Multivariate analysis of E-1559 temperature programmed desorption

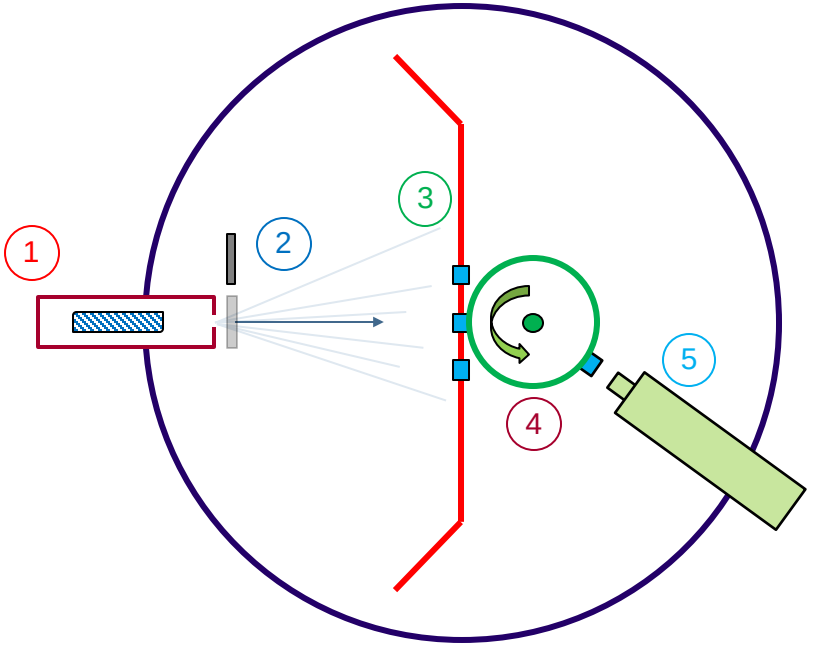
Bob Moision
The Aerospace Corporation

July 17, 2017

The ASTM E-1559 Test

- ASTM-E-1559 test is one of the cornerstone analytical test methods for contamination analysis
- The test is composed of two phases
  - ***Deposition:*** *A sample at a known temperature outgasses material that is collected onto quartz crystal microbalances (QCM), also with well defined temperatures*
  - ***Desorption:*** *Following deposition, the QCM is heated in a systematic manner and the temperature dependence of material desorbing from the surface is monitored*
- In the Aerospace Contamination Effects Research Test (CERT) chamber, there are 3 QCMs for deposition and desorption
  - *The center QCM is on a rotating stage and can be rotated to the entrance of a mass spectrometer*

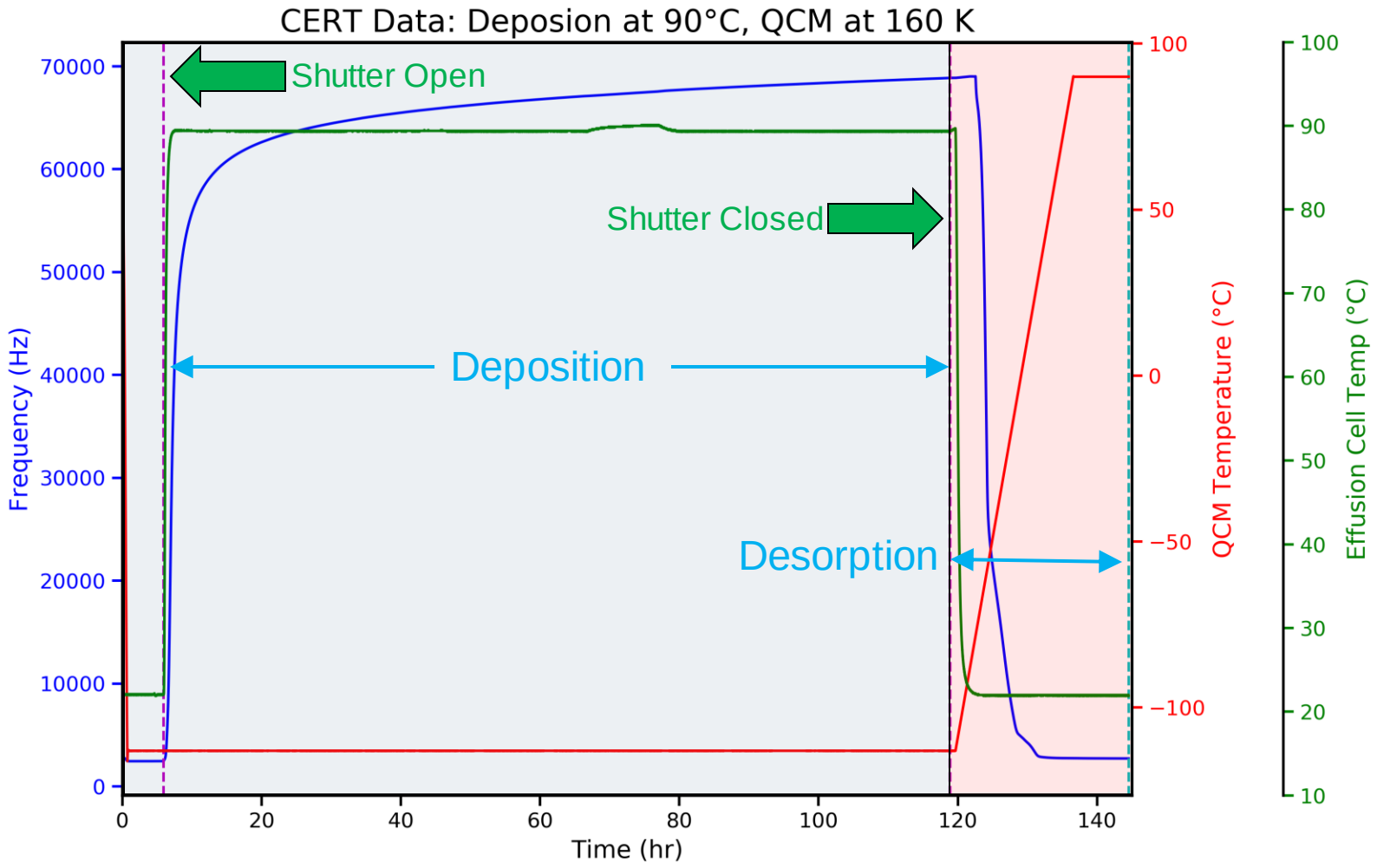
Contamination Effects and Research Test (CERT) Chamber Schematic



1. Effusion cell with sample
2. Shutter and outgassing material
3. QCMs and cold shroud
4. Center QCM on rotatable stage
5. Mass spectrometer

ASTM E-1559 Data

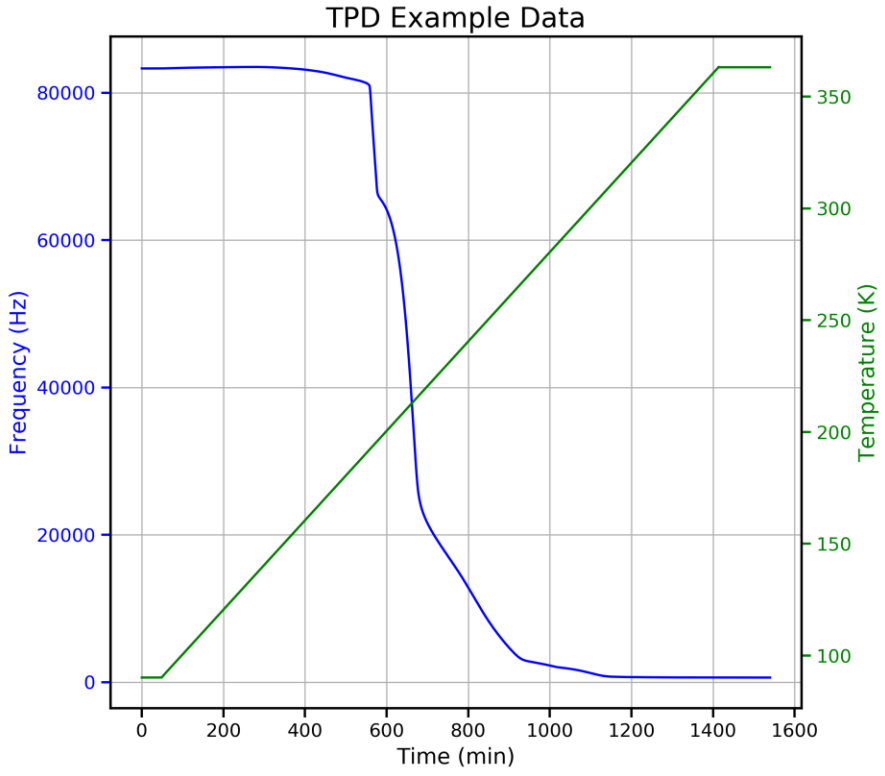
Deposition and Desorption phases

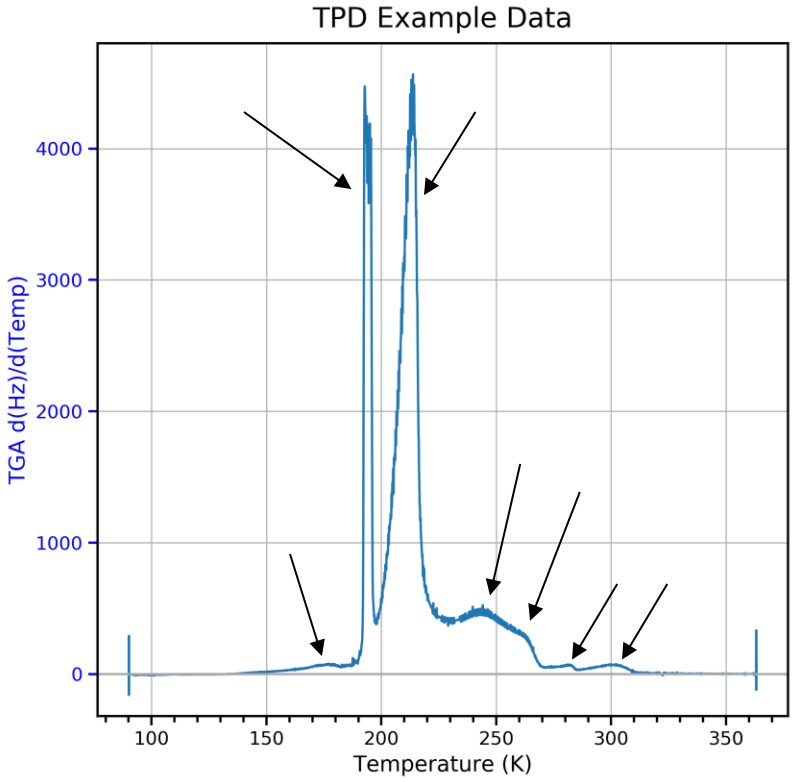


Temperature Programmed Desorption (TPD)

