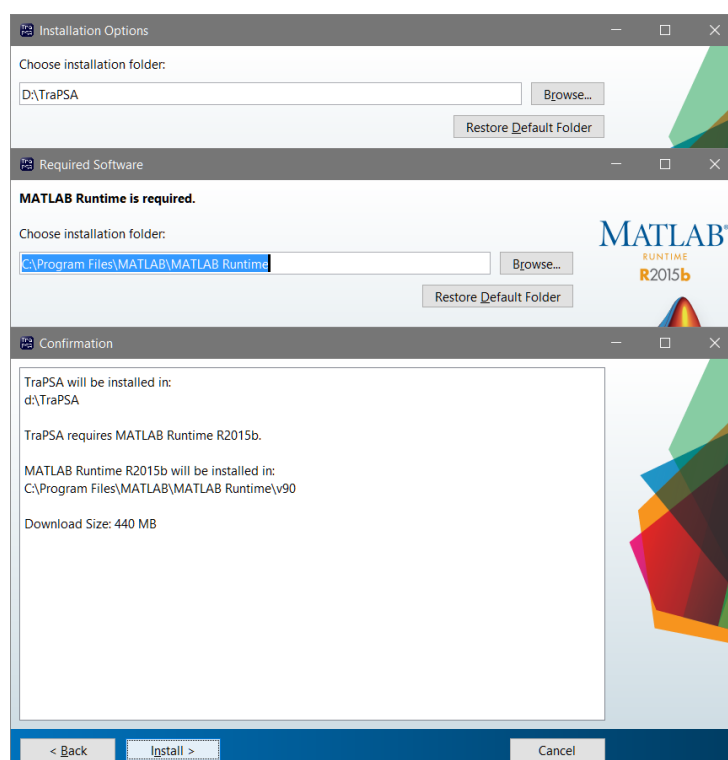


Figure 5.1 Installation TraPSA and MATLAB Runtime

5.3 Before Analysis

Path Setting:

1. Click **File** → **Path Setting** to open the path setting window.
2. Set the path folder for meteorological data, HYSPLIT (if it is installed), map data and output.
3. Set default map center.
4. Click **Confirm** to finish path setting.
5. There are no default values for the settings above so they must be specified.

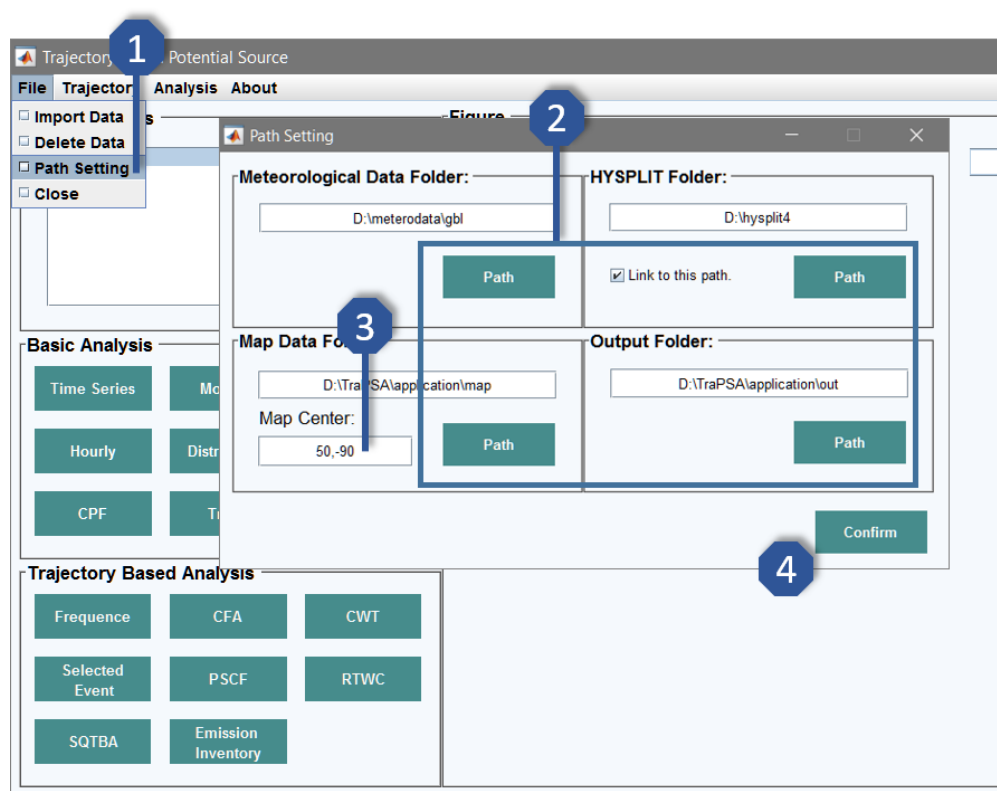


Figure 5.2 Path Setting

Parameters Setting:

1. Click **Analysis** → **Setting** → **Grid** to set up the grid cell size of the trajectory ensemble models, unit is degrees. TraPSA currently only supports uniform grid size. The default value of a grid is 0.5 degrees.
2. Click **Analysis** → **Setting** → **Weight** to open the grid weighting set-up window. Enter weight level number, click **v**.
3. Select weight level, enter weight range and weight factor, click **±** to confirm current weight level.

4. Select methods to apply weighting, click **Apply**. No default weighting values are supplied.
5. Click **Analysis** → **Setting** → **Parameter** to set up parameters used in the analysis, CPF criterion and minimum CPF wind speed (applied in CPF analysis, default values are 0.8 and 5 km/h), PSCF criterion (applied in PSCF analysis, default value is 0.8), peak number (applied in grid peak identification, controls the maximum displayed peak number, default value is 10) and map projection method (Robinson or Mercator, default is Robinson).

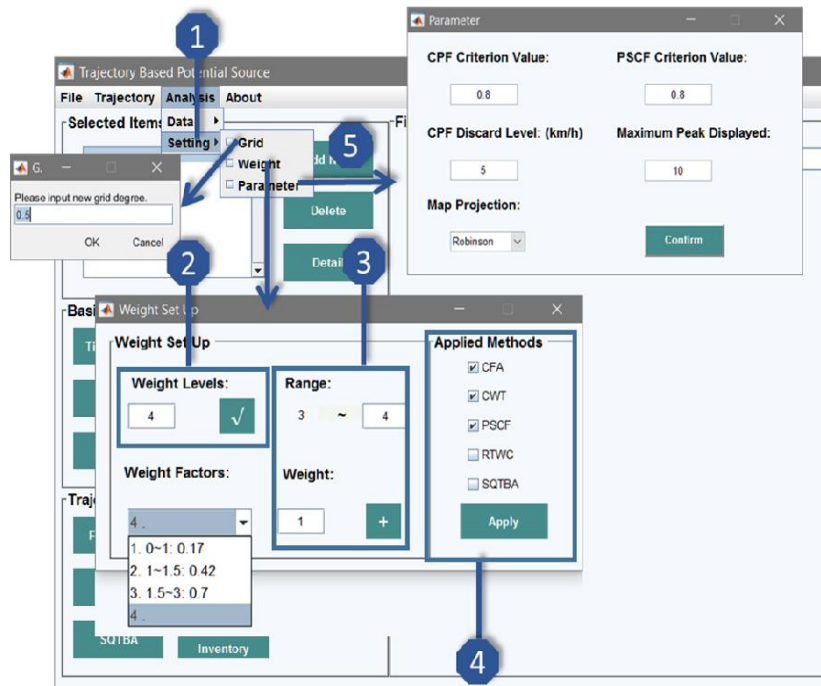


Figure 5.3 Parameter Setting

5.4 Monitoring Site Database

Database Establishment and Updating:

The site database will be established when pollutant data are imported for the first time. Trajectory data will be imported into the database after they are generated.

