Finnee, a Matlab toolbox for the untargeted analysis of datasets obtained using LC-MS with spectra recorded in full scan (MS1) mode

Guillaume L. Erny (guillaume.erny@fe.up.pt)

LEPABE - Laboratory for Process Engineering, Environment, Biotechnology and Energy, Faculty of Engineering, University of Porto, Porto, Portugal

1. Introduction

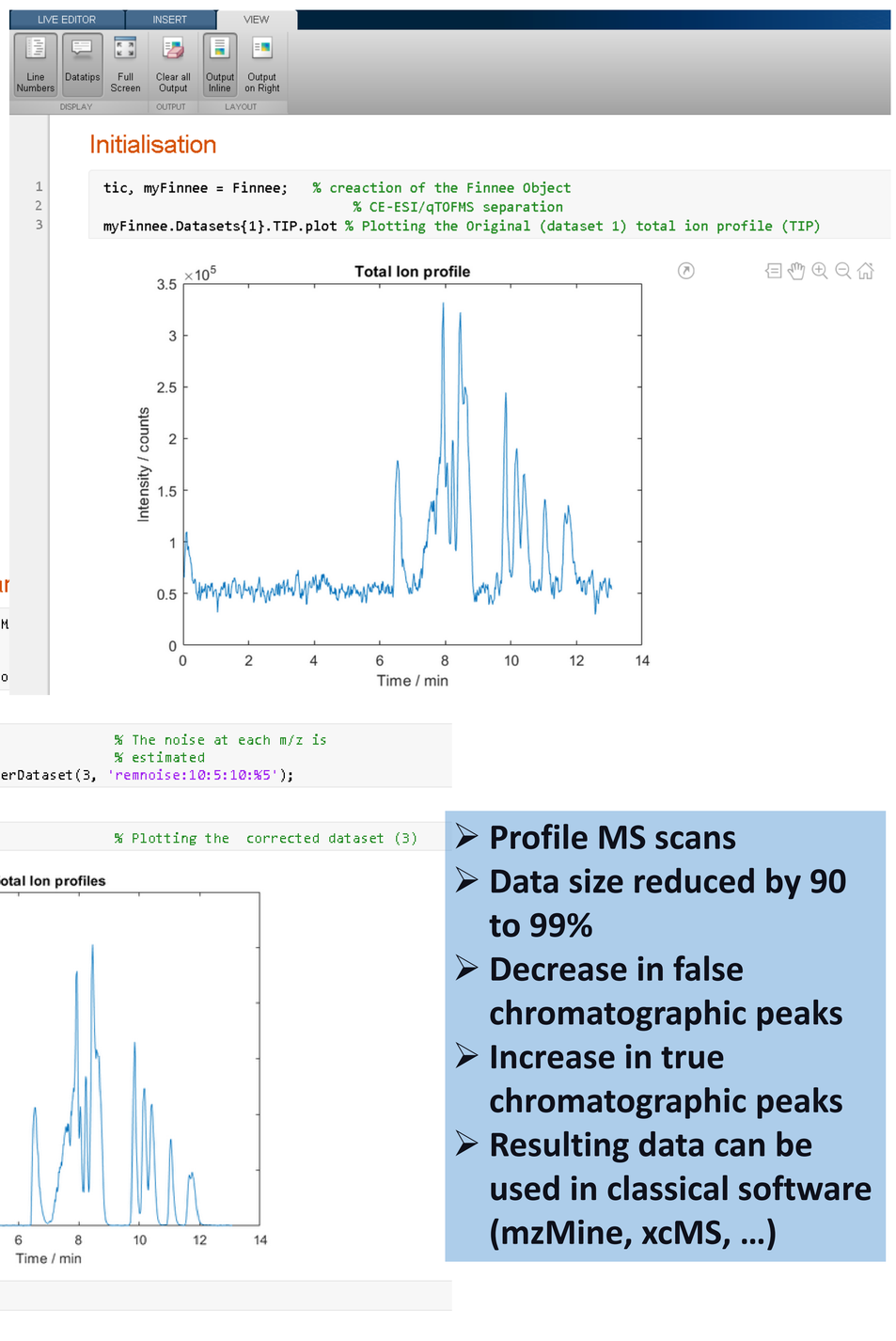
The combination of a separation technique with high-resolution mass spectrometry is the gold standard in chemical analysis. While instruments can be tune and/or resulting datasets filtered to only analyse specific ions resulting from known components (**targeted analysis**), such instruments can analyse thousand of chemicals spanning. Being able to process all the information resulting from the separation and without any prior knowledge or assumption (**untargeted analysis**) is extraordinarily appealing and central to differential metabolomics analysis.

