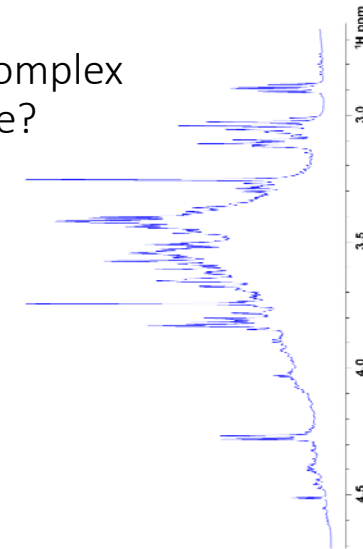


# Complex Mixtures and Multivariate Analysis

SNUG PG NMR  
Alan J R Smith

What is a Complex Mixture?



Caramel

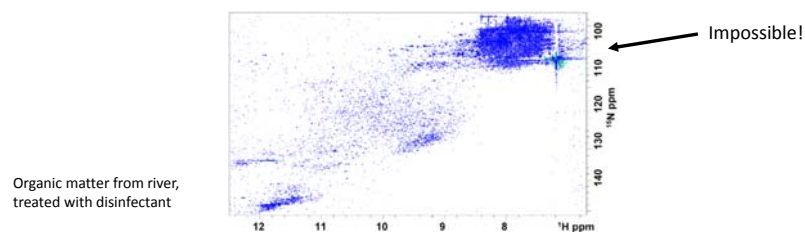
## Benefits to Mixture Analysis

- Often have no choice
- Chromatography is time consuming, expensive and usually a targeted process
- Compounds are in their native state, direct measurement of the mixture
- Minimal sample preparation



## Disadvantages to Mixture Analysis

- Often limited in terms of concentration of each individual component making 2D spectra difficult to acquire
- Complex samples result in complex spectra, longer to interpret (if even possible)



## How to Get Your Data into a Useable Format for Performing Multivariate Analysis

Generally useful for everyone, not just complex mixtures

## Topspin Heavy!

- Easier for me!
- Most if not all spectrometers in Scotland are Bruker, so you should be aware of topspin
- Topspin commands are shown in blue
- Mestrenova commands are shown in red

## Sample Prep & Experiment Parameters

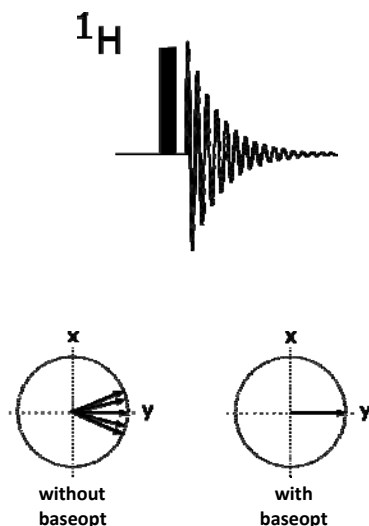
- Samples must be prepared identically
- Consider using an internal standard (e.g. DSS, TSS, could be in a capillary)
- Must have a fixed receiver gain if you want to compare absolute intensities