Also referred to as thermo-gravimetric analysis (TGA)

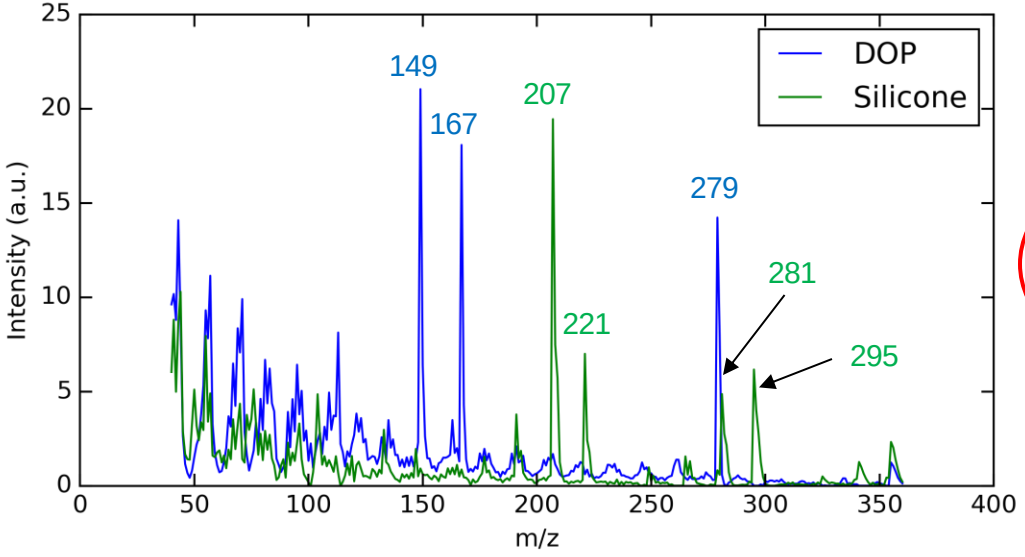
- The change in QCM signal with respect to QCM temperature shows peaks that represent different species leaving the QCM surface
- Because the QCM only measures changes in frequency it is unable to provide chemical identification of the desorbing species

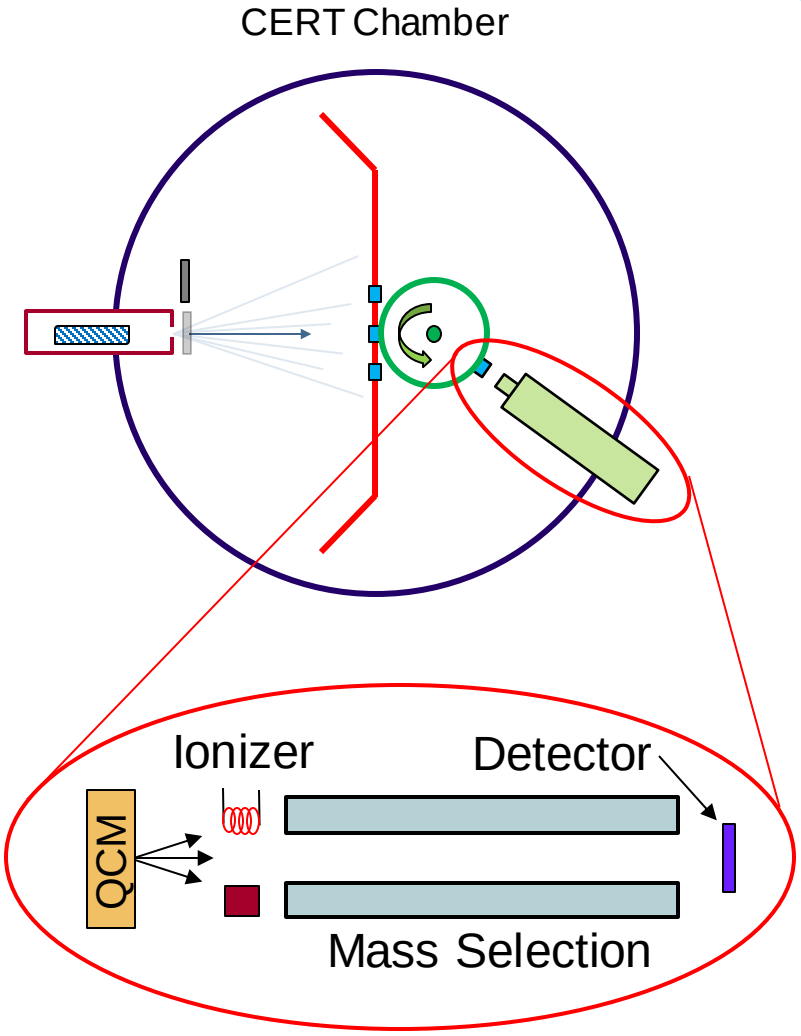




Quadrupole Mass Spectrometer

- A Quadrupole Mass Spectrometer (QMS) is used to analyze species leaving the QCM surface during TPD heating
  - *The QMS uses an electron impact (EI) ionizer that fragments the parent species*
  - *The QMS measures the mass of the ionized species*
  - *This forms a characteristic "mass fingerprint" for each species*

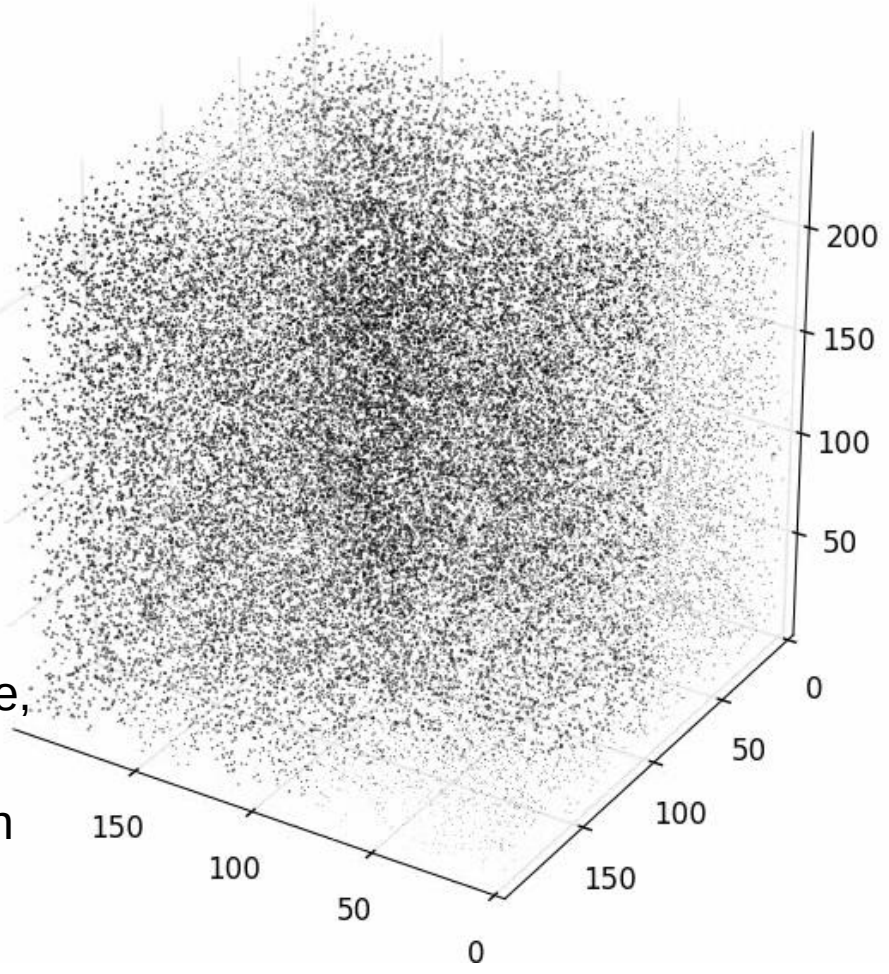




Too Much Data!

More data does not equal more useful information

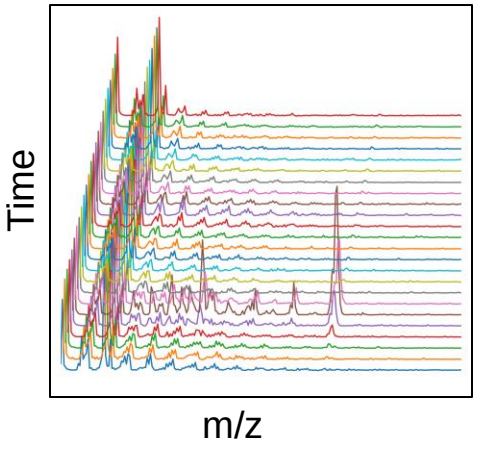
- Modern analytical instrumentation can produce immense amounts of data in a single experiment
- But most systems are characterized by only a handful of underlying processes or relationships
- MVA extracts information from data with multiple variables by using all variables simultaneously
  - *More efficient analysis of large datasets*
  - *Can improve signal to noise*
- MVA is used to more efficiently summarize, simplify and expose underlying trends
- MVA relies on data changing as a function of a second variable
  - *Mass spectra changing as a function of temperature)*



Multivariate Analysis

Non negative matrix factorization (NNMF)