1. Click **File** → **Import Data** to open the import window.
2. Enter site information, including site name, site location, missing data information and data file path. If the pollutant data need to be updated, please enter the identical site name.
3. Copy the date from data file and **v** to make sure the date format is identifiable.
4. Click **Import**. Site information will change after the data is successfully imported or updated. Right click to see detailed information if the panel is not big enough to display all the information.

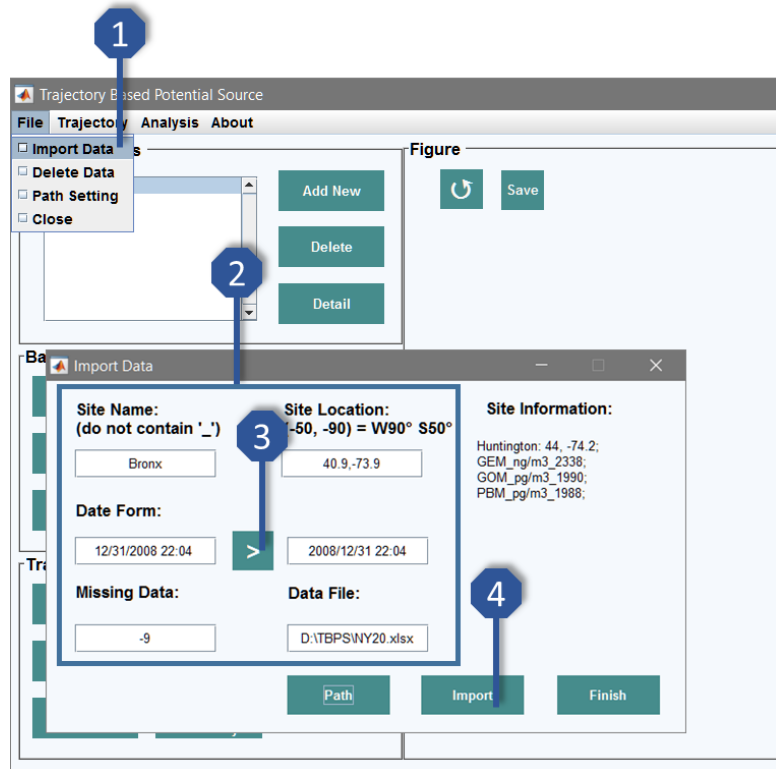


Figure 5.4 Site Database Establishment and Updating

Deleting Data:

1. Click **File** → **Delete Data** to open the delete data window.
2. Select site, species or trajectory items that need to be removed from the database, then click **Delete**.

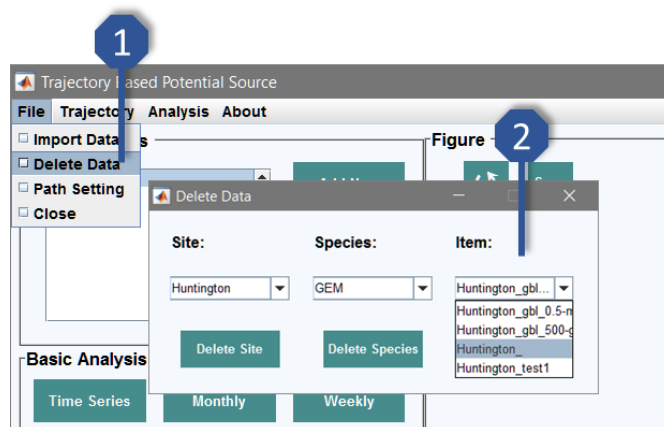


Figure 5.5 Deleting Data

Exporting Export:

1. Click **Analysis** → **Data** → **Pollutants** to show all the data in the pollutant database.
2. Click **Export** to generate an excel file containing the selected site pollutant data.
3. Click **Analysis** → **Data** → **Trajectory** to show the trajectory endpoints database.
4. Click **Export** to generate an excel file containing the selected trajectory item.
5. Click **Export Shapefile** to generate a shapefile for the selected trajectory item.

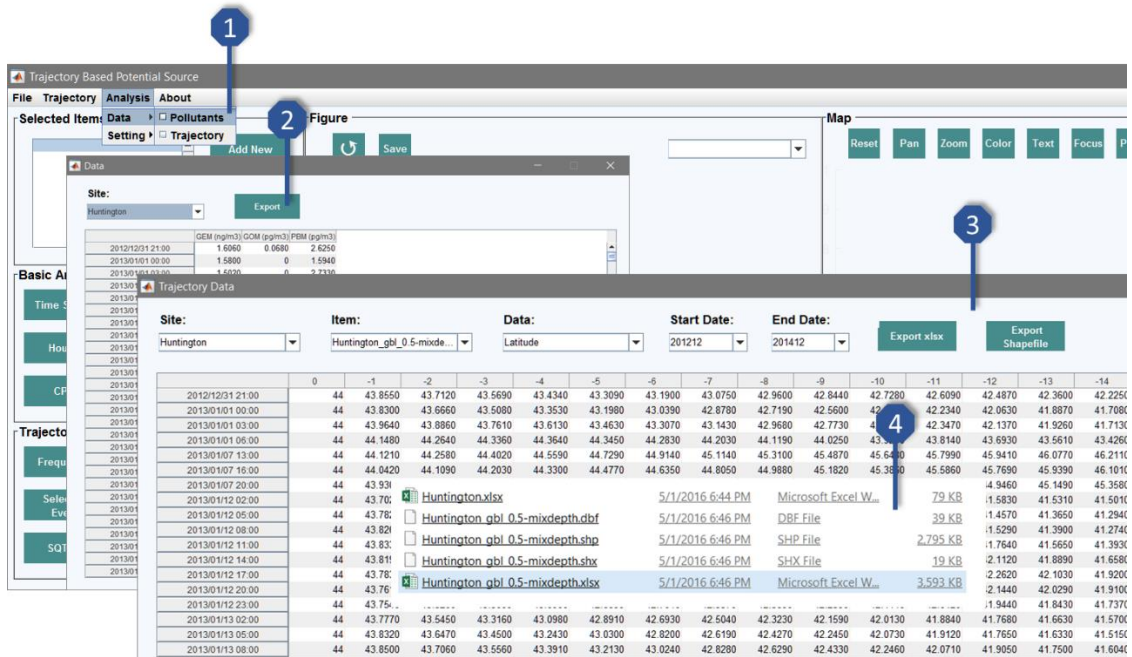


Figure 5.6 Exporting Data

5.5 Generating Trajectories

Set Up Trajectories:

1. Click **Trajectory** → **Set Up** to open the trajectory set-up window.
2. Set trajectory parameters, including site, start-end date, starting height and trajectory time in the left panel. Three different units can be used for trajectory height. Note that the height number should be < 1 if the fraction of mixed layer is used as the height unit. The unit is meter for other two options. The functions of two options will be introduced in the next section.
3. Set up meteorological data information in the right panel. **Name** is an arbitrary string used for labeling the trajectory item. The meteorological data filename is separated into three parts, date string, string before the date and string after the date. If more than one date file exists for one month, use '#' instead of the incrementing number.
4. Click **Check** to determine if the time period has been analyzed using the current setting to avoid repeated calculations. Click **Confirm** to generate a batch file for trajectory calculations.