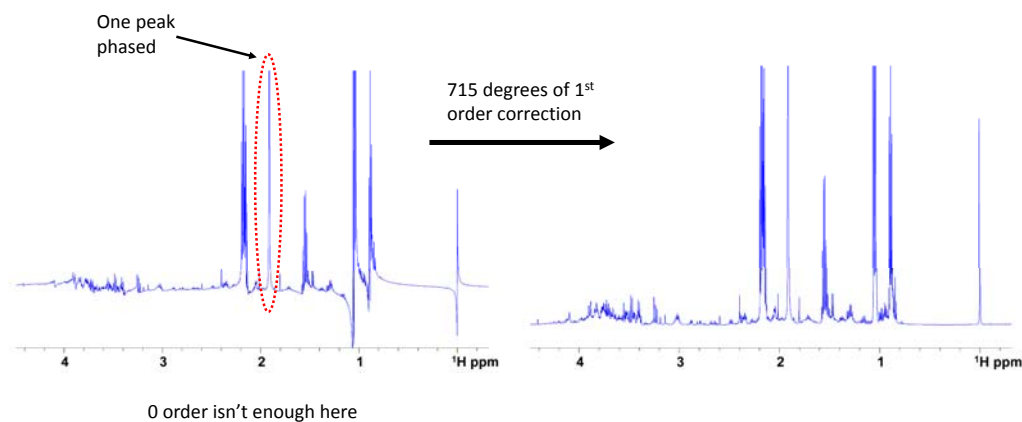
Glass capillary within NMR tube

## Sample Prep & Experiment Parameters

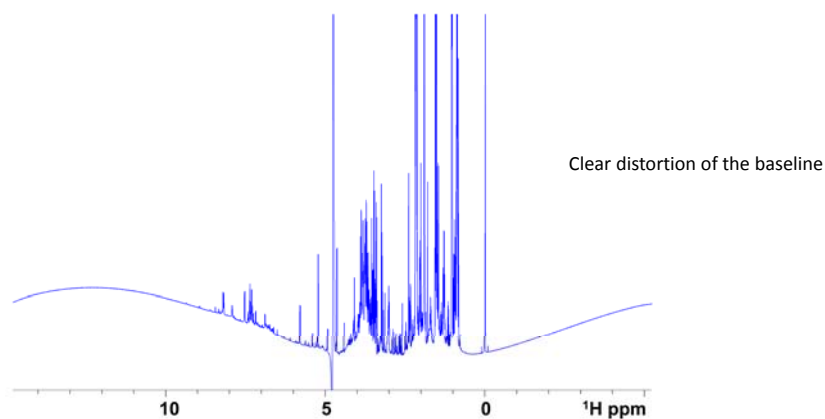
- Ensure relaxation delay (D1) is long enough for complete (or very close) relaxation (very long in inert atmosphere)
- If you're likely to need water suppression, can recommend the noesygppr1d (good baseline + suppression)
- Set Digmod to baseopt, this allows back prediction for chemical shift evolution during final pulse. The result being no need for 1<sup>st</sup> order phase corrections



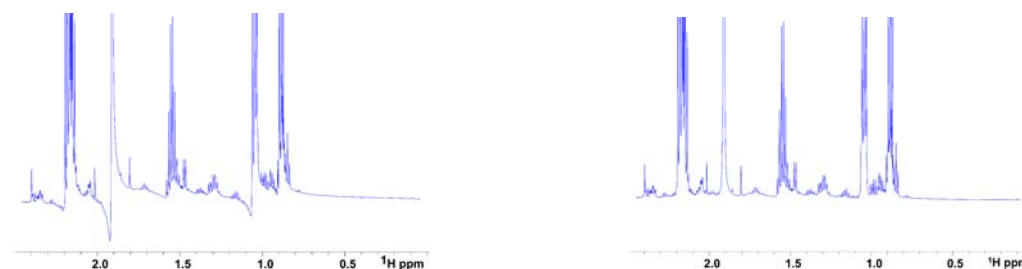
## Extreme 1<sup>st</sup> Order Phase Issues



## Extreme 1<sup>st</sup> Order Phase Issues



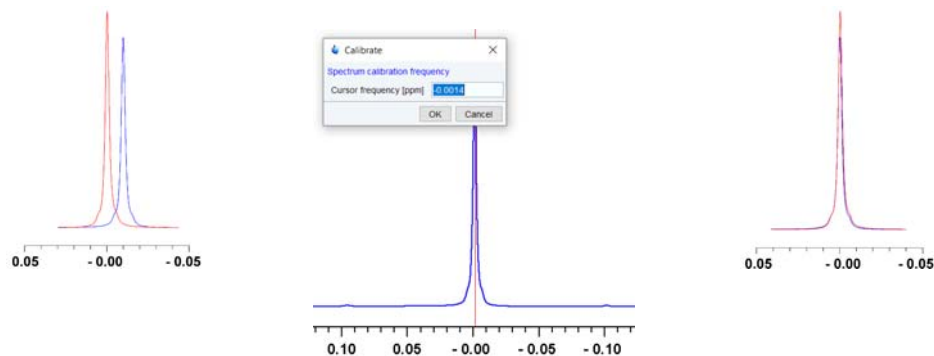
## Spectral Pre-processing – Phase Correction



