Graphical Rotation- "Boosting" the Interpretability of NIR Methods

Neal B. Gallagher
Eigenvector Research, Inc.
nealg@eigenvector.com
Charles E. (Chuck) Miller, N.B. Gallagher, J.M. Shaver
R. Bro (University of Copenhagen)

Outline

- Exploratory Analysis of NIR- Why?
- Special NIR Considerations
- “Effect” Spectra vs. pure component spectra
- Introduction to Rotation
  - Different types
  - Rotation ***for NIR***
- NIR Case Studies
- Conclusions/Summary

EIGENVECTOR
RESEARCH INCORPORATED

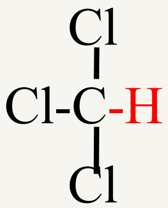
Exploratory Analysis in NIR

- Compliments deployed models
  - better understanding of system
    - physics and chemistry
    - i.e., the basis of deployed models
    - FDA, others: “What is NIR *really* monitoring?...”
  - Can enable improved understanding of NIR spectroscopy itself
    - Band assignment based on quantum chemistry cumbersome and inaccurate (anharmonicity)
- NIR advantages
  - Sampling versatility and non-destructiveness
    - Can study materials unaltered, in their “native” states!

EIGENVECTOR RESEARCH INCORPORATED

NIR Exploratory Considerations

- One “vibrating unit” can generate many bands!
- Bands from different analytes/phenomena typically ***not*** well-resolved!



$$\begin{array}{c} Cl \\ | \\ Cl\text{-}C\text{-}H \\ | \\ Cl \end{array}$$

combinations and overtones

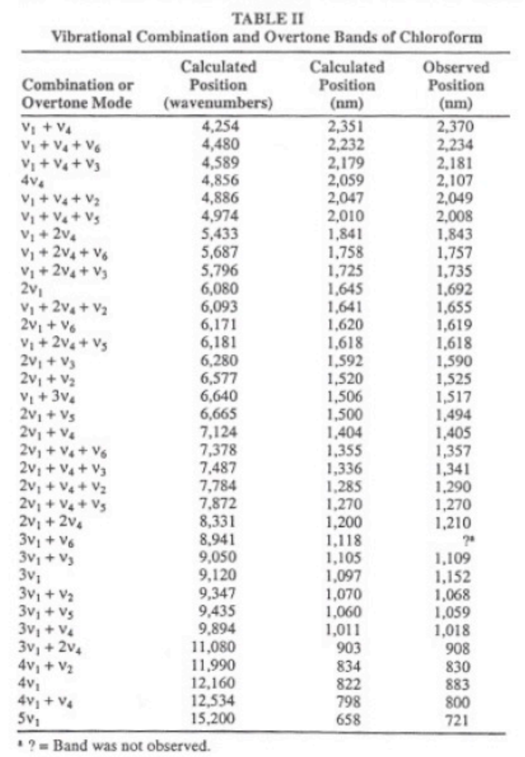
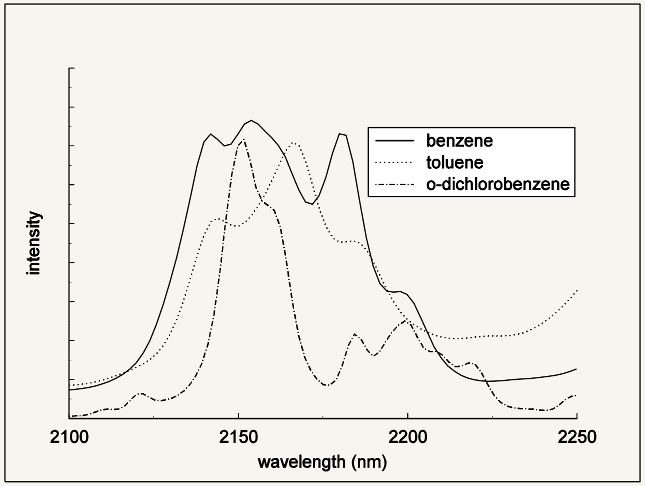