But classical chromatography peak picking approaches are inefficient, and specialised computerised tools must be designed. In full scan mode (MS1), such tools aim to (1) detect chromatographic-like peaks in the dataset (2) measure figure or merits as accurately as possible and (3) recognise peaks belonging to same ions between multiple datasets. Many freeware and software have been proposed, each balancing **computing speed**, **accuracy** and **completeness**.

2. Why Finnee?

Finnee stand for **Fin**d the **Nee**dle, it is a **MATLAB toolbox** [1] that aim to mine for peaks near the limit of quantification and uses chromatographic signal processing rules to insure good measurement of the peak's **chromatographic moments**. The toolbox includes many algorithms allowing to modify single datasets. For each transformation all spectra is recorded in a binary files allowing checking for data integrity. The toolbox is aimed to be versatile allowing easily to design, test and compare different algorithms

3. In-depth corrections of each Dataset[2]



- **Profile MS scans**
- **Data size reduced by 90 to 99%**
- **Decrease in false chromatographic peaks**
- **Increase in true chromatographic peaks**
- **Resulting data can be used in classical software (mzMine, xcMS, …)**

Figure 1. Matlab code and figure allowing to correct each dataset

4. PeakList and Visualisation[3]

Getting the peaklists

```
tic,% myFinnee = myFinnee.doCentroid(4, 'localmax:3:0'); % Profile to centroid conversion
myFinnee.Datasets{5}.Noise = myFinnee.Datasets{3}.AddInfo.Peak2PeakNoise;
ThrInt = 0; % Intensity Threshold
Deltams = 0.001; % Maximum difference in mz for points in successive scans to be marged in a common profile
PtsPerProf = 10; % Minimum number of points for a profile to be valid
myPeakList = PeakList(myFinnee.Datasets{5}, ThrInt, Deltams, PtsPerProf);
toc

myPeakList.Centergram % centergram representation
```