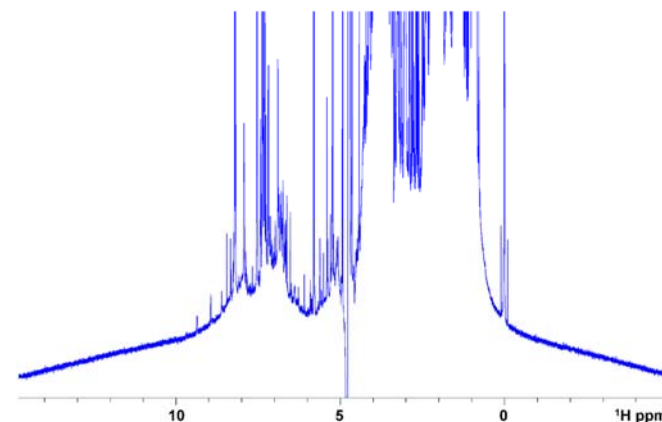
Commands/shortcuts – (.ph, shift + p) manual phasing  
 (apk0) automatic 0 order correction  
 (apk) automatic 0+1<sup>st</sup> order correction (try not to use this option)  
 (apkb) automatic phasing and baseline correction (topspin 4+)

## Spectral Pre-processing (alignment)

Commands/Shortcuts – (.md) opens multiple display  
(.cal, R) opens calibration window