TABLE II
Vibrational Combination and Overtone Bands of Chloroform

Combination or Overtone Mode	Calculated Position (wavenumbers)	Calculated Position (nm)	Observed Position (nm)
$\nu_1 + \nu_4$	4,254	2,351	2,370
$\nu_1 + \nu_4 + \nu_6$	4,480	2,232	2,234
$\nu_1 + \nu_4 + \nu_3$	4,589	2,179	2,181
$4\nu_4$	4,856	2,059	2,107
$\nu_1 + \nu_4 + \nu_2$	4,886	2,047	2,049
$\nu_1 + \nu_4 + \nu_5$	4,974	2,010	2,008
$\nu_1 + 2\nu_4$	5,433	1,841	1,843
$\nu_1 + 2\nu_4 + \nu_6$	5,687	1,758	1,757
$\nu_1 + 2\nu_4 + \nu_3$	5,796	1,725	1,735
$2\nu_1$	6,080	1,645	1,692
$\nu_1 + 2\nu_4 + \nu_2$	6,093	1,641	1,655
$2\nu_1 + \nu_6$	6,171	1,620	1,619
$\nu_1 + 2\nu_4 + \nu_5$	6,181	1,618	1,618
$2\nu_1 + \nu_3$	6,280	1,592	1,590
$2\nu_1 + \nu_2$	6,577	1,520	1,525
$\nu_1 + 3\nu_4$	6,640	1,506	1,517
$2\nu_1 + \nu_5$	6,665	1,500	1,494
$2\nu_1 + \nu_4$	7,124	1,404	1,405
$2\nu_1 + \nu_4 + \nu_6$	7,378	1,355	1,357
$2\nu_1 + \nu_4 + \nu_3$	7,487	1,336	1,341
$2\nu_1 + \nu_4 + \nu_2$	7,784	1,285	1,290
$2\nu_1 + \nu_4 + \nu_5$	7,872	1,270	1,270
$2\nu_1 + 2\nu_4$	8,331	1,200	1,210
$3\nu_1 + \nu_6$	8,941	1,118	?[a]
$3\nu_1 + \nu_3$	9,050	1,105	1,109
$3\nu_1$	9,120	1,097	1,152
$3\nu_1 + \nu_2$	9,347	1,070	1,068
$3\nu_1 + \nu_5$	9,435	1,060	1,059
$3\nu_1 + \nu_4$	9,894	1,011	1,018
$3\nu_1 + 2\nu_4$	11,080	903	908
$4\nu_1 + \nu_2$	11,990	834	830
$4\nu_1$	12,160	822	883
$4\nu_1 + \nu_4$	12,534	798	800
$5\nu_1$	15,200	658	721

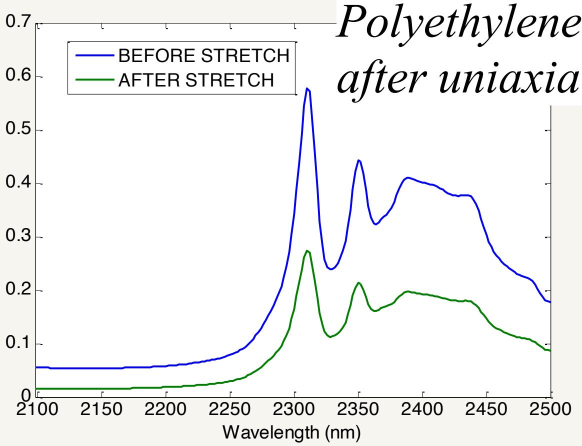
[a] ? = Band was not observed.

- however… model spectra are easy to obtain!