Raw Spectra



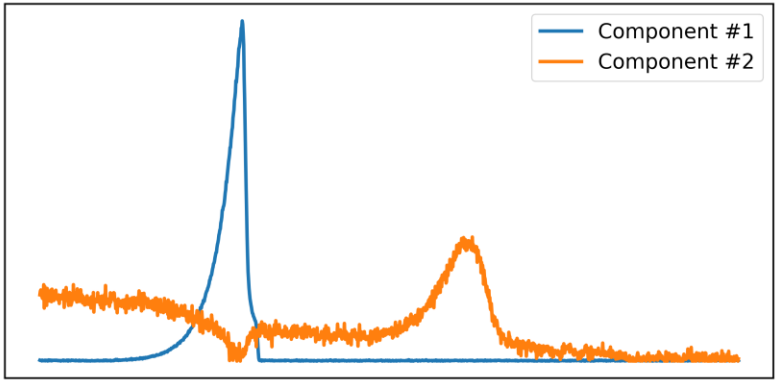
Data Matrix

Time

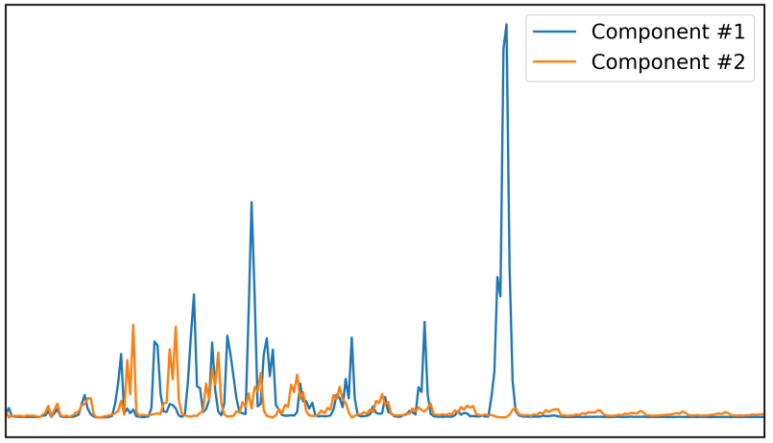
0	2	2	0	0	1
0	2	3	0	1	1
0	2	3	1	2	2
1	2	0	2	8	3
1	1	0	9	2	1
2	0	5	2	1	0

m/z

Component Concentration



Component Spectra

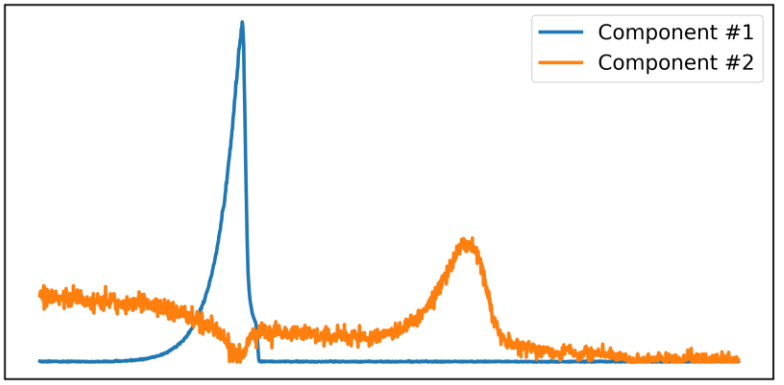


- NNMF decomposes the data matrix into a concentration and spectral components
  - *NNMF forces values in the "spectral" dimension to be positive*
  - *Results look like real spectra*
- The **only** user input is the number of components to be extracted

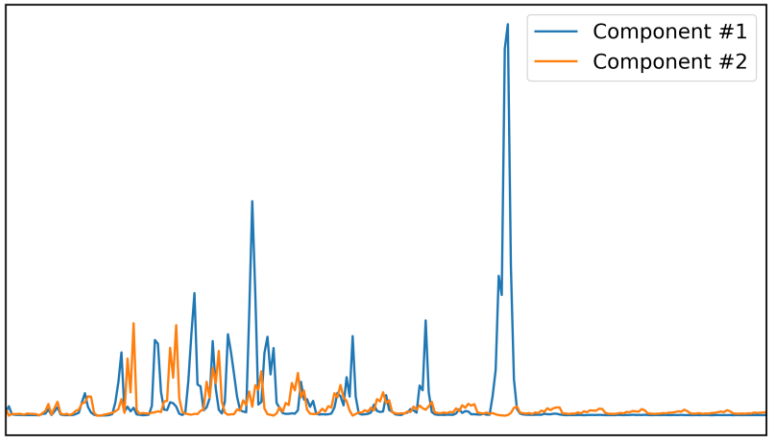
What MVA is NOT

- It is not magic
  - *Garbage in still equals garbage out*
  - *No amount of data massaging can make up for poor experimental design or poor quality data*
- It does not replace traditional approaches
  - *But traditional methods used in conjunction with MVA can provide a more efficient pathway to understanding your data*
- It is not difficult
  - *There are numerous commercial and free packages available*
  - *As with any approach, understanding the limitations of the method(s) will help to avoid pitfalls*

Component Concentration



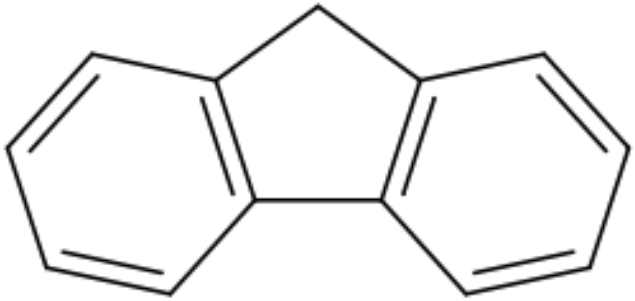
Component Spectra



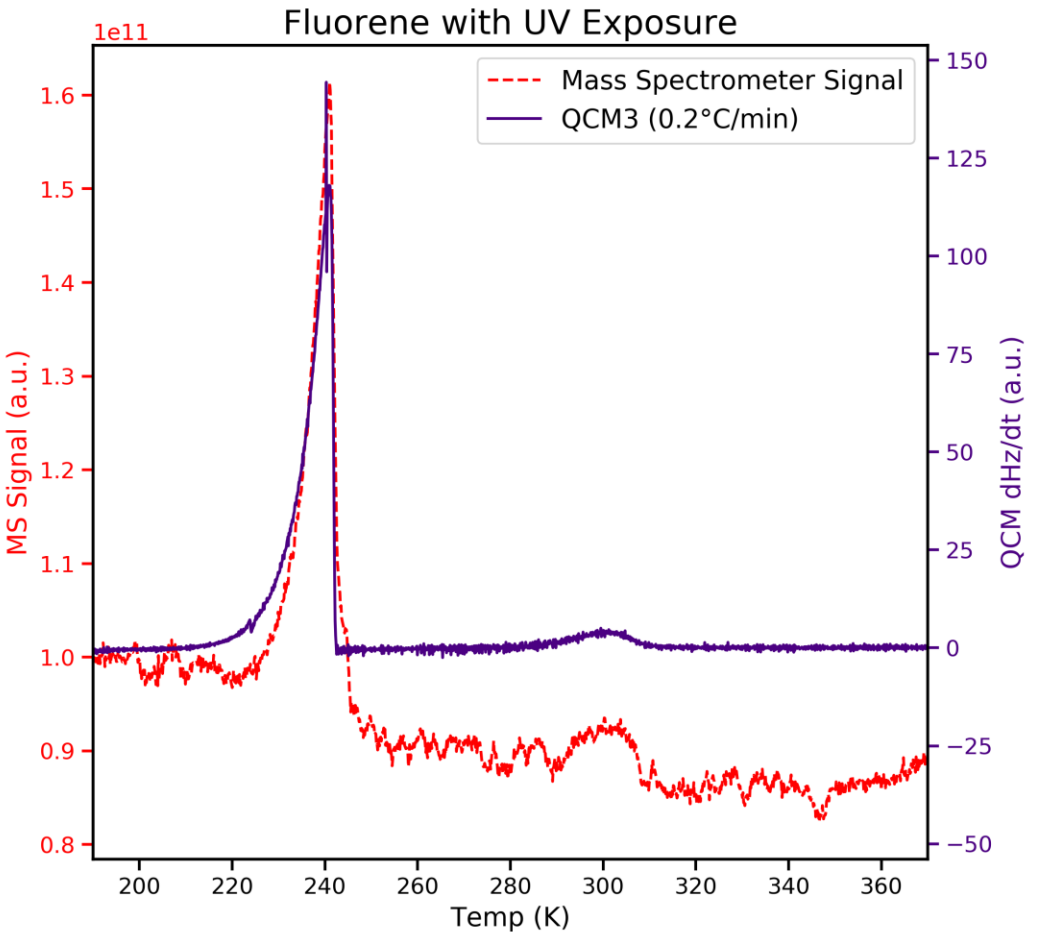
Fluorene Contamination

Can Fluorene polymerize in the presence of UV?

Fluorene $C_{13}H_{10}$



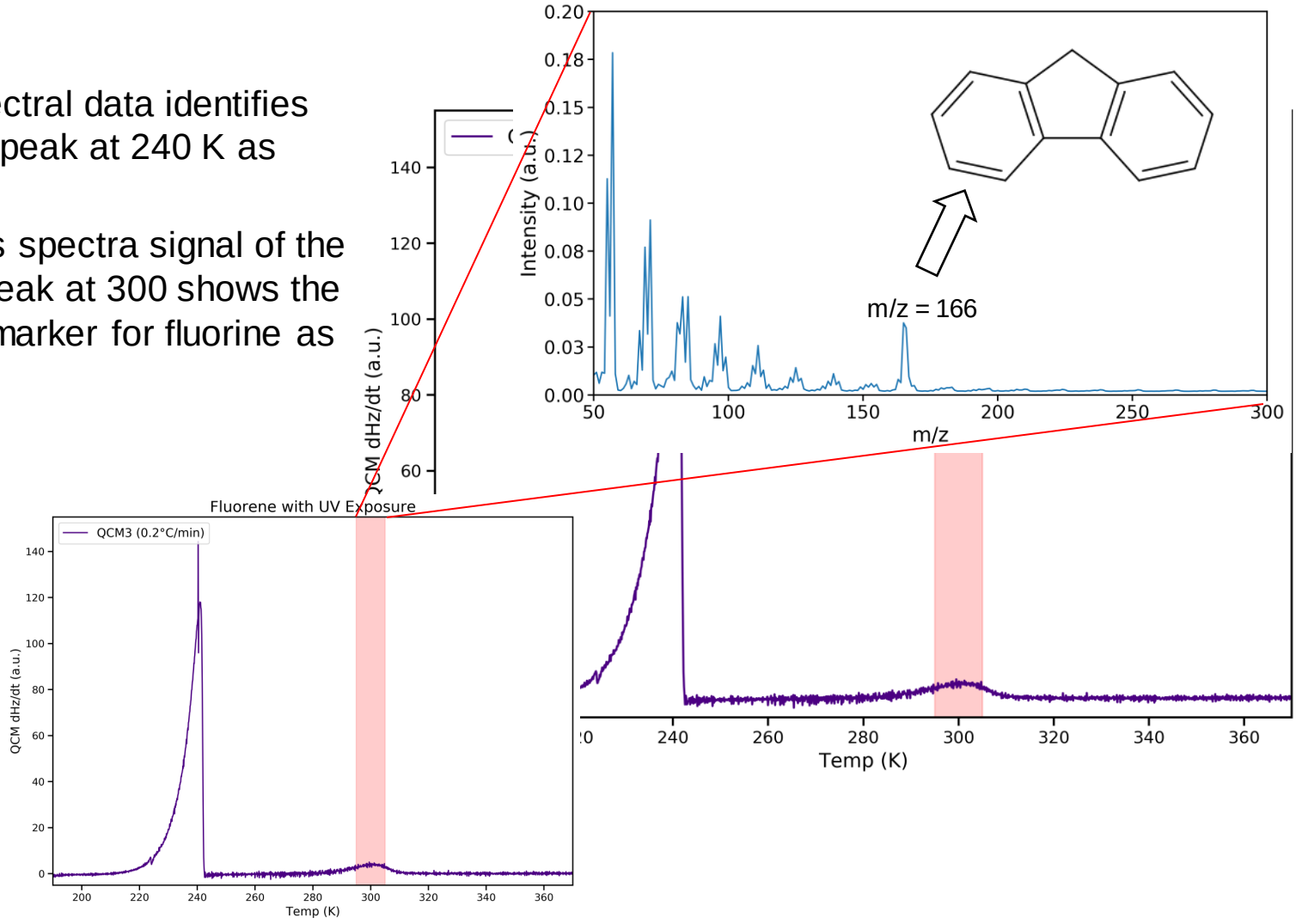
- Fluorene proposed to be used on a picosat
- Questions were raised on whether it could polymerize with UV exposure
- Fluorene was deposited onto QCMs while being irradiated with a UV lamp
- TPD of deposited material showed two peaks. Polymerization?