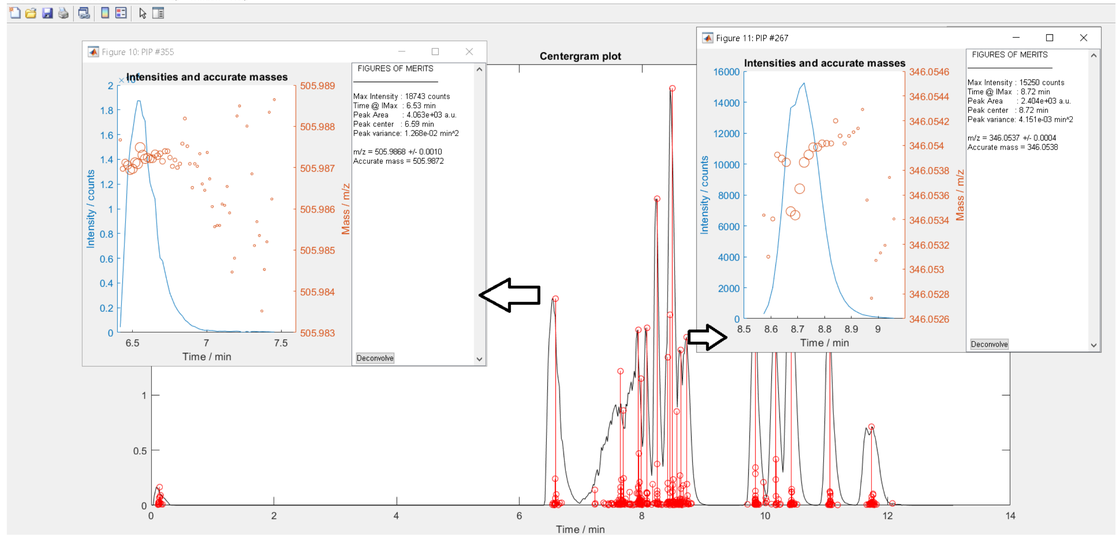
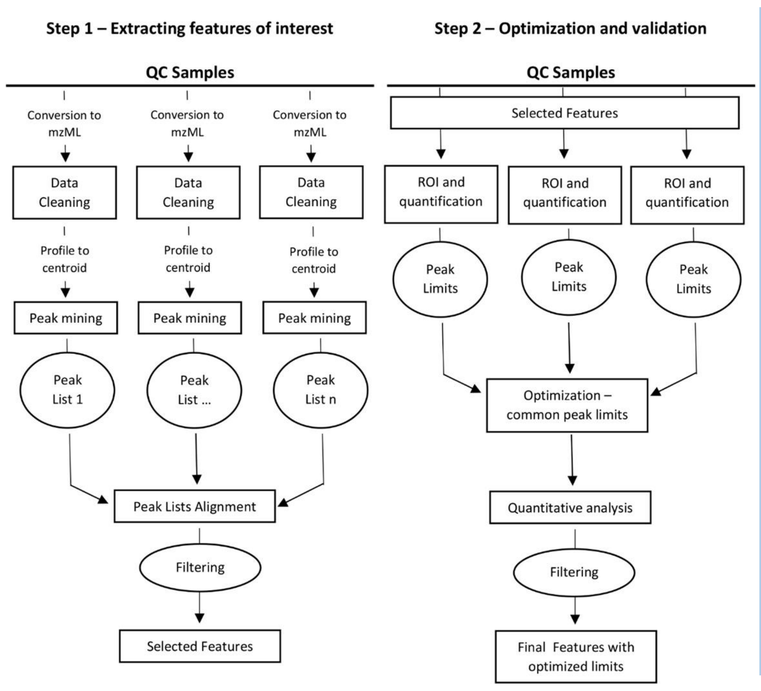


Figure 2. Centergram representation

5. Alignment of multiple datasets, Untargeted and TargeteAnalysis [4]



TWO STEPS APPROACH FOR UNTARGETED ANALYSIS

1. **Untargeted analysis with QC samples only**
   - **List of potential markers recognised by mz and time**
2. **Targeted analysis with QC samples and unknown**
   - **Improve analysis**
   - **Gain of time**
   - **Parallel analysis**

6. Conclusions and References

Finnee is **open source, free to use** and available in Github, Zenodo and Matlab central (https://github.com/glerny/Finnee2016). More information can be found in the Github wiki and in the Finnee Blog (https://finneeblog.wordpress.com/). A new version is being developed to use parallel computing on distributed servers.

[1] Erny *et al. Finnee - A Matlab toolbox for separation techniques hyphenated high resolution mass spectrometry dataset,* Chem. Int. Lab., **2016**, *155,* 138-144.
[2] Erny *et al. Background correction in separation techniques hyphenated to high-resolution mass spectrometry–Thorough correction with mass spectrometry scans recorded as profile spectra,* J. Chrom. A, **2017**, 1492, 98-105.
[3] Erny *et al. Algorithm for comprehensive analysis of datasets from hyphenated high-resolution mass spectrometric techniques using single ion profiles and cluster analysis*, J. Chrom. A, **2016**, 1429, 134-141
[4] Erny *et al. Mining for Peaks in LC-HRMS Datasets Using Finnee – A Case Study with Exhaled Breath Condensates from Healthy, Asthmatic, and COPD Patients*, ACS Omega, **2020**, 26, 16089–16098

Acknowledgements

This work was financially supported by the projects:
(i) UID/ EQU/00511/2019 - Laboratory for Process Engineering, Environment, Biotechnology and Energy – LEPABE funded by national funds through FCT/MCTES (PIDDAC);
(ii) POCI-01-0145-FEDER-029702 and POCI-01-0145-FEDER031297 funded by FEDER funds through COMPETE2020 – Programa Operacional Competitividade e Internacionalizaçã õ (POCI) and by national funds (PIDDAC) through FCT/ MCTES