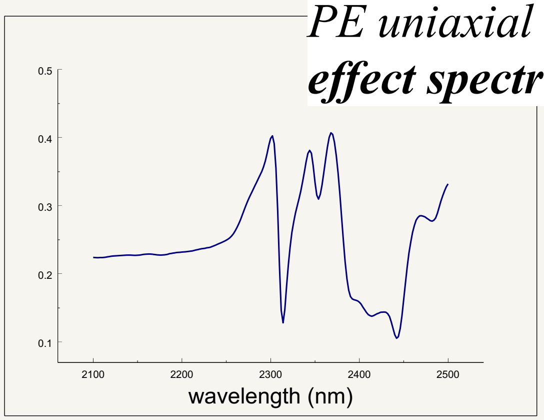


“Effect” Spectrum Concept

- Instead of a pure component spectrum
  - often due to physics instead of chemistry
  - temperature shifts, polymer alignment
- Often, a simple subtraction of two spectra can provide good insight
  - Reflects *differences* between two states
  - But, they require more thought: *Positive and Negative peaks* are meaningful
- Similar to interpretation of models built using *mean-centered* data



Polyethylene before and after uniaxial stretching



PE uniaxial stretching ***effect spectrum***

EIGENVECTOR RESEARCH INCORPORATED

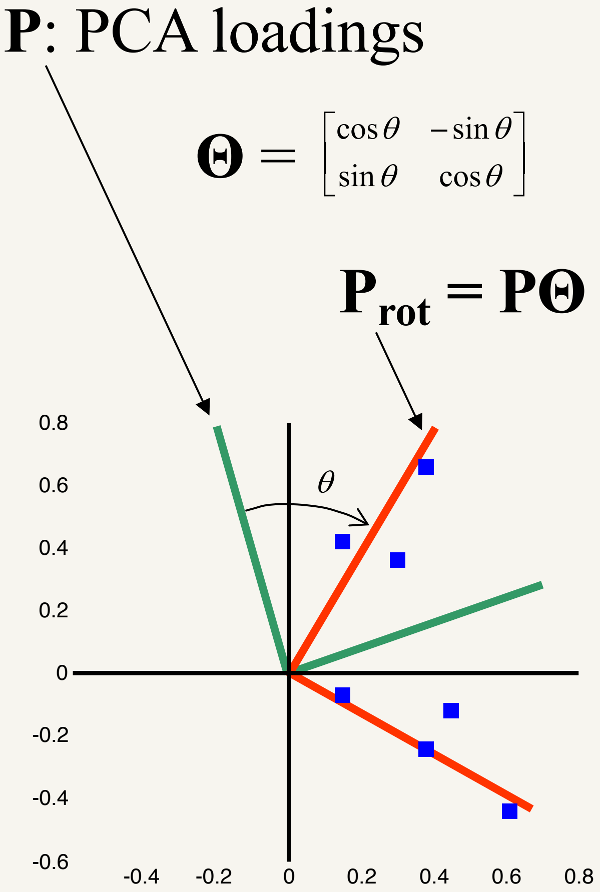
NIR Exploratory Checklist

- **Front End Loading (FEL1):**
  - **Experimental Design**
    - The data (to be modeled..)
    - “Special Data”: Pure components, model compounds, data from special experiments, etc….
  - **Background information:** samples, instrument, sampling interface, process stream, etc…
    - Prior knowledge to guide interpretations
  - ***Software that allows tracking of class information***
- **Attitude:** willingness to work with “effect spectra”
- **Mathematical Tools:**
  - Modeling Tools: PCA, CLS, MCR, PLS

EIGENVECTOR
RESEARCH INCORPORATED

Simple Orthogonal Rotation

- PCs from PCA constitute only one of infinitely many possible basis sets for a data set
    - Can be ***rotated*** to a more interpretable basis
    - However, must specify a ***target*** for this rotation
- Thurstone (1947), later authors: “Rotate PCs to attain simple structure”
    - Each loadings vector has only a few non-zero elements
    - “exclusivity”
    - In sociology, psychology- allows grouping of variables into underlying phenomena!