Fluorene Contamination

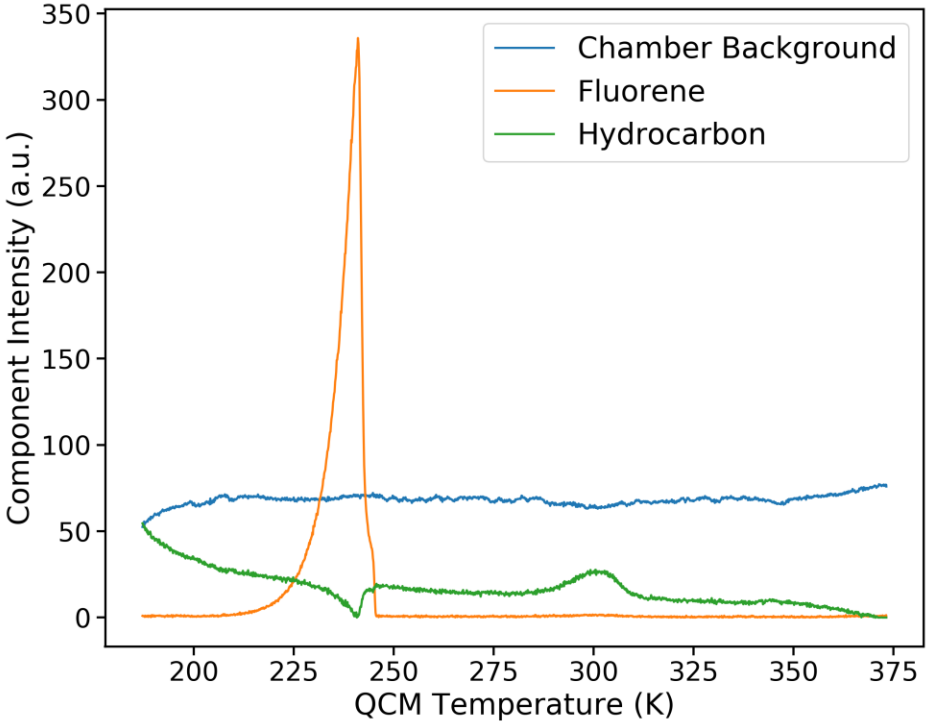
Can Fluorene polymerize in the presence of UV?

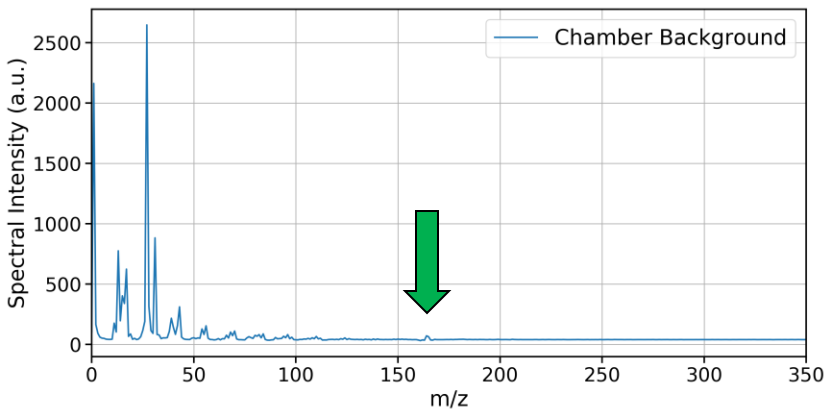
- Mass spectral data identifies the large peak at 240 K as fluorine
- The mass spectra signal of the smaller peak at 300 shows the 166 m/z marker for fluorine as well

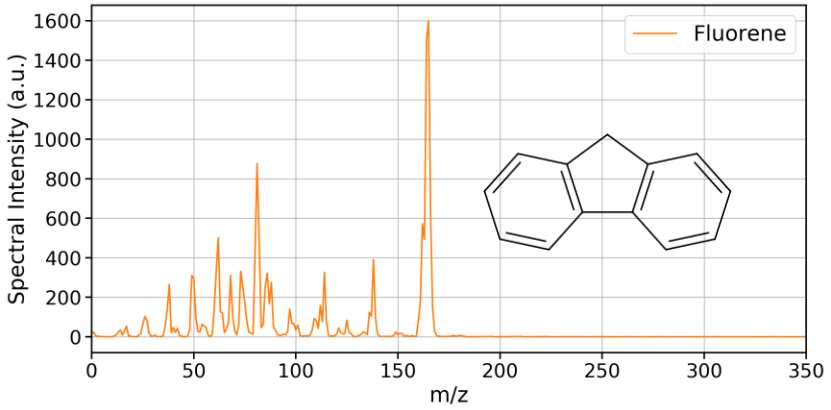


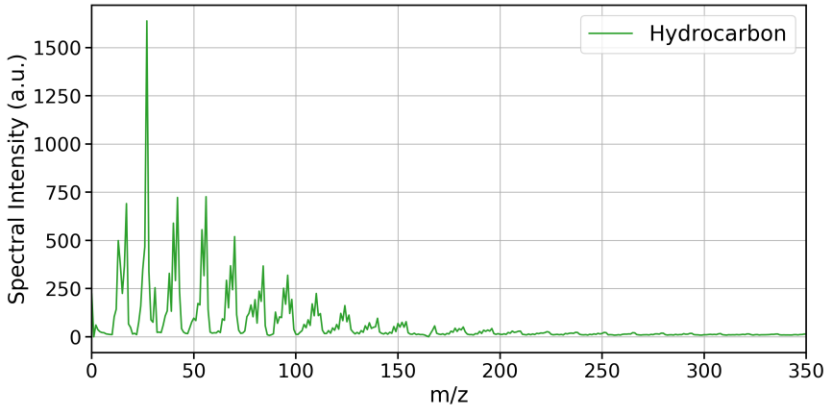
Fluorene Contamination

NNMF analysis identifies peaks





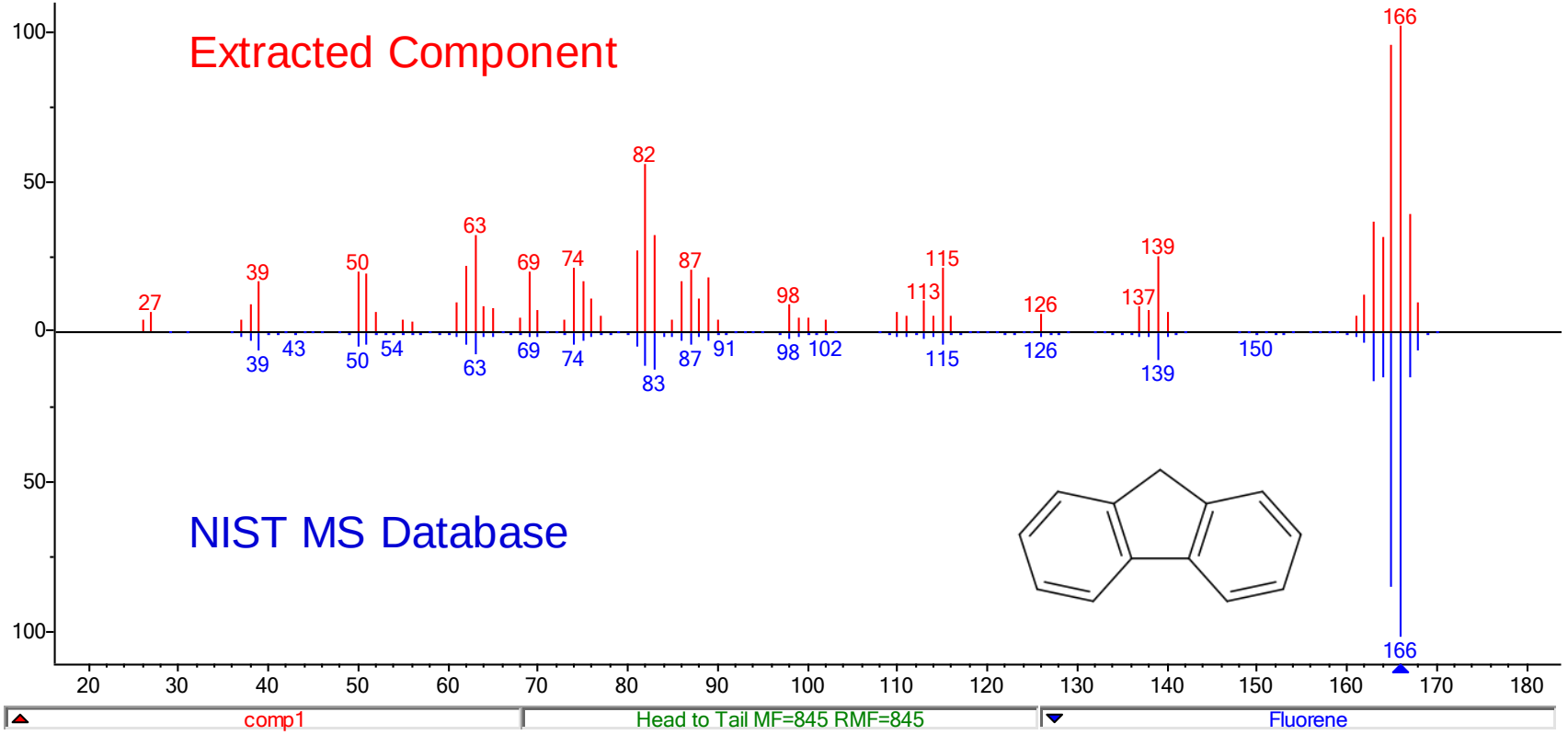




- Fluorene sample used for analysis was labeled as 98% pure
- MVA analysis clearly separates flourene from the hydrocarbon impurity and chamber background
- Heavyhanded use of fluorene in sample cell resulted in fluorene chamber contamination