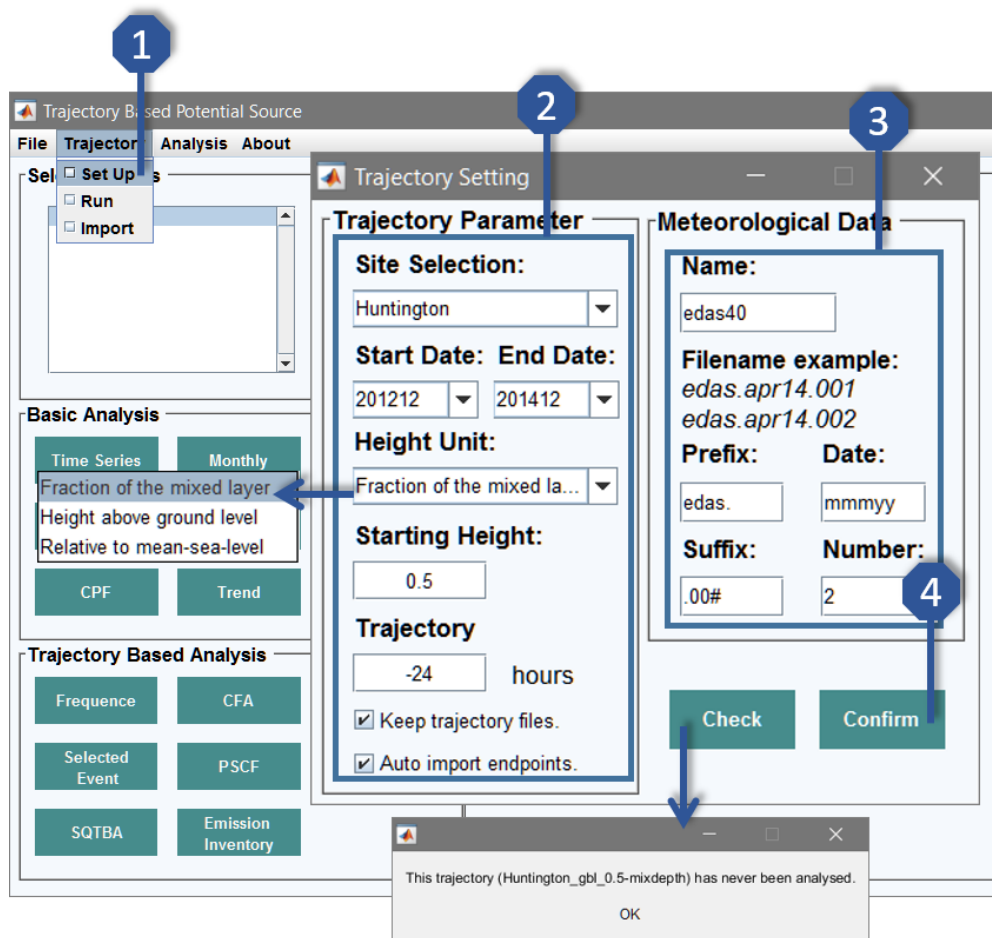


Figure 5.7 Trajectory Setting

Run Trajectories:

1. Click **Trajectory** → **Run** to start HYSPLIT trajectory analysis. This step may take a relatively long time depending on the number of data points, back trajectory time, meteorological data and computer configuration. It generally takes 3 to 10 minutes to run 3000 data points for 72 h back trajectories.
2. If option **Auto import endpoints** is selected, the endpoints files generated by HYSPLIT will be automatically imported into TraPSA.
3. If option **Keep trajectory files** is not selected, the original files generated by HYSPLIT will be automatically deleted from the output folder.

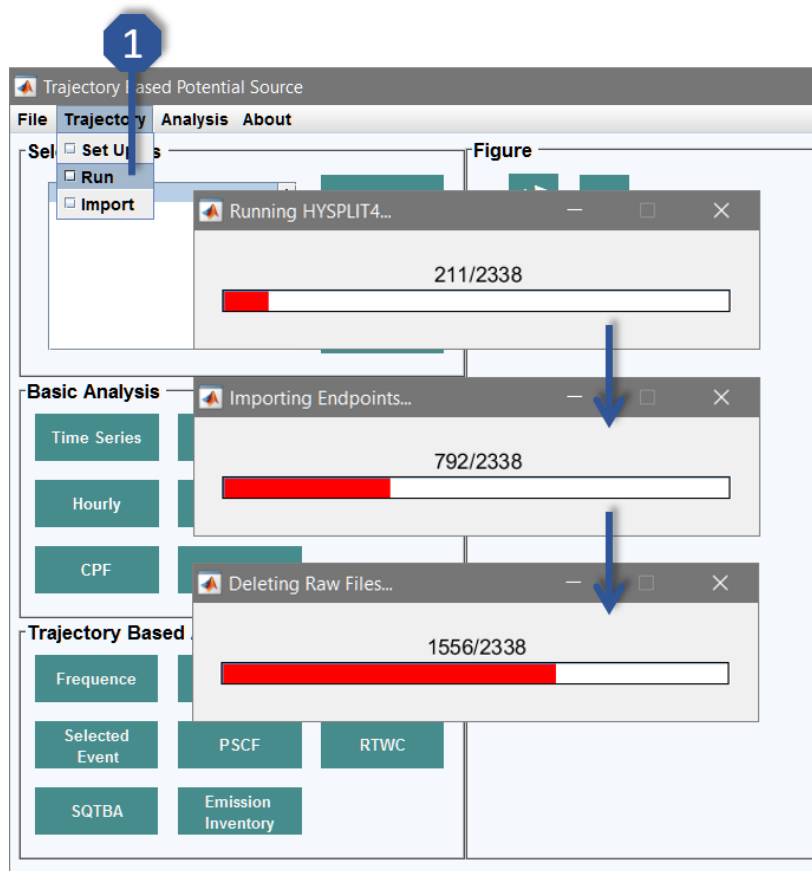


Figure 5.8 Generating Trajectories

Importing Trajectories:

It is highly recommended that TraPSA be used to generate trajectory endpoints as it can perform trajectory analysis in a smart and fast way. However, TraPSA also supports importing original HYSPLIT output files. These endpoints files must have identical date information corresponding to the concentration data in the database, otherwise TraPSA cannot identify the files.

1. Click **Trajectory** → **Import** to import trajectory files generated by TraPSA if the option **Auto import endpoints** is not applied, or to import already existing HYSPLIT output files.
2. Select the site and trajectory item that needs to be imported, or click **Add New** to create a new trajectory item for importing.
3. Click **Select Files** to select the endpoint file, then click **Import**. Note that the trajectory time should be the maximum back trajectory time of all the selected files.