$\mathbf{P}$: PCA loadings

$$\boldsymbol{\Theta} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

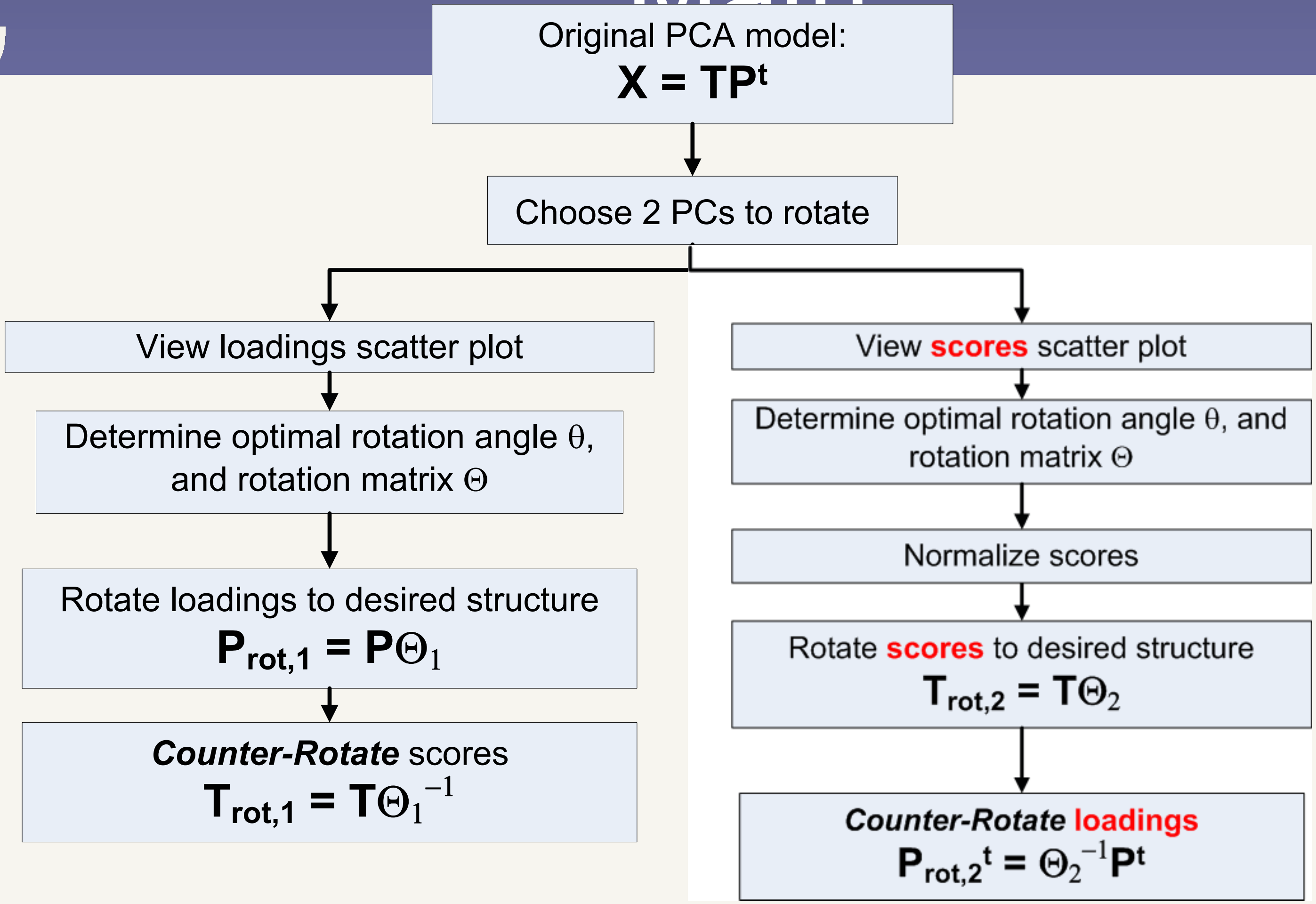
$$\mathbf{P}_{\mathbf{rot}} = \mathbf{P}\boldsymbol{\Theta}$$

Rotation- Options

- **Graphical rotation**
  - typical to use two vectors (PCs) at a time
- ***Orthogonal* Rotation**
  - Preserves orthogonality of vectors & explained variance structure
  - ***Varimax*** (Kaiser, 1958): specific target : ***max***imize ***vari***ance of loadings; ***Quartimax*** (1954), ***Equimax*** (1961)
  - ***"Graphical Rotation"*** (Windig 1987, mass spectrometry)
    - Target based on preferred structure of ***loadings***
- ***Oblique* Rotation**
  - Orthogonality ***not*** preserved, but more flexible
  - Lawton and Sylvestre (1971) →curve resolution/MCR
    - Loadings all positive- more intuitive
  - ***Simplimax:*** Kiers (1994) Maximize "simplicity" of rotated pattern