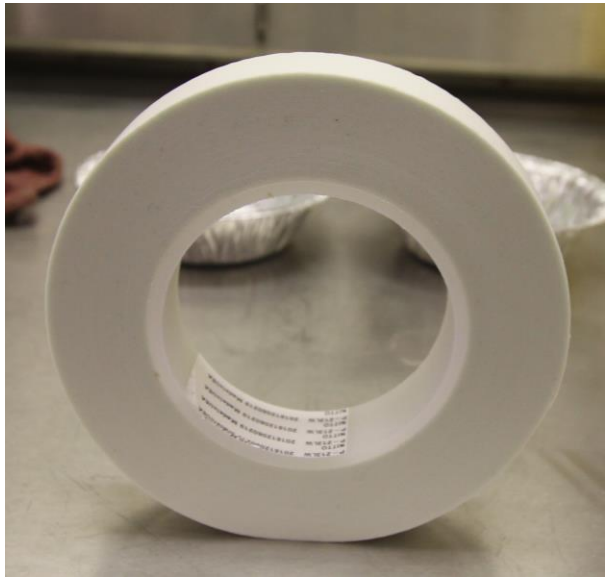
Fluorene Contamination

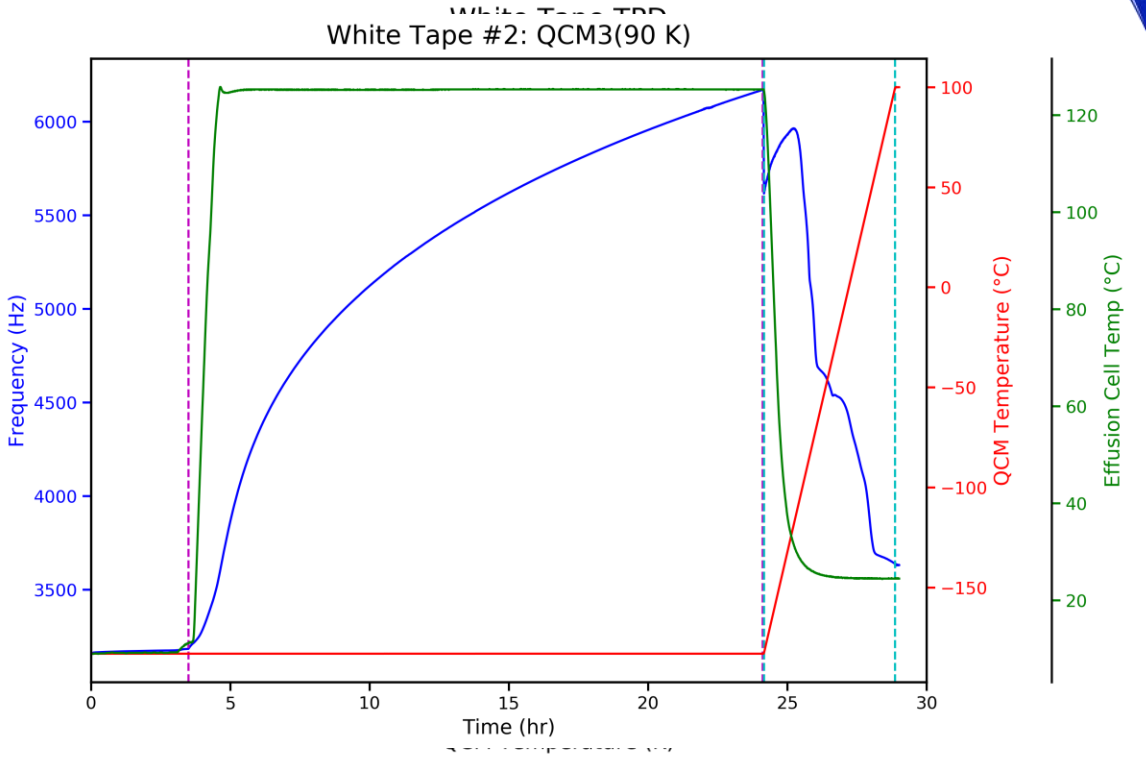
Component spectra compared to NIST MS database



White Fiberglass Tape TPD

Sticky tape has a number of chemical components

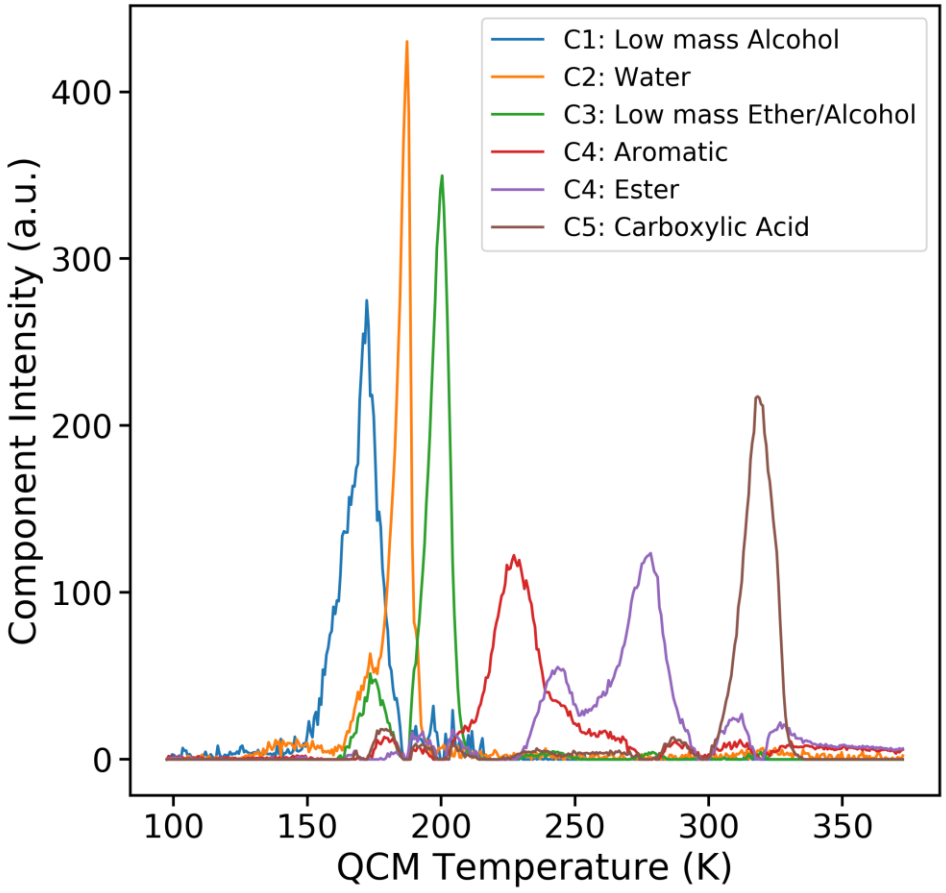


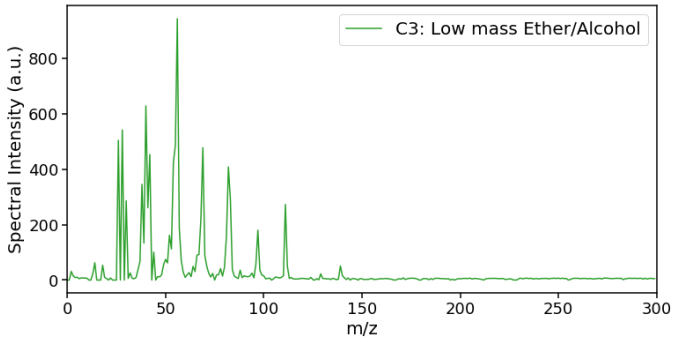


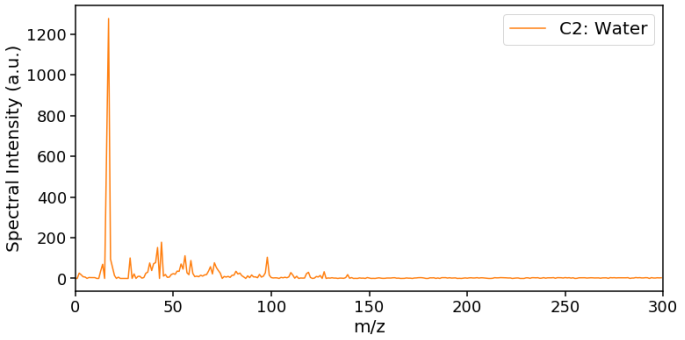
- QCM and mass spectral signal for white tape TPD shows good agreement
- Can MVA techniques be used to deconvolute the overlapping peaks?