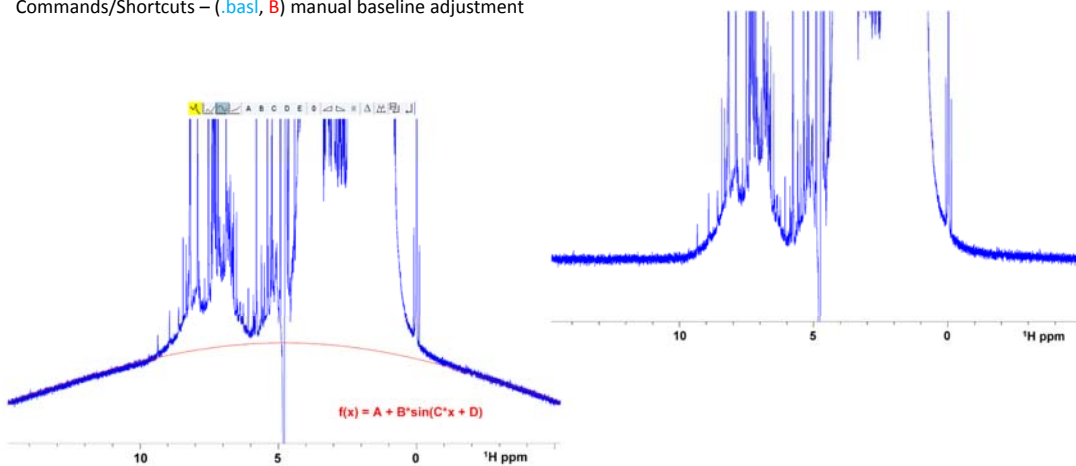


## Spectral Pre-processing (Baseline Correction)



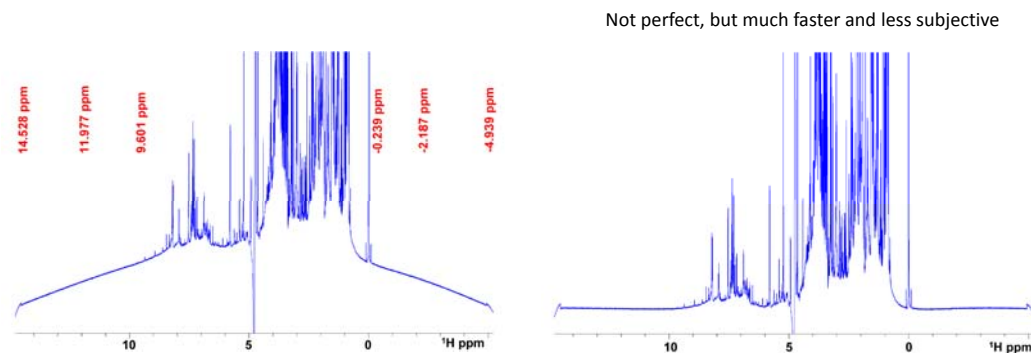
## Spectral Pre-processing (Baseline Correction)

Commands/Shortcuts – (.basl, B) manual baseline adjustment

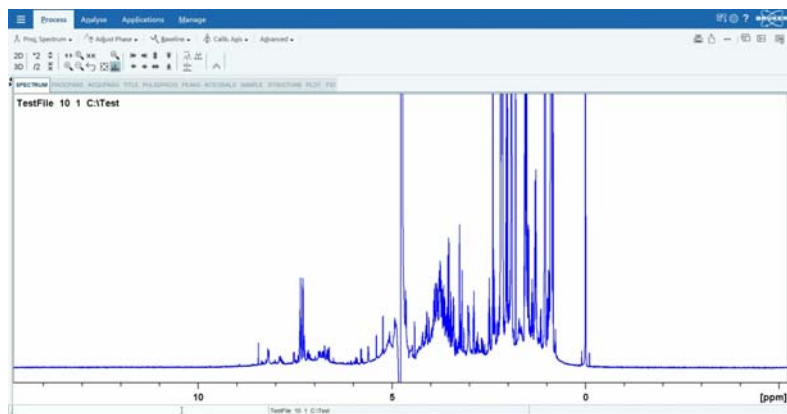


## Spectral Pre-processing (Baseline Correction)

Commands/Shortcuts – (.baslpts, Processing/autobaseline/multipt) manually select points which should be flat baseline.



# Spectral Binning (Bucketing)



Topspin  
Enter integrals menu (.int)

Select import integration regions

Select premade bin list (need to make in excel)