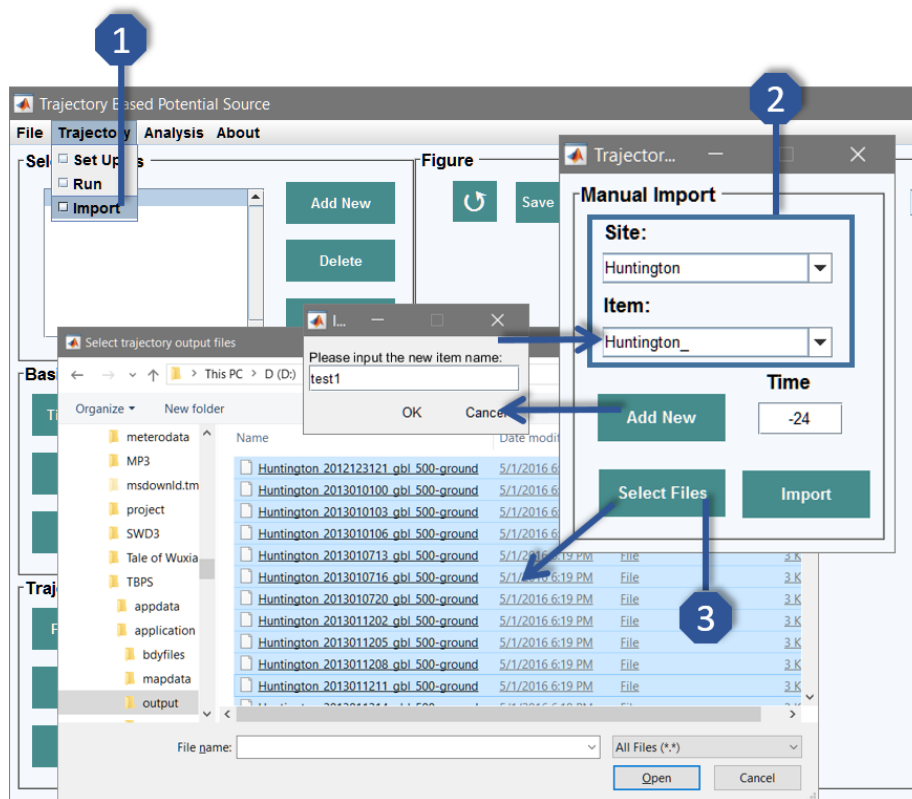


Figure 5.9 Manually Importing Trajectories

5.6 Item to be Analyzed

1. Click **Add New** to add items to be analyzed.
2. The item to be analyzed will be extracted from the database which includes one pollutant species, one trajectory item, time periods (any specific time points or by season, weekdays or day period) and normalization method. These features can be selected in the corresponding panels.
3. A time period without a trajectory item is not allowed to be added. Also make sure the selected trajectory time is less than the maximum back trajectory time. Then click **Add**.
4. Name the featured item to be analyzed with arbitrary string, click **OK**. Then add another item.
5. Click **Finish** to close the window when all items are added. These items will be shown in **Selected Items**.

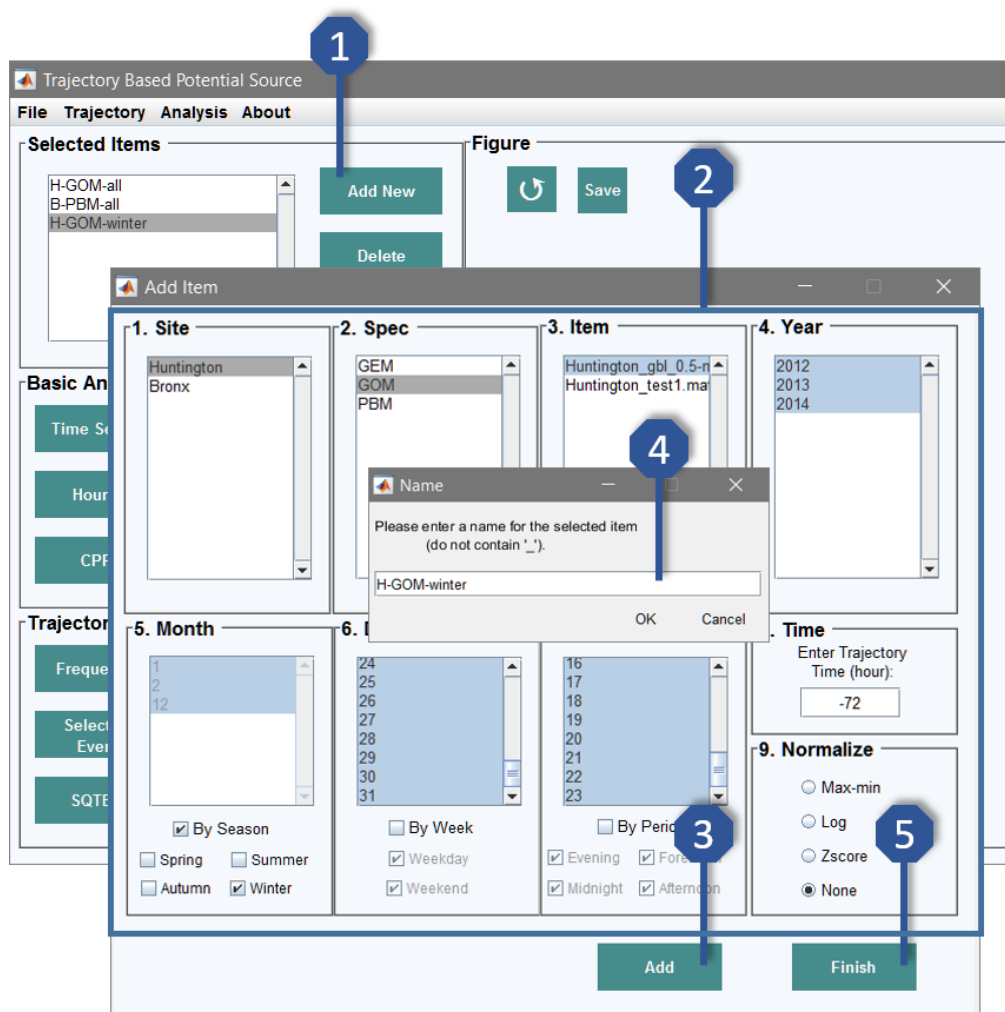


Figure 5.10 Select Item to be analyzed

5.7 Basic Analysis

1. Click **Select Items**. Multiple options are allowed; however, no multiple-site methods can be used in basic analysis. Therefore, all the basic analysis will be carried out and displayed separately.
2. Click any basic method, and the results of selected items will be displayed in the figure panel.
3. Click the **pop-up menu** in the figure panel to switch figures. **Right click** on the pop-up menu to delete the current figure.
4. Click **Save** in the figure panel to save the current figure. The figure will be saved as a '.png' file and associated data will be saved as a text file in the output folder.

