## IsoTrack and It’s Parameters

*Yaroslav Lyutvinskiy. 20.12.2018*

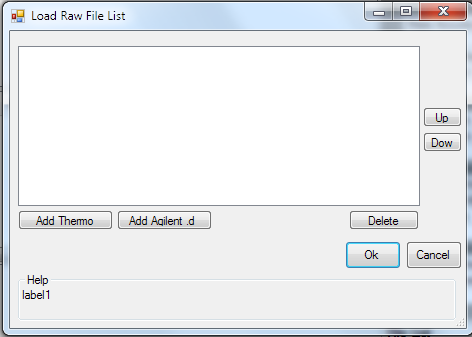
Here is a description of IsoTrack parameters and recommendations for short-pHilic (7 min gradient) and long-pHilic (19 min gradient) protocols. Not all the parameters have optimal default values, so I will provide some recommendations.

### General

**Processes** – By default, it is equal to number of processing cores of PC. Normally, shall not be changed.

### Input

**File List** – Is only input parameter for untargeted analysis. To fill “File List” press the button on “File List” string. Then will appear dialog box to form file list:

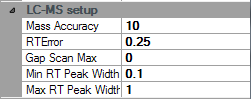
Here you can select to add to file list some Thermo or Agilent raw files. You should not mix in one file list Thermo and Agilent raw files. Also, you should not mix in single run files of negative and positive polarity together. Untargeted analysis designed to work well only with data contains no isotopic labeling. It makes low sense to mix together files where you expect different retention times for the same component.

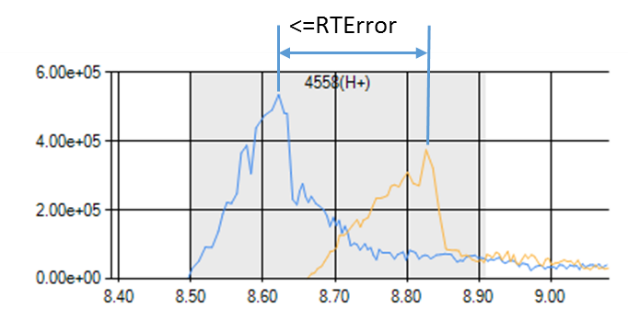
Summing together raw files selected for untargeted analysis shall be:

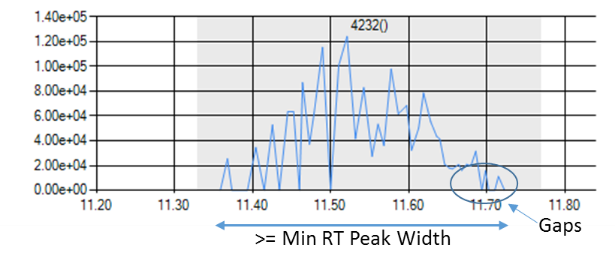
* Of the same vendor
* Of the same polarity
* No isotopically labeled data
* Of the same protocol and experiment

Order of raw file in the dialog box will define order of files in reports. Here you can change it with buttons “Up” and “Down”. Pressing these buttons will change position of selected files in the list up and down respectively.

### LC-MS setup

**Mass accuracy** – mass spectrometer mass accuracy in parts per million (ppm). It is normally 10 ppm for Orbitrap and 20 ppm for Agilent TOF-MS.

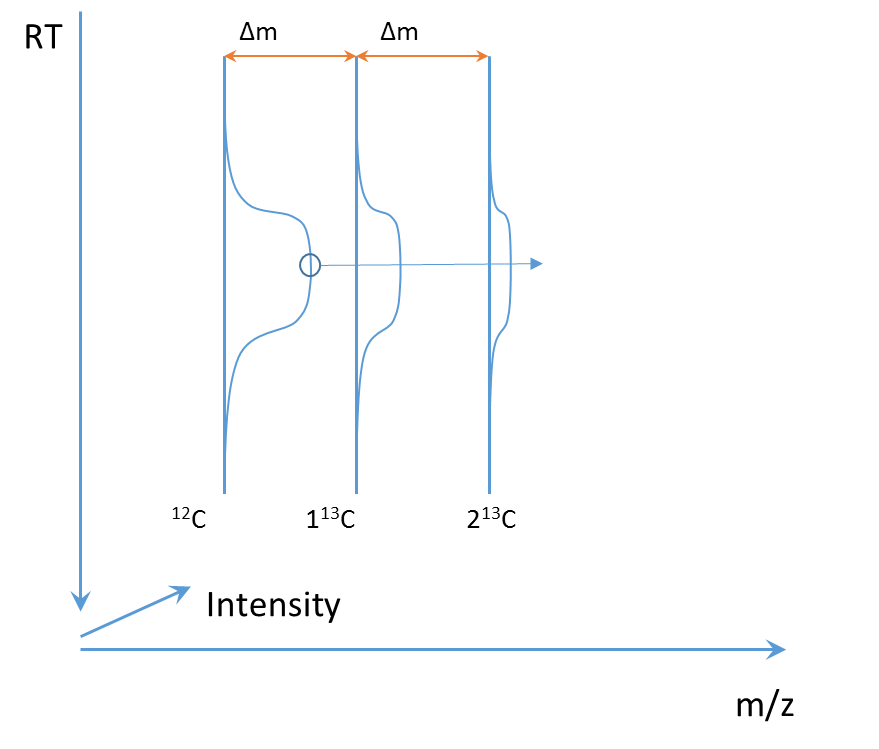
**RTError** – Difference in retention time allowed for chromatographic peaks of the same component in different LC-MS runs. Measured on the apexes of peaks. Reasonable values are 0.5 for long pHilic method and 0.25 for short pHilic method.