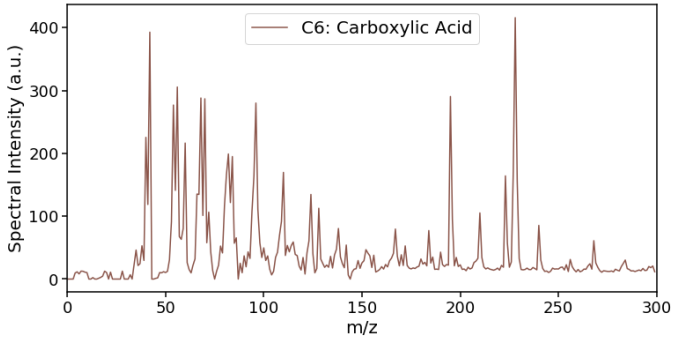
White Fiberglass Tape TPD

NNMF identifies individual chemical components



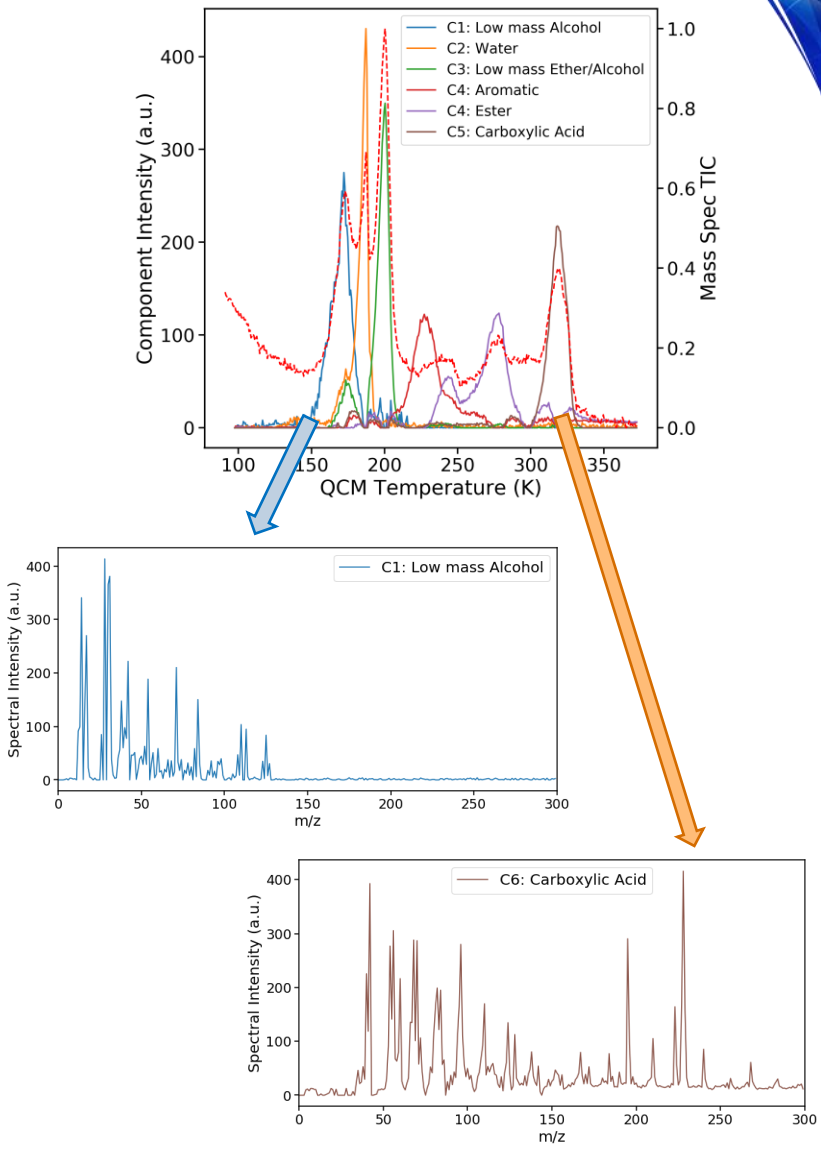






Conclusions

- The addition of a mass spectrometer to analyze materials desorbing from QCM surfaces provides chemical insight that is unavailable with the QCM alone
- The use of multivariate techniques for analysis of TPD data can help to reduce a large amount of mass spectral information into simple and understandable components
- Both commercial and free multivariate software packages are available
  - *Freeware: Python (scikit-learn), R, Octave*
  - *Commercial: MATLAB (PLS Toolbox), CAMO (Unscrambler)*

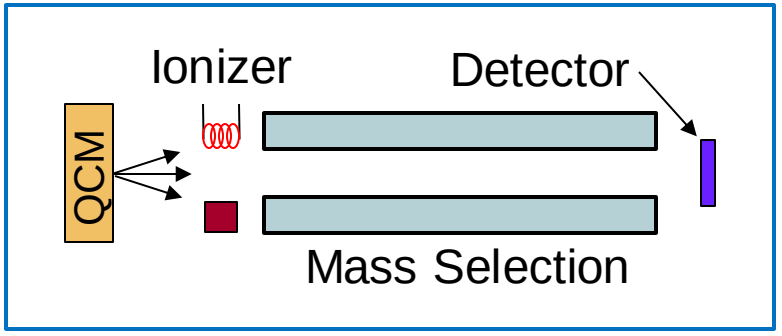


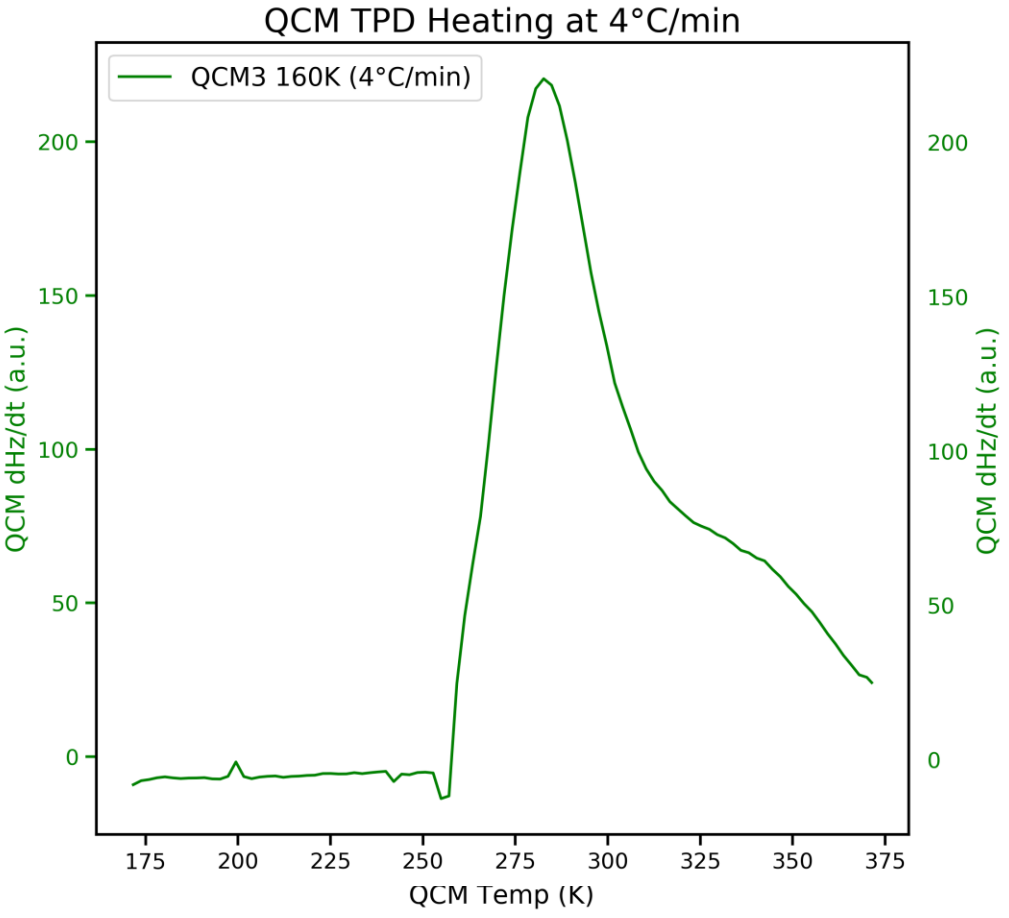
Backup Slides

QCM and Mass Spec Signal Discrepancy

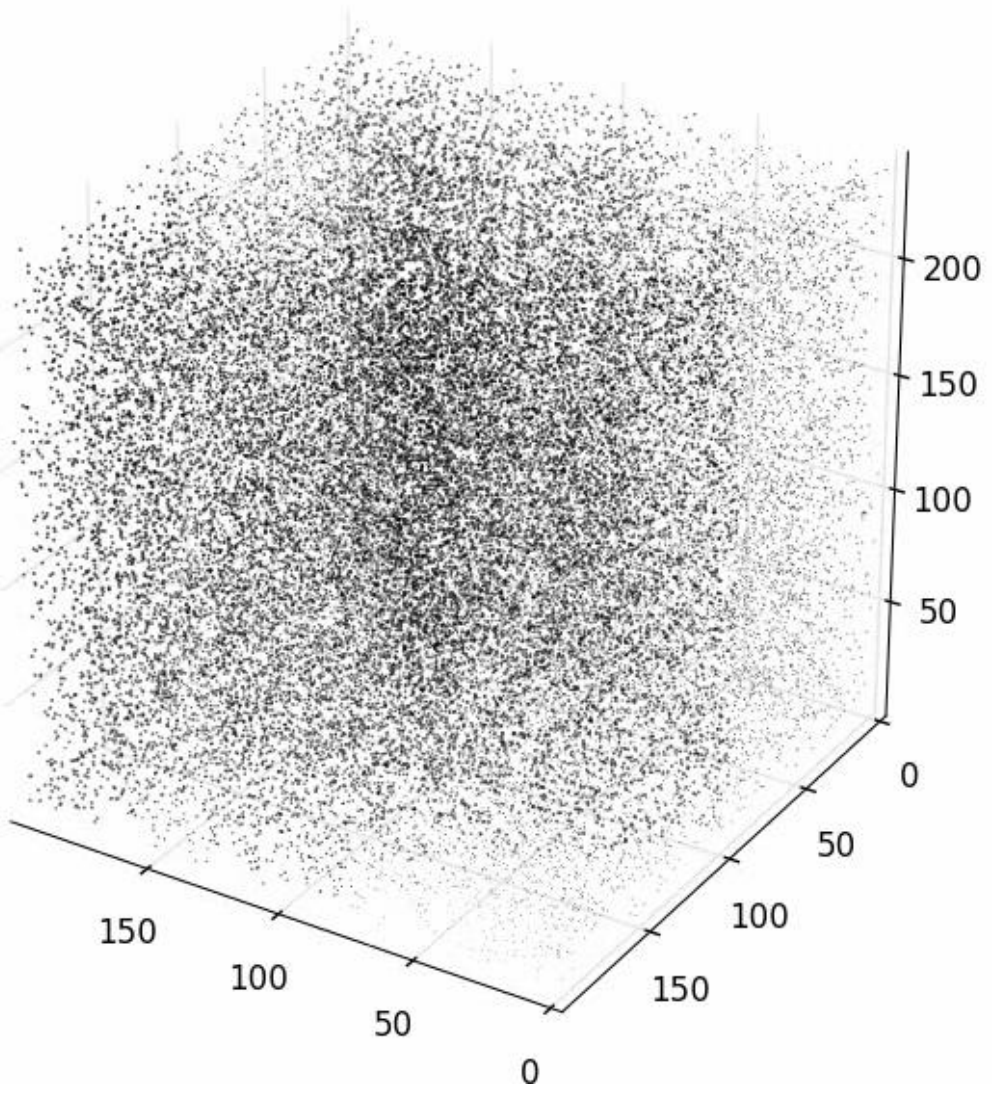
A quick aside…

- During the TPD of some materials, the QCM and mass spectral signals showed large differences
- The entrance to the mass spectrometer is located about ½” from the front surface of the QCM
- The mass spectrometer can only detect molecules that have desorbed from the QCM surface
- Viscoelastic effects can result in QCM signal artifacts during TPD