EIGENVECTOR
RESEARCH INCORPORATED

Graphical Rotation with PCA-Math



Original PCA model:

$\mathbf{X} = \mathbf{TP}^t$

Choose 2 PCs to rotate

View loadings scatter plot

Determine optimal rotation angle θ, and rotation matrix Θ

Rotate loadings to desired structure

$\mathbf{P}_{\mathrm{rot},1} = \mathbf{P}\Theta_1$

Counter-Rotate scores

$\mathbf{T}_{\mathrm{rot},1} = \mathbf{T}\Theta_1^{-1}$

View **scores** scatter plot

Determine optimal rotation angle θ, and rotation matrix Θ

Normalize scores

Rotate **scores** to desired structure

$\mathbf{T}_{\mathrm{rot},2} = \mathbf{T}\Theta_2$

Counter-Rotate **loadings**

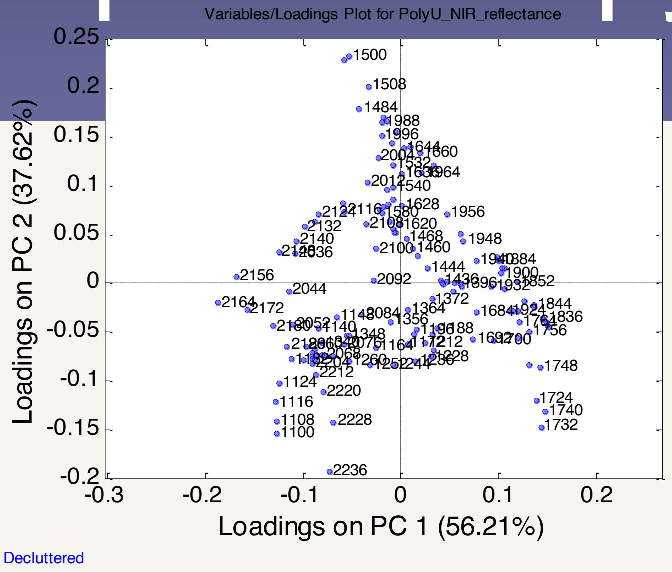
$\mathbf{P}_{\mathrm{rot},2}^t = \Theta_2^{-1}\mathbf{P}^t$

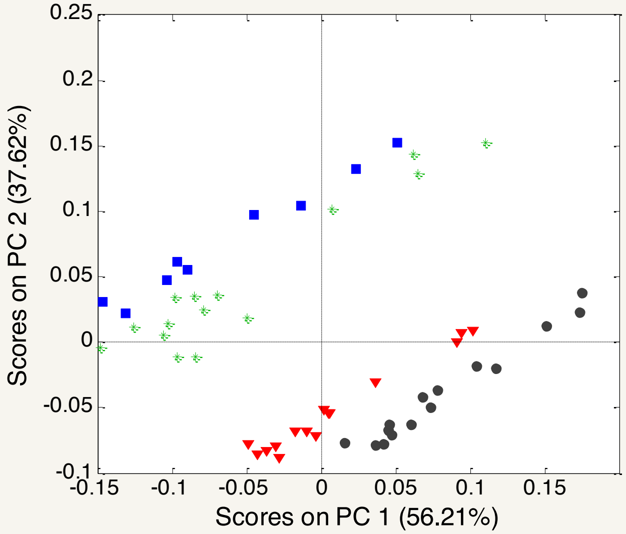
Rotation is NOT just "forcing" what you want: ***must counter-rotate*** *the other vector set in order to explain the original data!*

EIGENVECTOR RESEARCH INCORPORATED

Rotation in NIR Spectroscopy

- A rotation target based on structure in ***loadings*** can be problematic!
  - Pure components & effects often do ***NOT*** involve isolated, discrete bands!
- So- why not use a rotation target based on structure in ***scores***?....