Save

Go to integrals tab

Highlight the integrals

Right click to export as a .CSV

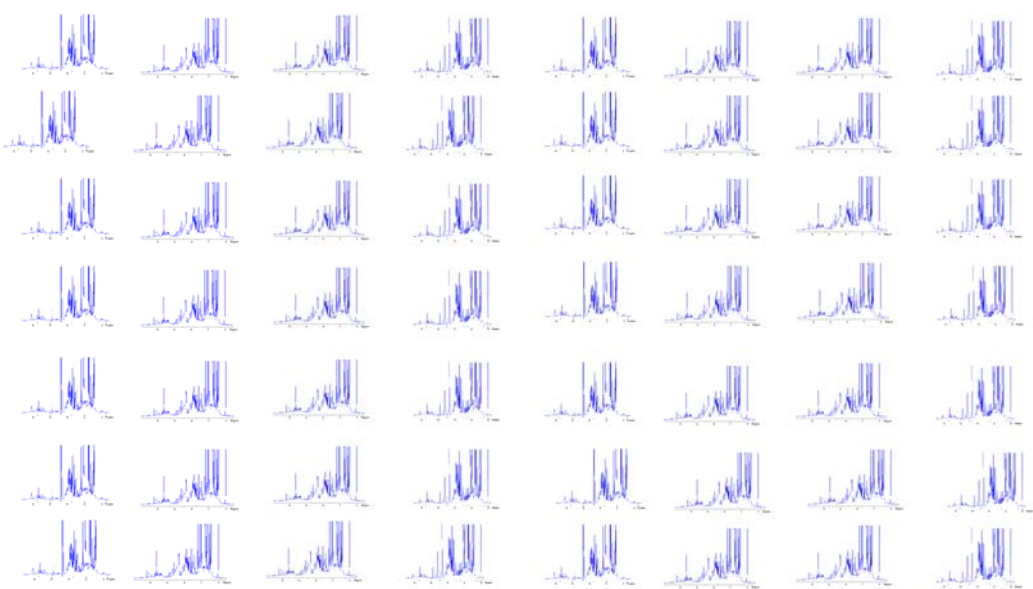
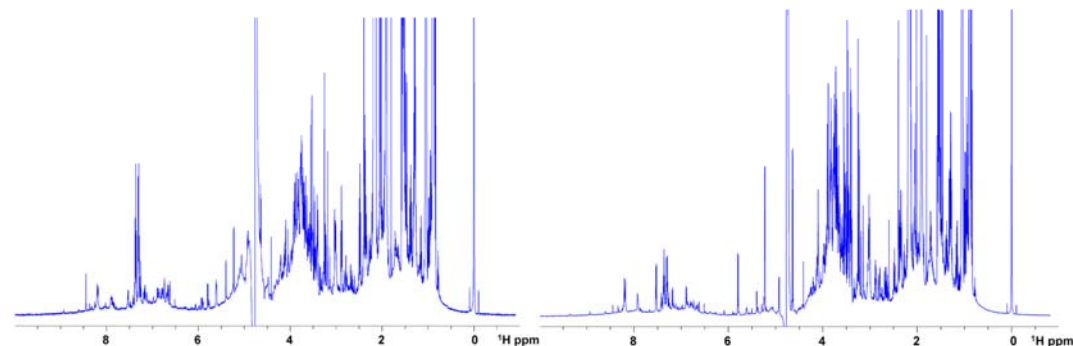
Mestrenova

Processing menu/more

Binning

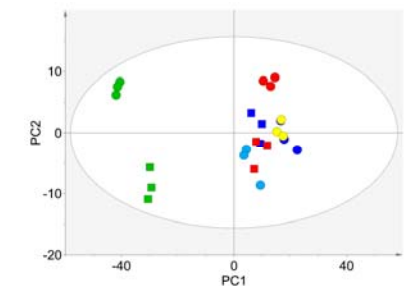
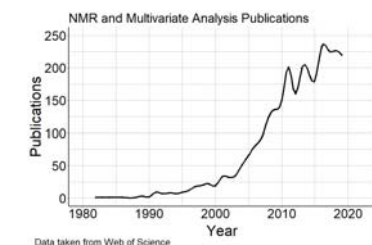
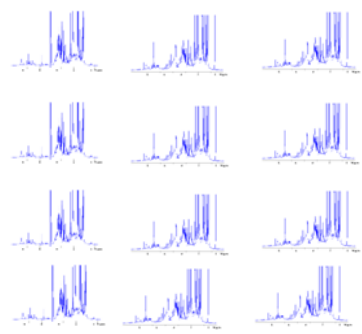
# Multivariate Analysis

- Mostly performed on 1D datasets due to time constraints of 2D acquisition



# Multivariate Analysis

- Highly graphical in its approach, allowing fast interpretation of complex data.



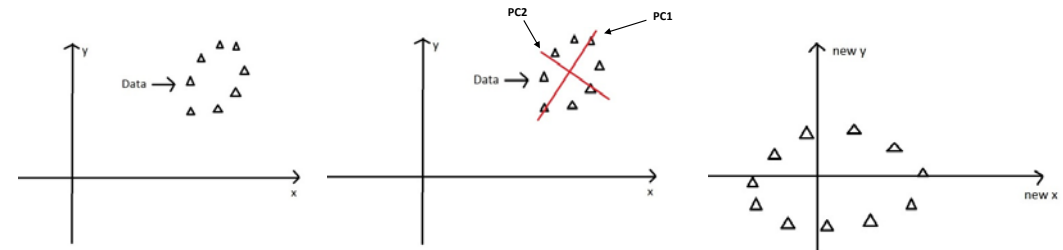
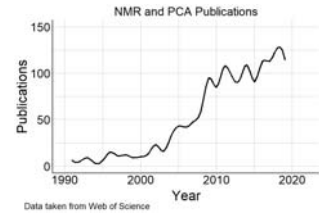
# Multivariate Analysis