**Gap Scan Max** – Maximum gap scan number. I.e. Maximum number of sequential scans where LC-MS trace intensity is zero. Reasonable value is 2.

**Min RT Peak Width** – Minimum width of the signal in RT domain to be considered as a peak. Shorter signals are to be ignored. Reasonable values are 0.1 for long pHilic method and, probably, 0.05 for short pHilic method.

**Max RT Peak Width** – Minimum width of the signal in RT domain to be considered as a peak. Longer signals have to be deconvoluted by wavelet analysis otherwise considered to be noise and ignored. Reasonable values are 2.0 for long pHilic method and, probably, 1.0 for short pHilic method.

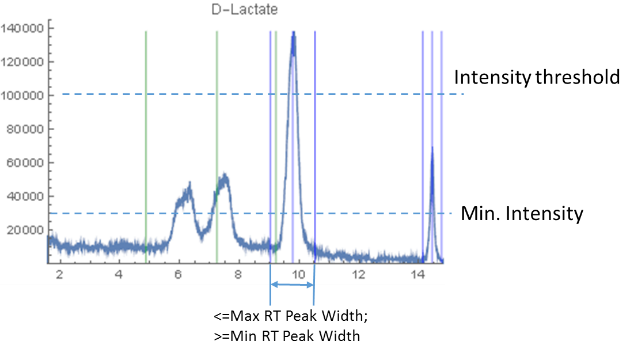
### Isotope Check

****Beside of peak detecting IsoTrack also check if found peak has corresponding stable isotope peaks. Peaks, recognized as stable isotope peaks as on the picture, will not be included in final report.

**C13 to Check** – Number of C13 Isotope peaks to check

### Wavelet peak detection

Wavelet peak detection is intended to extract peaks from traces where are more than one peak observed, like on the picture below.

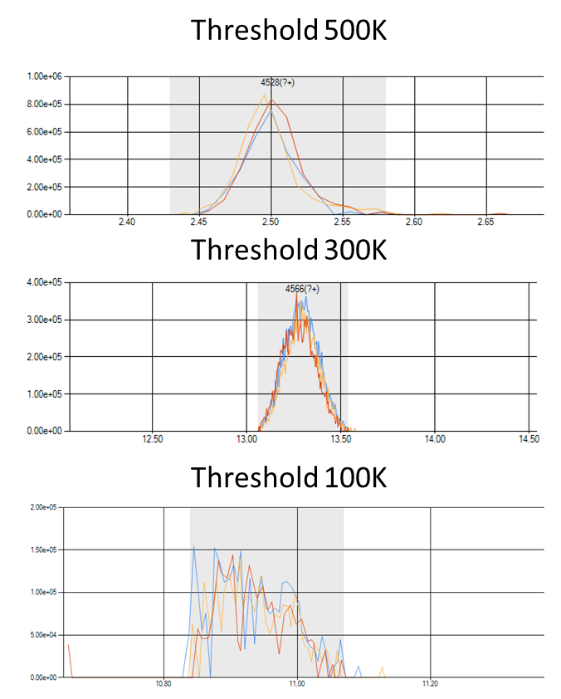
Maximum peak width to be considered is set by **Max RT Peak Width** parameter. Peaks, which are wider then maximum will be considered as background and excluded from report.

Minimum peak width to be considered is set by **Min RT Peak Width** parameter. Peaks, which are narrower then minimum will be considered as random spikes and excluded from report.

**Min. Intensity** – Minimum intensity for wavelet recognized peaks. In my opinion it should not be bigger than “Intensity threshold” parameter (See section “Task Related Misc.”), since it will discriminate wavelet detected peaks over trace peaks. It could be reasonably lower than “Intensity threshold”. Value of 30000 should allow you to gather everything from Thermo files. Value of 100000 will give you most of notable peaks. I use 30000.

**S/N ratio** – signal to noise ratio. 1.5 is quite sensitive value.

### Task Related Misc.

**Common In Files** – This number means in how much of the files from file list have to be present peak to be treated as not random and to be included in final report. For example, if you do untargeted analysis on the dataset of 4 files and you have set this parameter to 4, you require every peak to be present in each of the files. If you set it to 3 you allow peak to be absent in one of the files, and so on. I use this parameter at least at the number of replicates in data set.

**Intensity threshold** – That is a main sensitivity parameter for untargeted analysis. Peak have to have at least one point above this threshold to be included in final report. Here on the picture are some examples of the quality of peaks, which becomes to be available with different threshold values. The value of parameter, you are choosing, depends on what kind of peaks you are hunting for.

### Outputs

 **DB3 File** – This file contains detailed information about found peaks, traces and isotopes. That is a SQLite single file relational database. This file is mandatory, since it is used for IsoTrack internal inter-process communications.

**Target List – out** – Final report in tab-separated text format. It contains list of the files used in Untargeted processing session, and, after table of found compounds. Columns in the table are following:

NAME – Unique identifier of compound (usually “Target #N”)

ADDUCT – Usually empty. Idea is that NAME and ADDUCT can specify particular signal in list of compounds

DESC – compound description, free text. Filled with strings like “RT -X, MZ -Y”

MZ – mass-to-charge value of first pure 12C isotope signal of compound

RTMIN – Minimum retention time of LC peak of compound

RTMAX – Maximum retention time of LC peak of compound

MODE – “+” for positive mode, “-” for negative mode

C13TOCHECK – Number of 13C isotopes checked