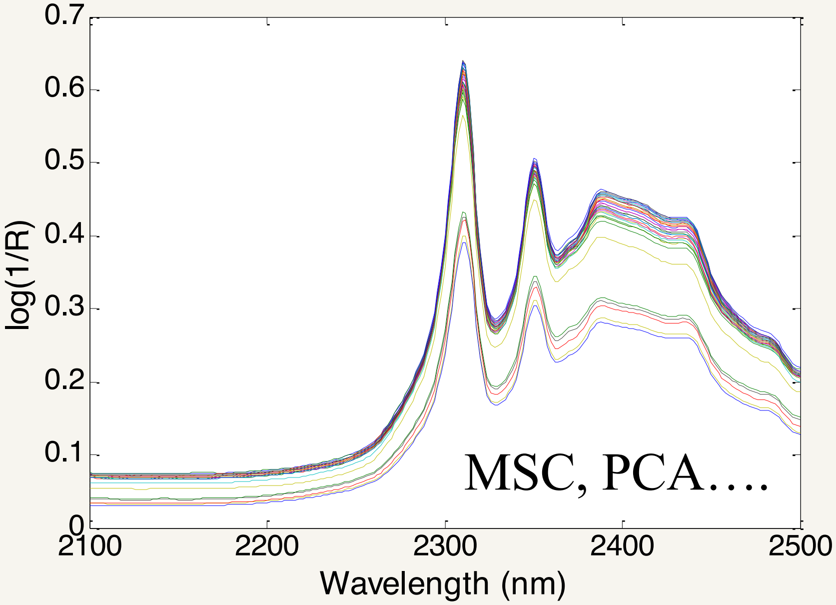




EIGENVECTOR
RESEARCH INCORPORATED

Case Study 1- Polyethylene blend films

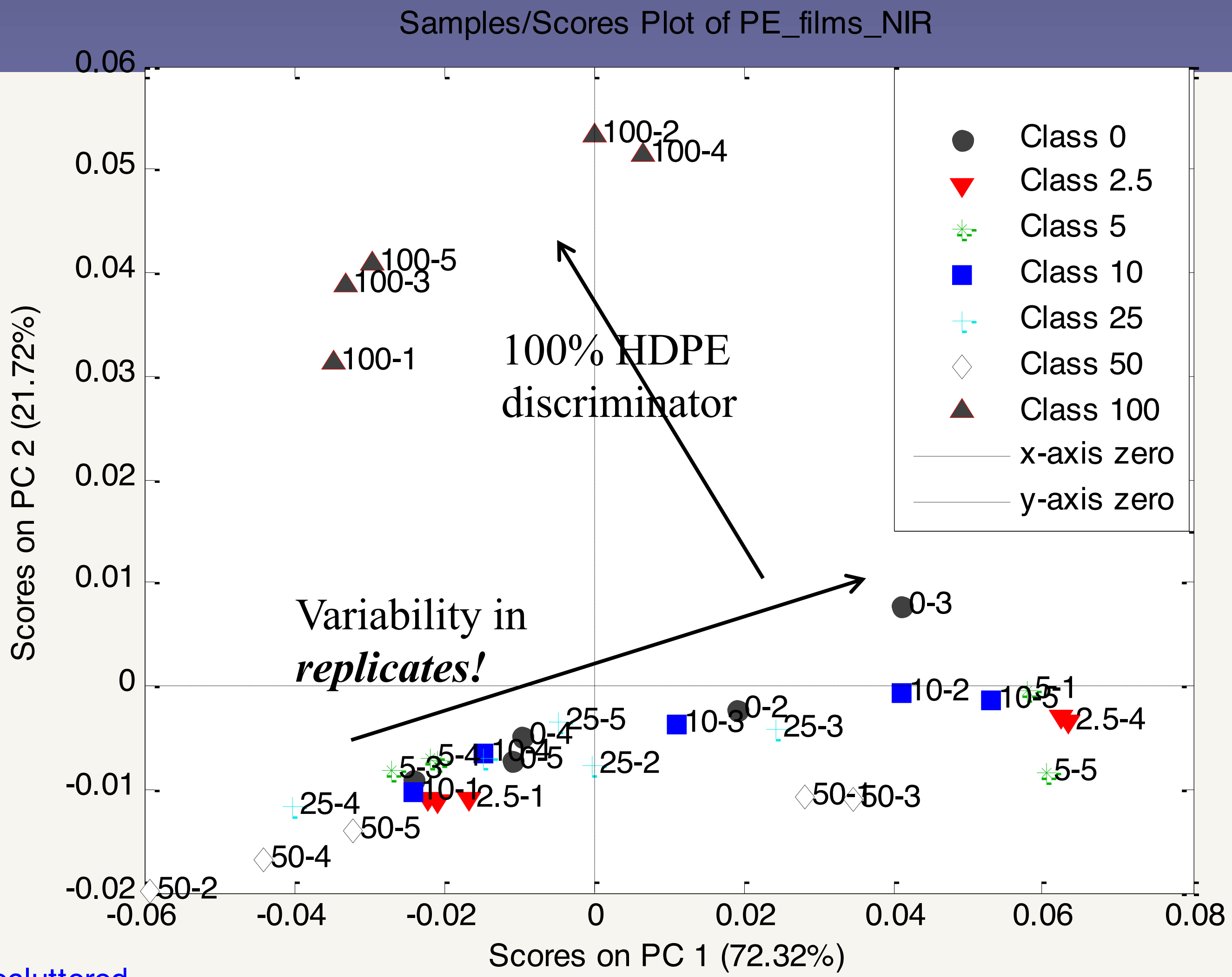
- 7 different blends of low- and high-density polyethylene (Statoil A/S)
  - 0, 2.5, 5, 10, 25, 50 and 100% HDPE
- Extruded into thin films
  - 25-40 μm thick
- 5 replicates of each composition analyzed by NIR (Technicon 500)