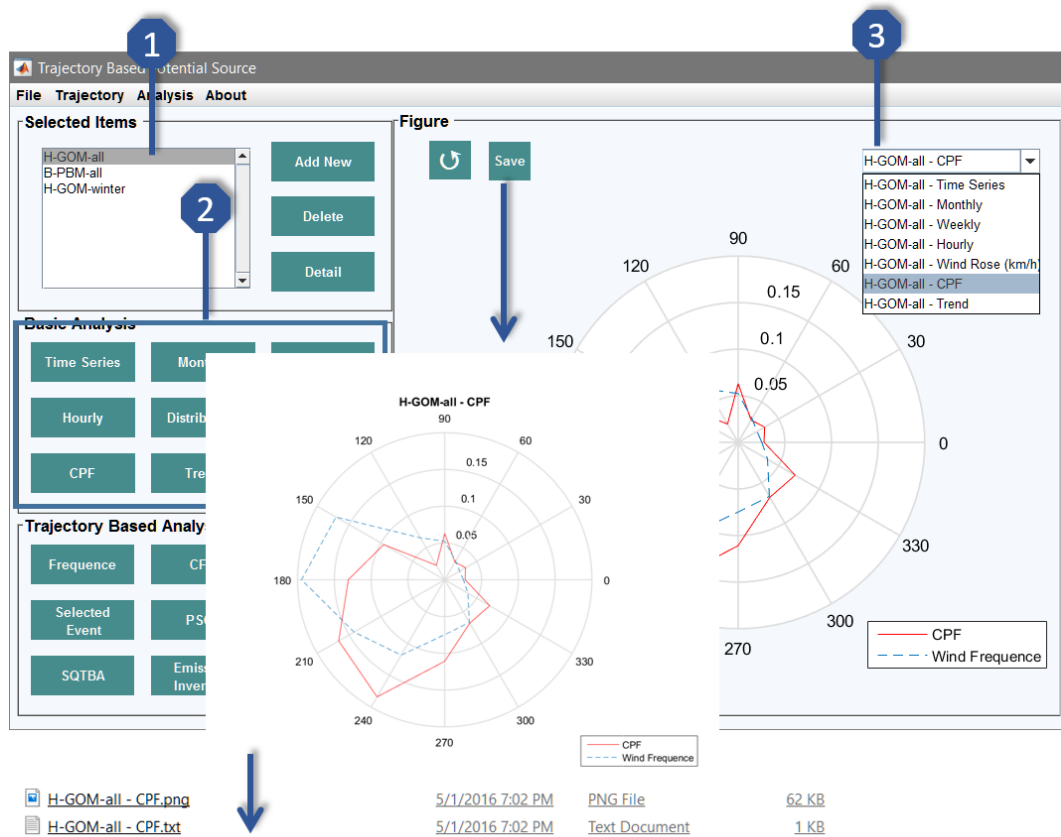


Figure 5.11 Basic Analysis

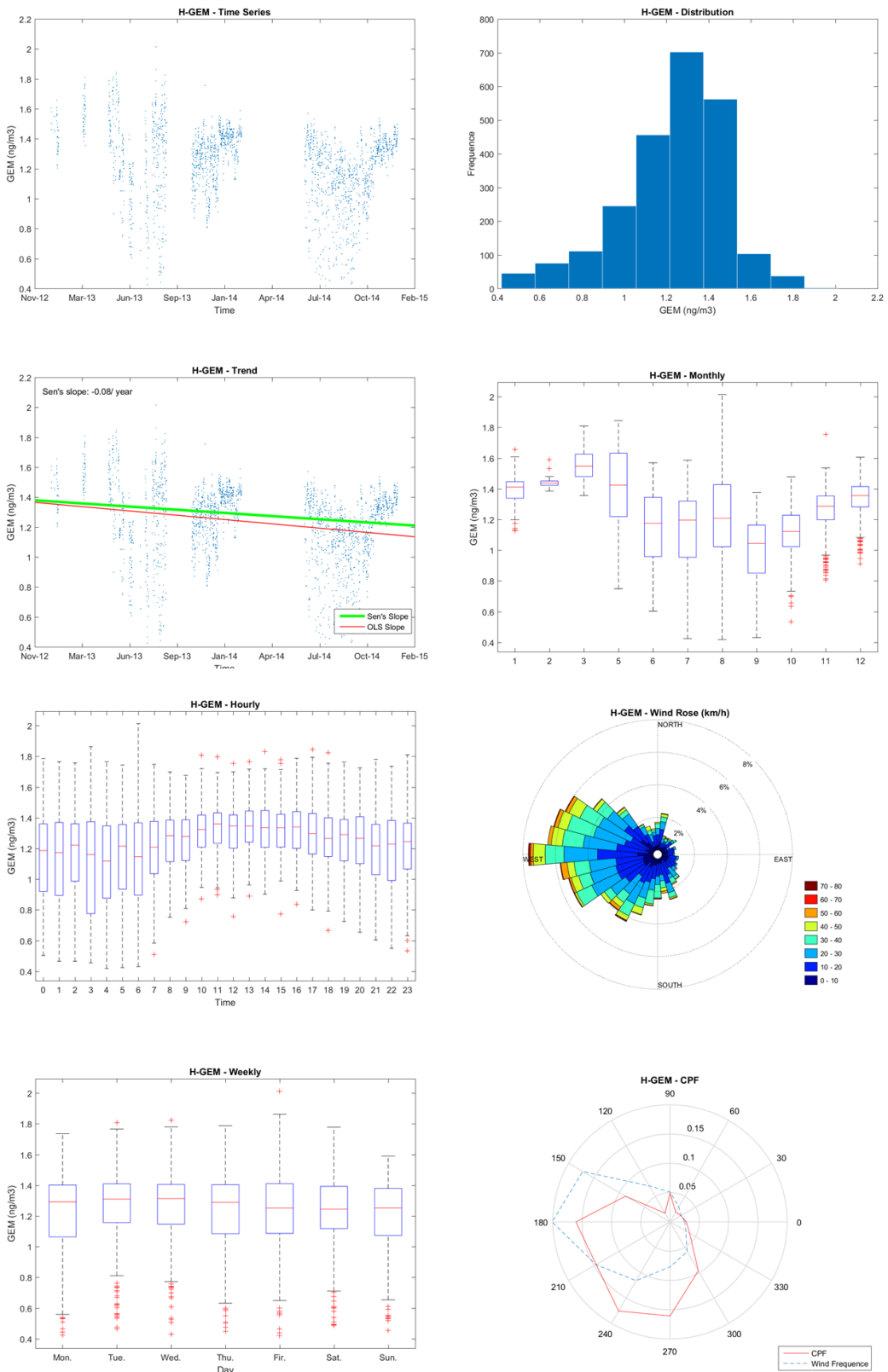


Figure 5.12 Example Results for Basic Methods

5.8 Trajectory Ensemble Receptor Models

1. Click **Select Items**. Multiple selections are allowed for trajectory based analysis.
2. Click any trajectory ensemble method. If multi-site is selected, a dialogue box will be displayed, click **No** to apply multi-site methods.
3. The results of trajectory ensemble model will be created as a new layer on the map panel.