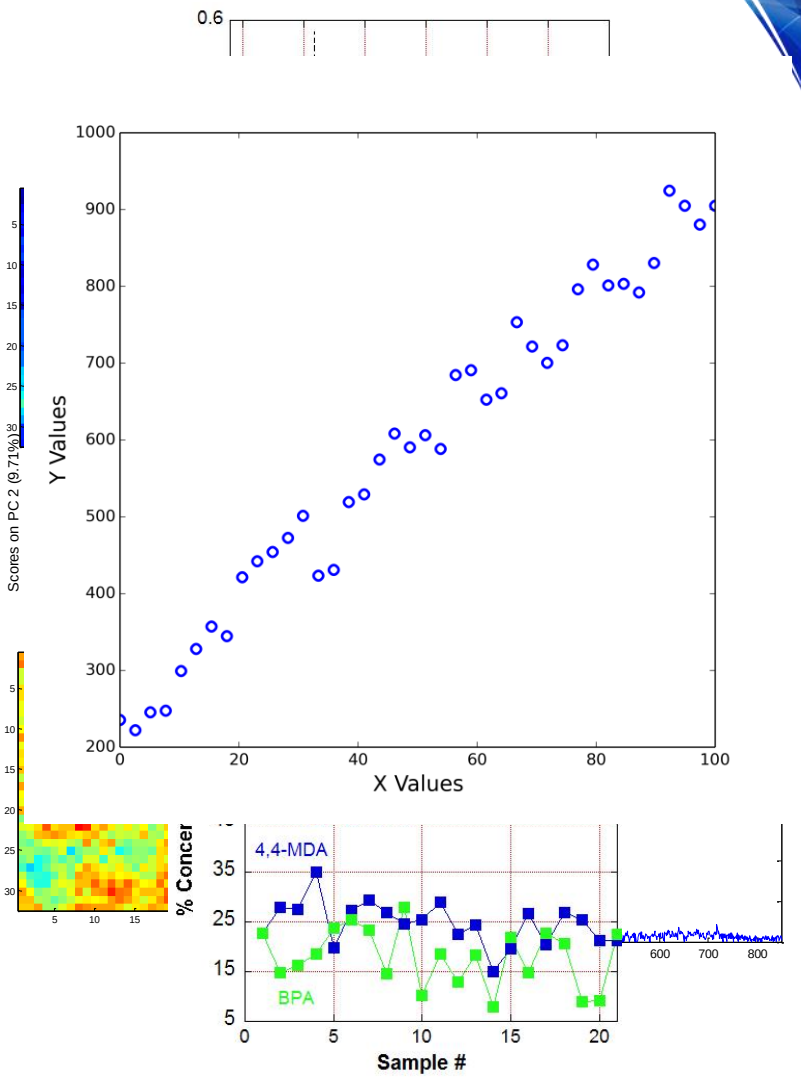


Dimensionality Reduction



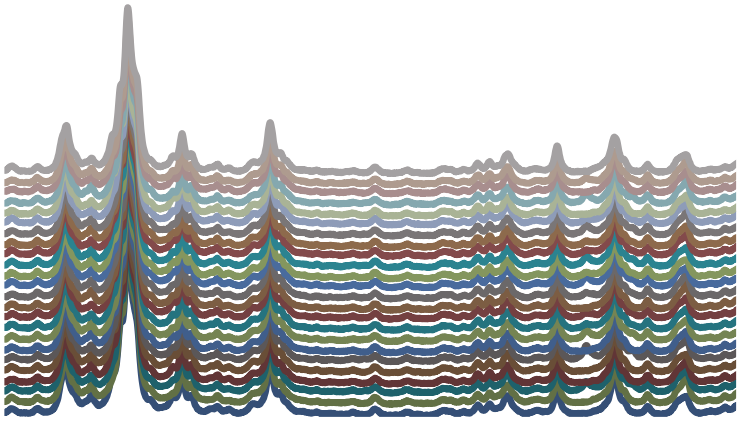
What is Multivatiate Analysis?

- Multivariate analysis (MVA) provides a means of summarizing a large number of variables into a smaller number of statistical variables
- MVA is a general term and includes approaches that can be used for:
  - *Identification: What species or components are present*
  - *Classification: Does my data fit into a category? Is it an outlier?*
  - *Quantification and Prediction: How does my data relate to a library of known samples*
- It relies on data changing as a function of a second variable

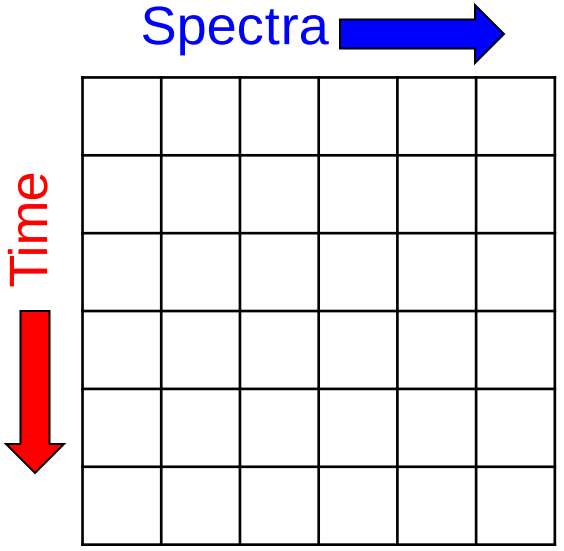


MVA Getting Started

- Need data suitable for MVA analysis
  - *Spectral changes as a function of time, spatial location, depth, etc…*

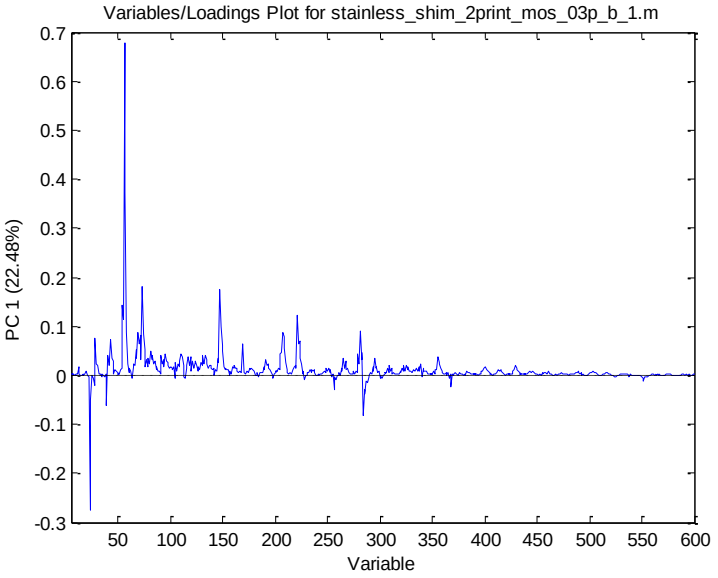


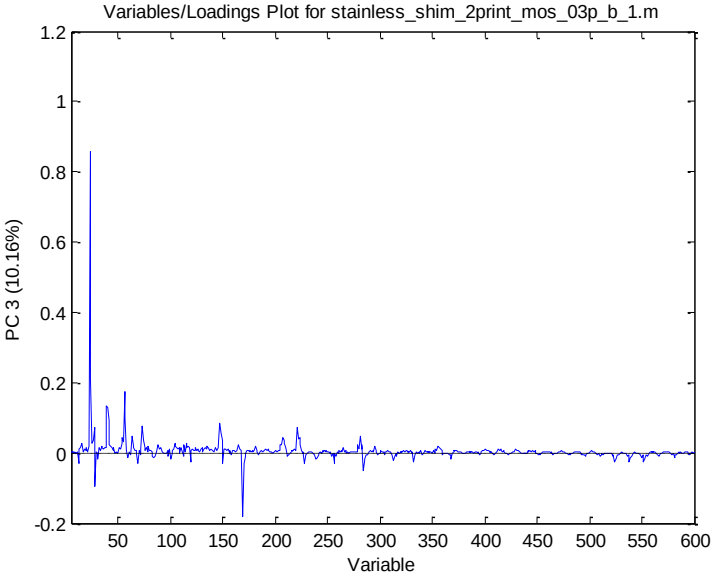
- MVA makes extensive use of matrix algebra
  - *Data needs to be arranged in row and column format prior to MVA analysis*



Principal Component Analysis (PCA)

- Data matrix is diagonalized yielding eigenvectors (loadings) and eigenvalues (scores) that capture maximum variance
- Scores and loadings mathematically constrained to be orthogonal
  - *Resulting spectra can have positive or negative intensity and have no physical/chemical meaning*
  - *Can be difficult to interpret*
- No input from user



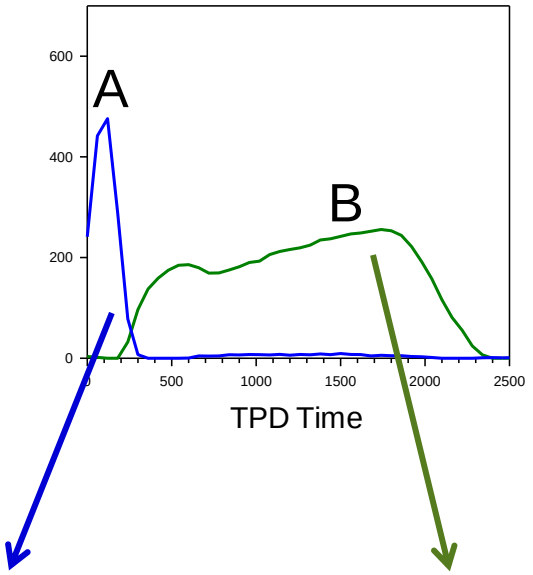


General Processing

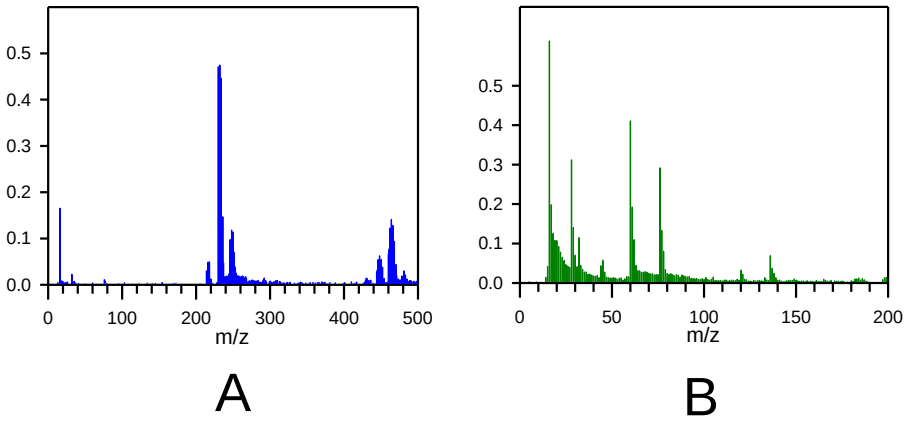
- For all mass spec data, the higher resolution data is binned into 1 amu width bins
  - *Why? simple to implement computationally*
  - *Python has fairly fast data structures (called dictionaries or hash tables) to deal with this type of data and make sorting, slicing, etc.. simple and rapid*
  - *Avoids most issues pertaining to calibration (If a peak shifts slightly during collection, is it really shifting or is a new peak growing in? How much does it need to shift before you call it a "new" feature instead of a continuation of a previous peak)*
  - *Even with the loss of resolution, these methods consistently provide a great deal of insight*
- Data is frequently Poisson scaled prior to MVA analysis
  - *This scaling adjusts each variable (m/z channel) so that the level of noise is equal for all variables*
  - *This tends to enhance higher m/z signals that provide more characterization relative to the lower m/z signals*