- Example of two techniques (there are 100's)
- One which compares sample to sample (PCA)
- Another which delves into the spectra (STOCSY)

\* read more at <https://bit.ly/2QW09op>

# PCA (Principal Component Analysis)

- Transforms data into a new coordinate system to highlight the variance in the system
- With NMR data, each integral would be an individual dimension



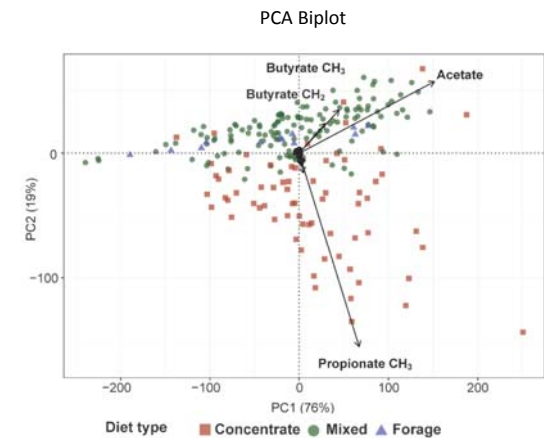
# Example Dataset



- Goal is to link specific metabolites to feedstock and try to gain some understanding of methane production
- Cows fed 1 of 3 diets (concentrate, mixed and forage)
- 128 <sup>1</sup>H spectra

# PCA Example

- Rumen methanogens principally use H<sub>2</sub> and CO<sub>2</sub> to produce CH<sub>4</sub>
- Acetate and Butyrate production results in H<sub>2</sub>
- Concentrate diet is associated with an increase in propionate
- This is linked to a decrease in methane production



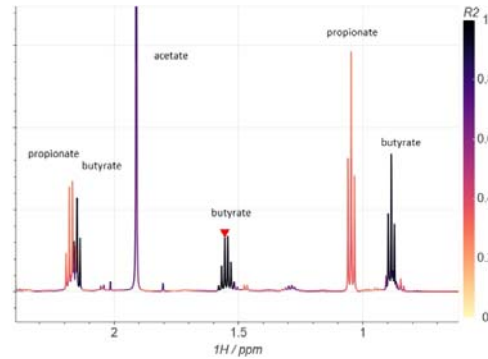
# STOCSY (Statistical TOCSY)

- One “driver” peak is selected
- R<sup>2</sup> values are calculated across spectral range
- R<sup>2</sup> values are then colour coded for visualisation purposes, highlighting correlations between metabolites

STOCSY can be performed in the following-

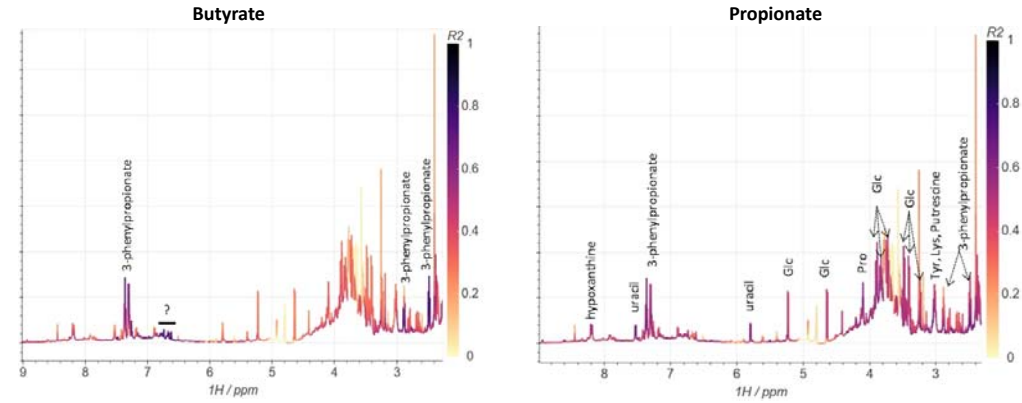
MWASTools (R package)  
iSTATS (R package)  
muma (R package)