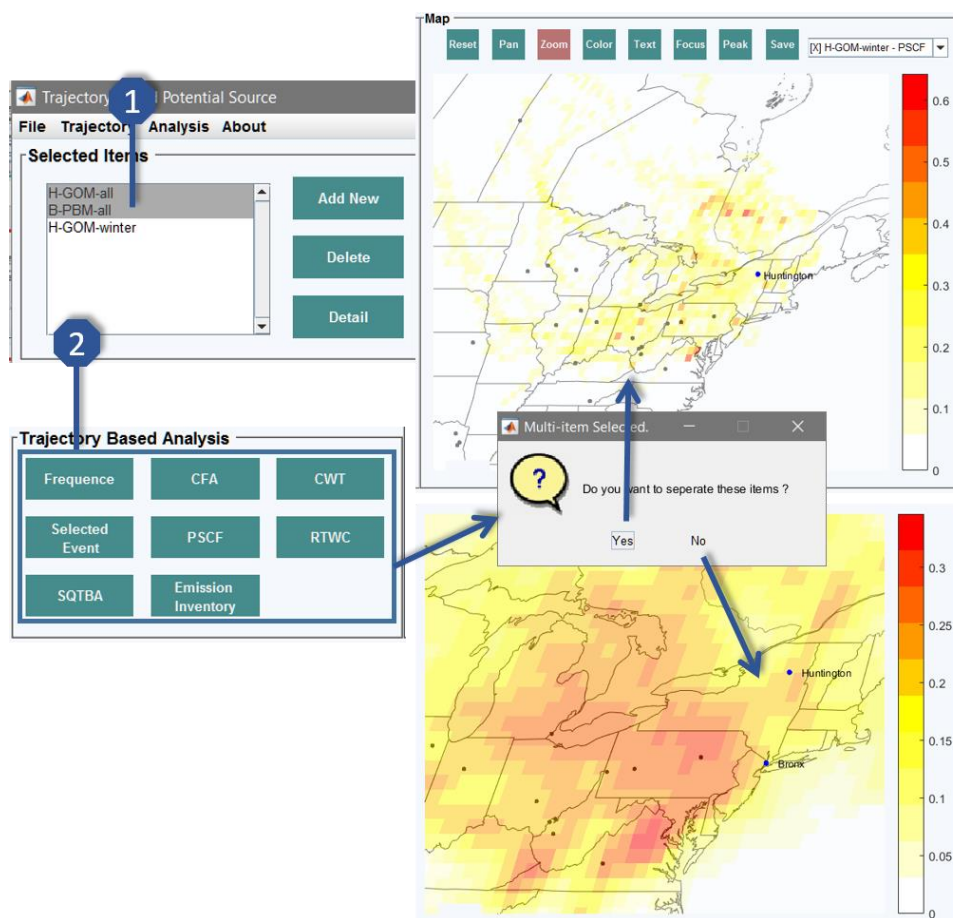


Figure 5.13 Trajectory Ensemble Analysis

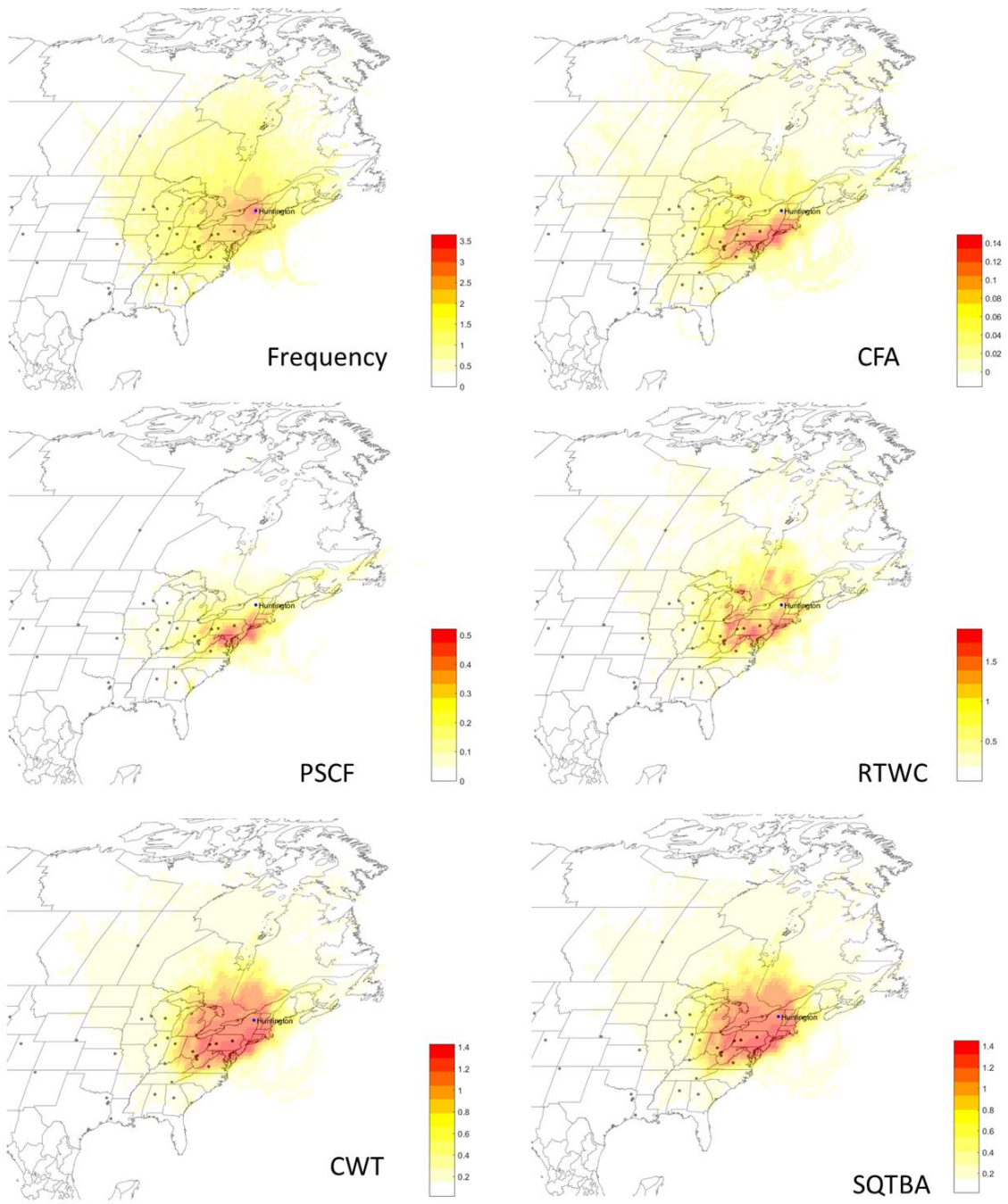


Figure 5.14 Example Results for Trajectory Based Methods

5.9 Event Analysis

1. Click **Selected Event** to open the event selection window.
2. Any time point can be selected as an event. Highest or lowest events can be selected using the special events determination function. Click **v** to add events. **Right click** could delete the selected event.
3. Click **Confirm** to proceed.
4. Three figures will be created as shown below: trajectory lines will be created as a new layer on the map panel; the data points, which are selected as events, will be marked on the corresponding time series (the time series figure will be created if it does not exist); 3D trajectory lines will be generated.



Figure 5.15 Event Analysis

5.10 Emissions Inventory

1. Click **Emissions Inventory** to import the inventory file.
2. The import file should be organized in the format shown below.
3. The emissions inventory map will be created as a new layer on the map panel.

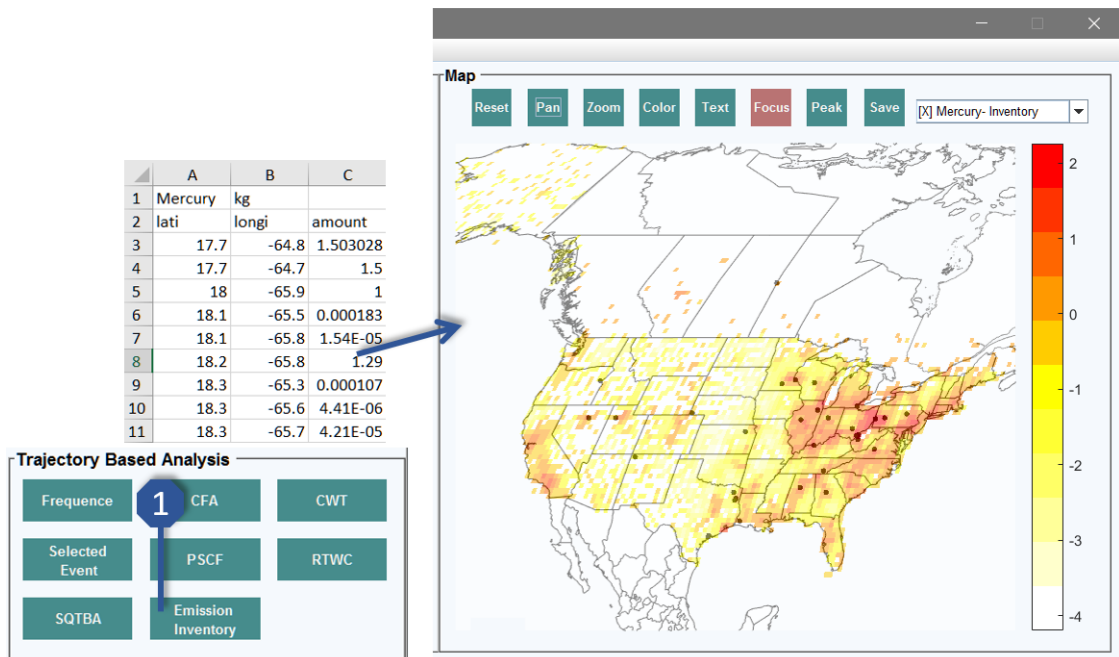


Figure 5.16 Emission Inventory

5.11 GIS Editing

Objects Editing:

1. The **Pop-up menu** on the right top of the map can be used to switch between map layers. **[X]** indicates the status of the layer is **on**, the layer is displayed on the map; **[O]** indicates the status of the layer is **off**, the layer is not displayed on the map. Left click to switch the status, **right click** to delete the layer. Note that the items will be stacked on the top when the status is changed from off to on.
2. Click **Reset** to re-draw the map; click **Pan** to move the displayed map; click **Zoom** to zoom in and zoom out; click **Color** to adjust the color bar.
3. Click **Text** to add text or marker at a specific point on the map. TraPSA allows the users to arbitrary determine the position of text or point if the **location** is empty. The marker will not be created if **Marker** is **n**.
4. Click any item on the top of the map to adjust its properties. The possible adjustments are shown below.