Self Modeling Curve Resolution (SMCR)
Multivariate Curve Resolution (MCR)
Self Modeling Mixture Analysis (SMMA)

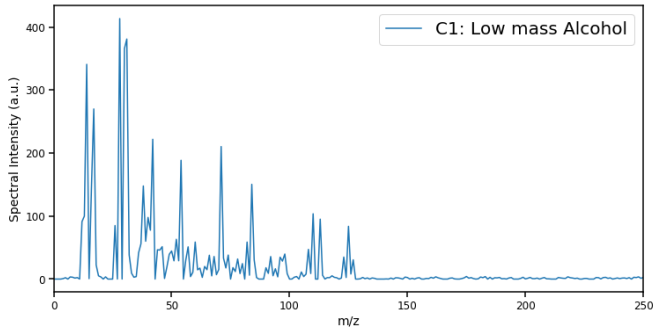
- Model provides results related to physical data
  - *Results look like real spectra*
  - *Easy to interpret*
- MCR uses minimal a priori information to decompose the data matrix into chemically meaningful factors

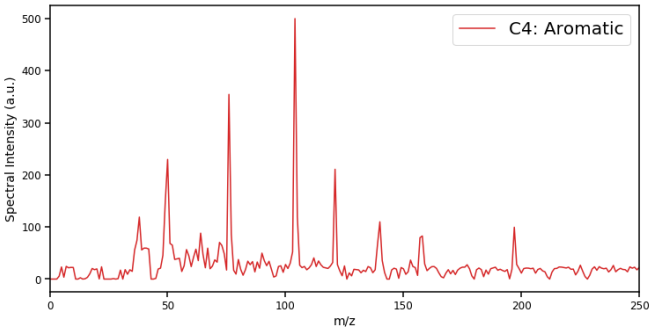


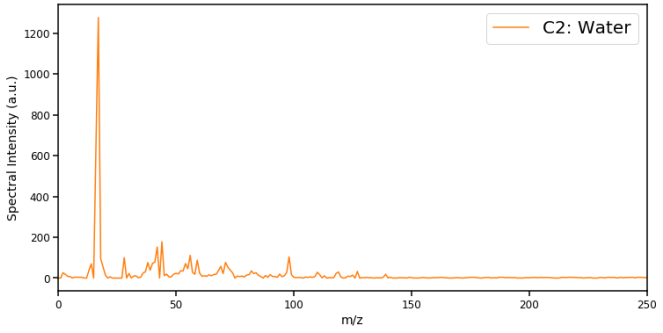
- MCR deconvolutes original data into a concentration matrix and associated "pure" component spectra
- User must supply number of components to be modeled

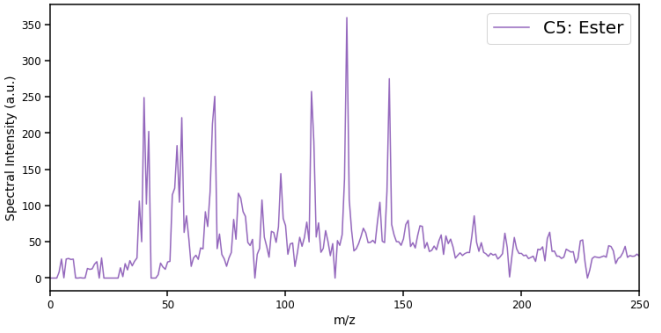


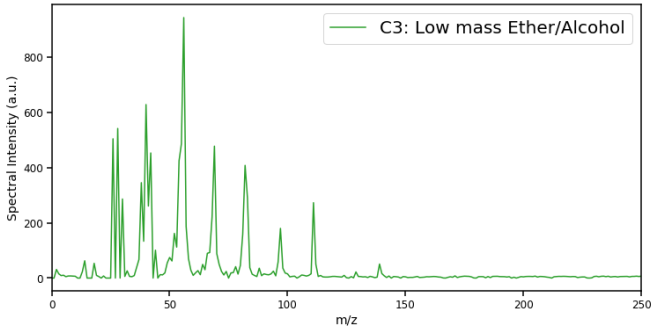
White Tape Component Spectra

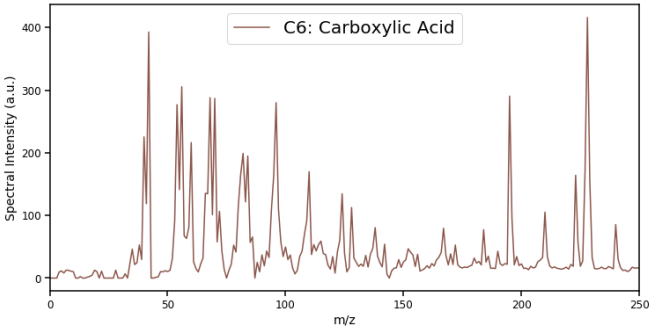




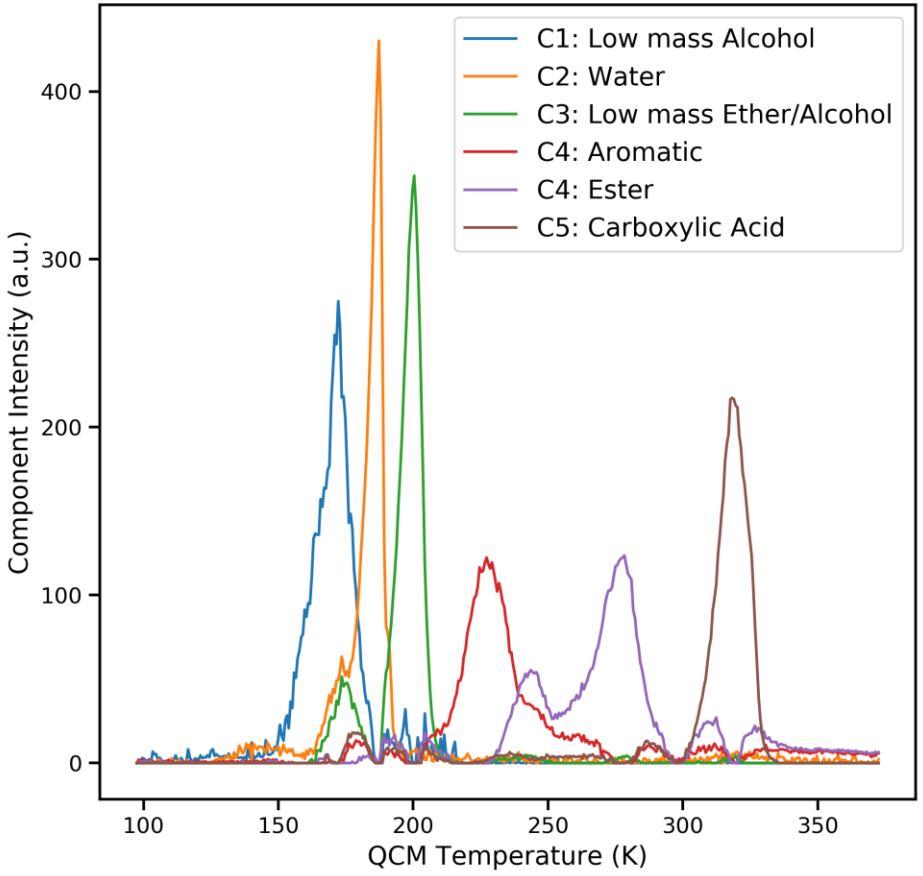


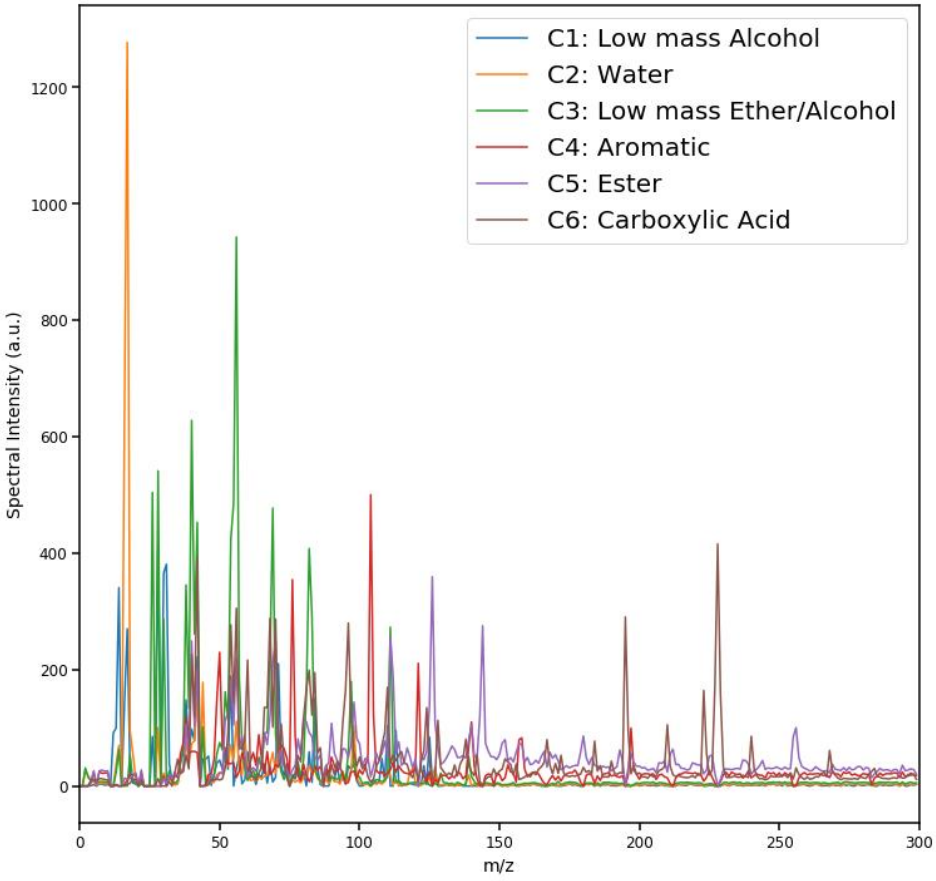






White Tape Component Spectra





Fluorene Contamination

Comparison of extracted components shows peak differences