Visualisation was performed using Bokeh (a python package), could use Plotly (R + python)

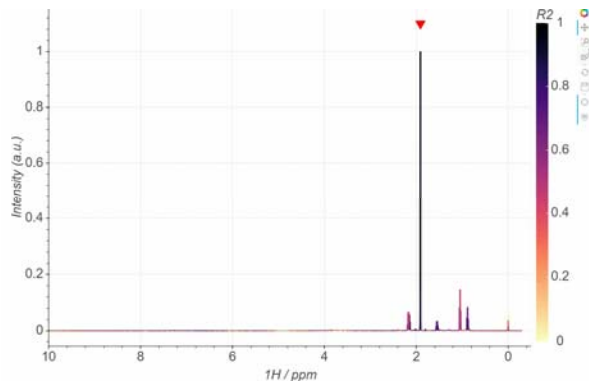


# STOCSY

Propionate has much stronger correlations to metabolites such as glucose and certain amino acids.



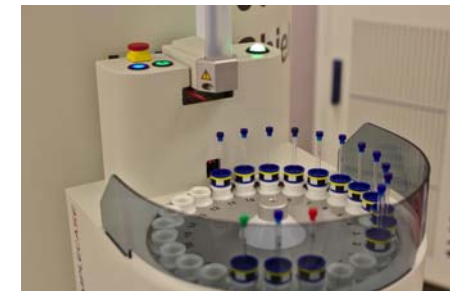
# STOCSY



Plotly and Bokeh produce interactive plots, allowing you to interrogate as if it were an NMR spectrum.

# Take Home Message

- You **must** prepare samples appropriately
- Setup the **spectrometer** correctly, worthwhile setting up a dedicated experiment in automation
- Pre-processing is **key**, you don't want to artificially increase variation in the dataset



## Examples and Useful Links

Navigating freely-available software tools for metabolomics analysis – DOI [10.1007/s11306-017-1242-7](https://doi.org/10.1007/s11306-017-1242-7)

HRMAS-NMR spectroscopy and multivariate analysis meat characterisation – DOI [10.1016/j.meatsci.2012.06.034](https://doi.org/10.1016/j.meatsci.2012.06.034)

<sup>1</sup>H-NMR with Multivariate Analysis for Automobile Lubricant Comparison – DOI [10.1111/1556-4029.13471](https://doi.org/10.1111/1556-4029.13471)

Assessment of <sup>1</sup>H NMR spectroscopy and multivariate analysis as a technique for metabolite fingerprinting of *Arabidopsis thaliana* – DOI [10.1016/S0031-9422\(02\)00705-7](https://doi.org/10.1016/S0031-9422(02)00705-7)

Using Compact <sup>1</sup>H NMR, NIR, and Raman Spectroscopy Combined with Multivariate Data Analysis to Monitor a Biocatalyzed Reaction in a Microreaction System – DOI [10.1155/2018/5120789](https://doi.org/10.1155/2018/5120789)

Analysis of Scotch Whisky by <sup>1</sup>H NMR and chemometrics yields insight into its complex chemistry – DOI [10.1016/j.foodchem.2019.125052](https://doi.org/10.1016/j.foodchem.2019.125052)

Multivariate Analysis of 2D <sup>1</sup>H, <sup>13</sup>C methyl NMR Spectra of Monoclonal Antibody Therapeutics to Facilitate Assessment of Higher Order Structure – DOI [10.1021/acs.analchem.7b03571](https://doi.org/10.1021/acs.analchem.7b03571)

## Acknowledgements

Thanks to -

- Dusan Uhrin and Will Kew for the STOCSY spectra/data, J Palarea-Albaladejo for the PCA
- R Bica, A Macrae and R.J Dewhurst for the cow rumen study
- You, for listening!