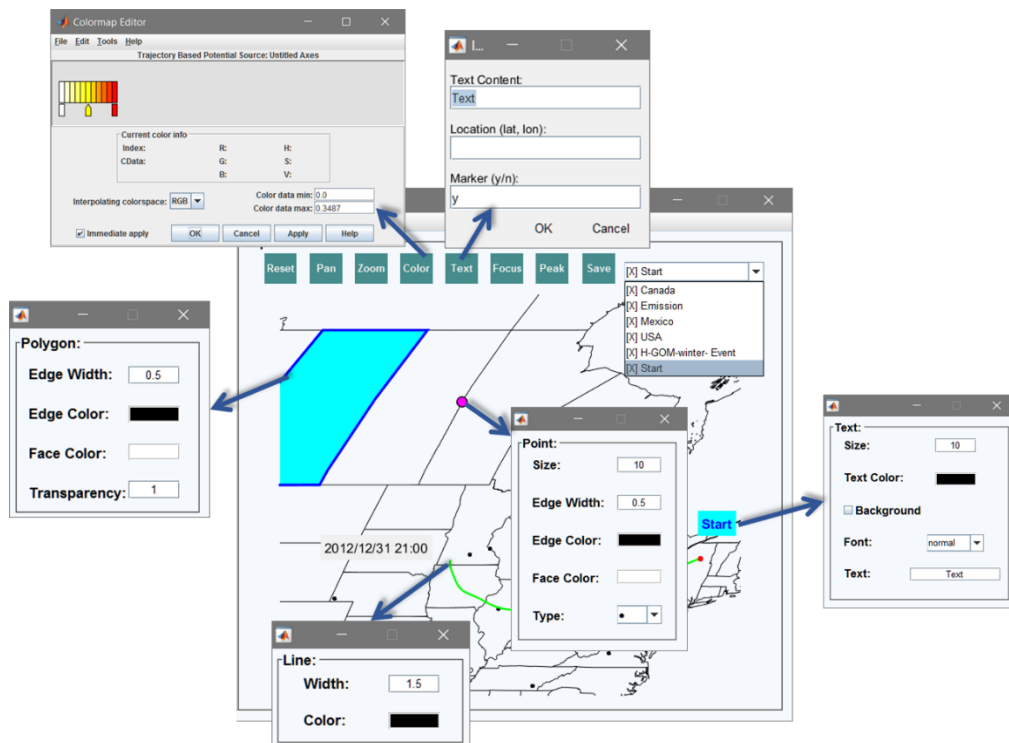


Figure 5.17 GIS Objects Editing

Grid Analysis:

Click **Focus** to apply grid smoothing; click **Peak** to identify and display the peak of the current grid. The results after applying Focus and Peak are shown below.

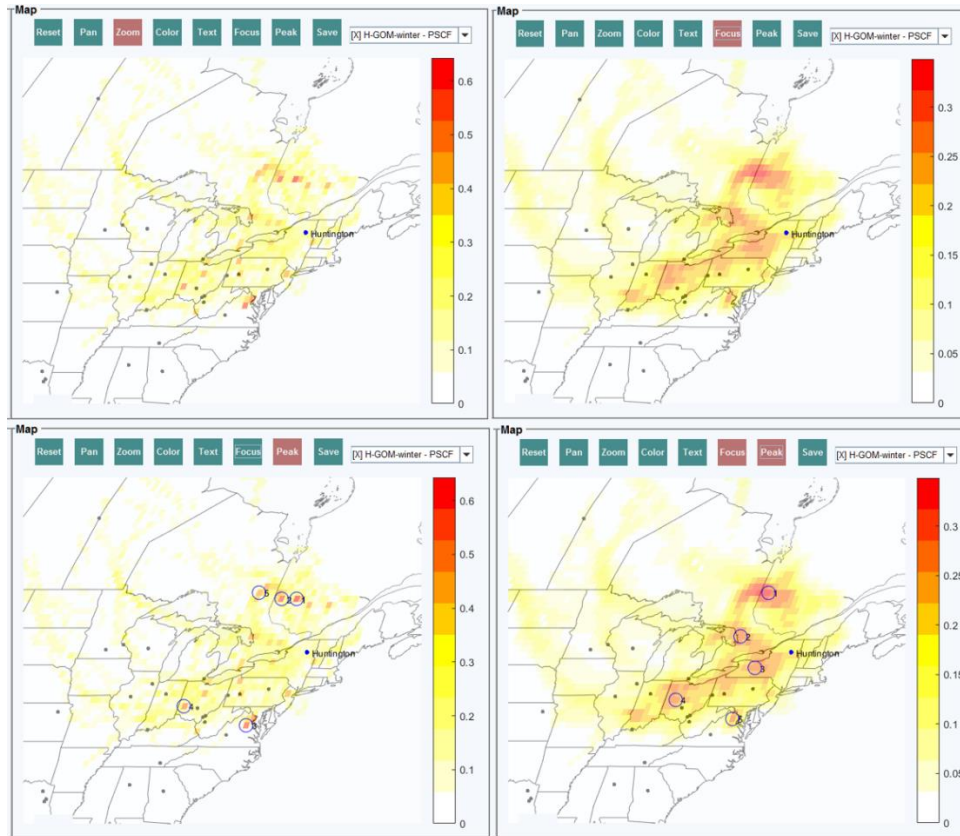


Figure 5.18 Grid Analysis

5.12A Final Result

A final result of multi-site PSCF analysis using Bronx and Huntington GOM data is shown below. Click **Save** on the map panel to generate a '.png' file for the currently displayed map. The displayed grid (including original and smoothed grid) will be saved as a Geo TIFF file ('.tif'). The displayed trajectory line will be saved as one shapefile ('.shp'), and the displayed peak point will be saved as a text file that includes latitude, longitude and intensity.

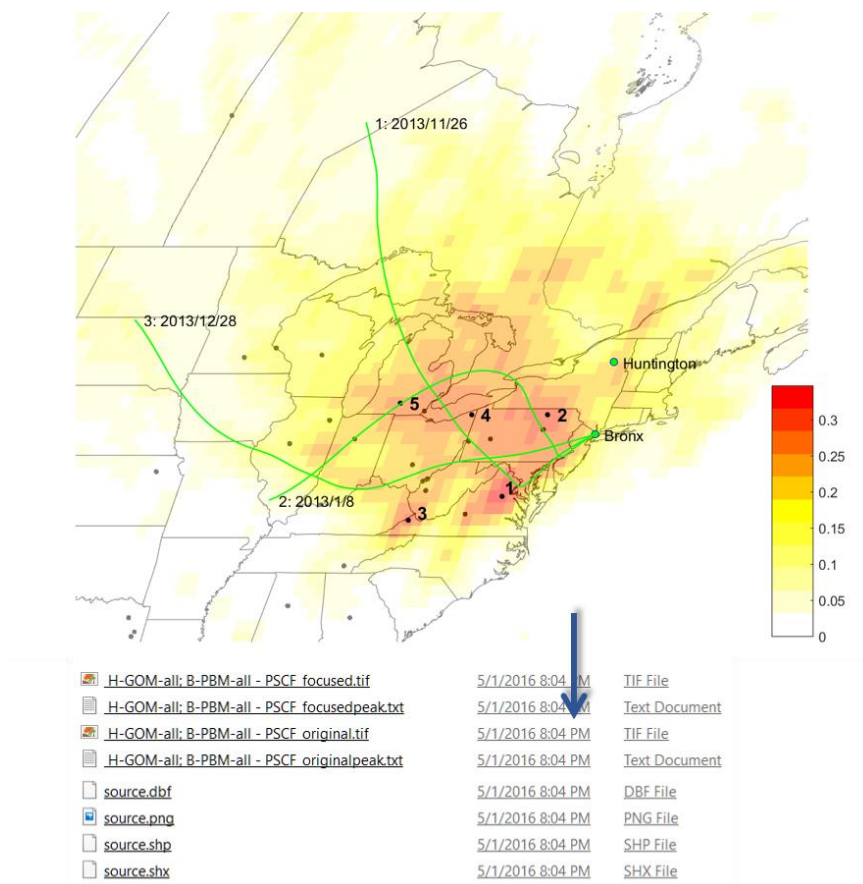


Figure 5.19 Multi-site PSCF Analysis Result of GOM data at Bronx and Huntington

6. Reference

1. Choi, H. D.; Holsen, T. M.; Hopke, P. K., Atmospheric mercury (Hg) in the Adirondacks: Concentrations and sources. *Environmental Science & Technology* **2008**, *42*, (15), 5644-5653.
2. Squizzato, S.; Masiol, M., Application of meteorology-based methods to determine local and external contributions to particulate matter pollution: A case study in Venice (Italy). *Atmospheric Environment* **2015**, *119*, 69-81.
3. Liu, W.; Hopke, P. K.; Han, Y.-j.; Yi, S.-M.; Holsen, T. M.; Cybart, S.; Kozlowski, K.; Milligan, M., Application of receptor modeling to atmospheric constituents at Potsdam and Stockton, NY. *Atmospheric Environment* **2003**, *37*, (36), 4997-5007.
4. Han, Y.-J.; Holsen, T. M.; Lai, S.-O.; Hopke, P. K.; Yi, S.-M.; Liu, W.; Pagano, J.; Falanga, L.; Milligan, M.; Andolina, C., Atmospheric gaseous mercury concentrations in New York State: relationships with meteorological data and other pollutants. *Atmospheric Environment* **2004**, *38*, (37), 6431-6446.
5. Zhou, L., Comparison of two trajectory based models for locating particle sources for two rural New York sites. *Atmospheric Environment* **2004**, *38*, (13), 1955-1963.
6. Keeler, G. J.; Samson, P. J., SPATIAL REPRESENTATIVENESS OF TRACE-ELEMENT RATIOS. *Environmental Science & Technology* **1989**, *23*, (11), 1358-1364.
7. Cheng, I.; Zhang, L.; Blanchard, P.; Dalziel, J.; Tordon, R., Concentration-weighted trajectory approach to identifying potential sources of speciated atmospheric mercury at an urban coastal site in Nova Scotia, Canada. *Atmospheric Chemistry and Physics* **2013**, *13*, (12), 6031-6048.
8. Stohl, A., Trajectory statistics-A new method to establish source-receptor relationships of air pollutants and its application to the transport of particulate sulfate in Europe. *Atmospheric Environment* **1996**, *30*, (4), 579-587.
9. Rutter, A. P.; Snyder, D. C.; Stone, E. A.; Schauer, J. J.; Gonzalez-Abraham, R.; Molina, L. T.; Marquez, C.; Cardenas, B.; de Foy, B., In situ measurements of speciated atmospheric mercury and the identification of source regions in the Mexico City Metropolitan Area. *Atmospheric Chemistry and Physics* **2009**, *9*, (1), 207-220.
10. Helsel, D.; Hirsch, R., Statistical methods in water resources—Hydrologic analysis and interpretation: US Geological Survey Techniques of Water-Resources Investigations, book 4, chap. In A3: 2002.
11. Tabari, H.; Aghajanloo, M.-B., Temporal pattern of aridity index in Iran with considering precipitation and evapotranspiration trends. *International Journal of Climatology* **2013**, *33*, (2), 396-409.
12. Bhavsar, S. P.; Gewurtz, S. B.; McGoldrick, D. J.; Keir, M. J.; Backus, S. M., Changes in mercury levels in Great Lakes fish between 1970s and 2007. *Environmental science & technology* **2010**, *44*, (9), 3273-3279.
13. Sen's slope.
14. Huang, J. Y.; Choi, H. D.; Hopke, P. K.; Holsen, T. M., Ambient Mercury Sources in Rochester, NY: Results from Principle Components Analysis (PCA) of Mercury Monitoring Network Data. *Environmental Science & Technology* **2010**, *44*, (22), 8441-8445.
15. Weiss-Penzias, P. S.; Gustin, M. S.; Lyman, S. N., Sources of gaseous oxidized mercury and mercury dry deposition at two southeastern U.S. sites. *Atmospheric Environment* **2011**, *45*, (27), 4569-4579.

16. Kabashnikov, V. P.; Chaikovsky, A. P.; Kucsera, T. L.; Metelskaya, N. S., Estimated accuracy of three common trajectory statistical methods. *Atmospheric Environment* **2011**, *45*, (31), 5425-5430.
17. Han, Y.-J.; Holsen, T. M.; Hopke, P. K., Estimation of source locations of total gaseous mercury measured in New York State using trajectory-based models. *Atmospheric Environment* **2007**, *41*, (28), 6033-6047.
18. Pérez, I. A.; Artuso, F.; Mahmud, M.; Kulshrestha, U.; Sánchez, M. L.; García, M. Á., Applications of Air Mass Trajectories. *Advances in Meteorology* **2015**, *2015*, 1-20.
19. Stein, A.; Draxler, R.; Rolph, G.; Stunder, B.; Cohen, M.; Ngan, F., NOAA's HYSPLIT atmospheric transport and dispersion modeling system. *Bulletin of the American Meteorological Society* **2015**, *96*, (12), 2059-2077.
20. Hopke, P. K., Recent developments in receptor modeling. *Journal of Chemometrics* **2003**, *17*, (5), 255-265.
21. Hopke, P. K., Review of receptor modeling methods for source apportionment. *J Air Waste Manag Assoc* **2016**, *66*, (3), 237-59.
22. Han, Y.-J.; Holsen, T. M.; Hopke, P. K.; Yi, S.-M., Comparison between back-trajectory based modeling and Lagrangian backward dispersion modeling for locating sources of reactive gaseous mercury. *Environmental science & technology* **2005**, *39*, (6), 1715-1723.
23. Cheng, I.; Xu, X.; Zhang, L., Overview of receptor-based source apportionment studies for speciated atmospheric mercury. *Atmospheric Chemistry and Physics* **2015**, *15*, (14), 7877-7895.
24. Fu, X. W.; Feng, X.; Liang, P.; Deliger; Zhang, H.; Ji, J.; Liu, P., Temporal trend and sources of speciated atmospheric mercury at Waliguan GAW station, Northwestern China. *Atmospheric Chemistry and Physics* **2012**, *12*, (4), 1951-1964.
25. Fu, X. W.; Feng, X.; Shang, L. H.; Wang, S. F.; Zhang, H., Two years of measurements of atmospheric total gaseous mercury (TGM) at a remote site in Mt. Changbai area, Northeastern China. *Atmospheric Chemistry and Physics* **2012**, *12*, (9), 4215-4226.
26. Lee, G.; Kim, P.; Han, Y.; Holsen, T.; Lee, S., Tracing Sources of Total Gaseous Mercury to Yongheung Island off the Coast of Korea. *Atmosphere* **2014**, *5*, (2), 273-291.