MSC, PCA….

Ref: C. Miller, Appl. Spectrosc., 47(2), 222 (1993).
PE_films_NIR.mat

Polyethylene- PC 1 & 2 scores



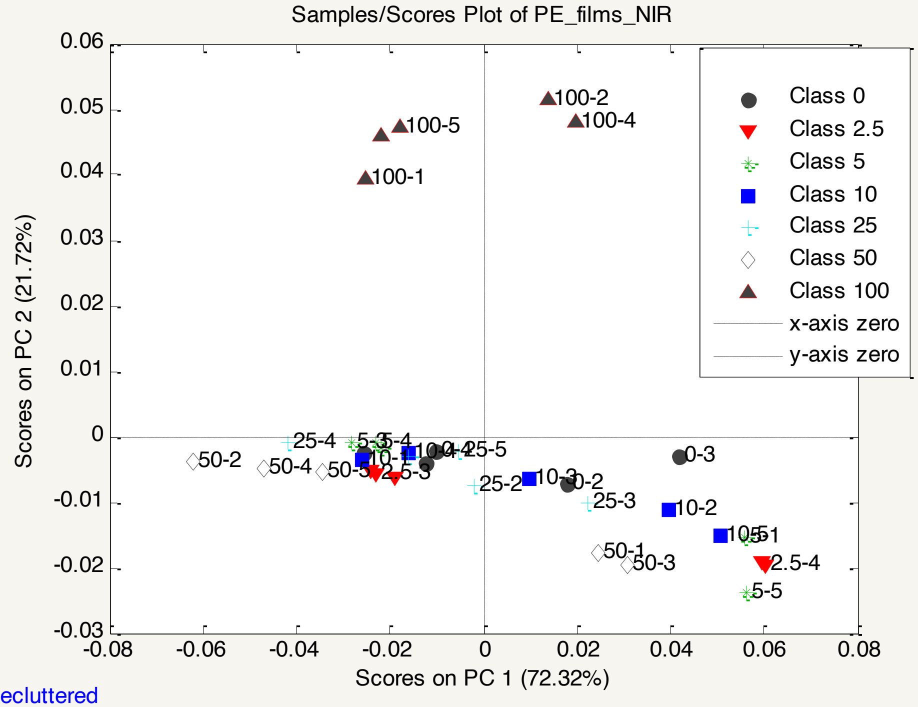
Samples/Scores Plot of PE_films_NIR

100% HDPE discriminator

Variability in ***replicates!***

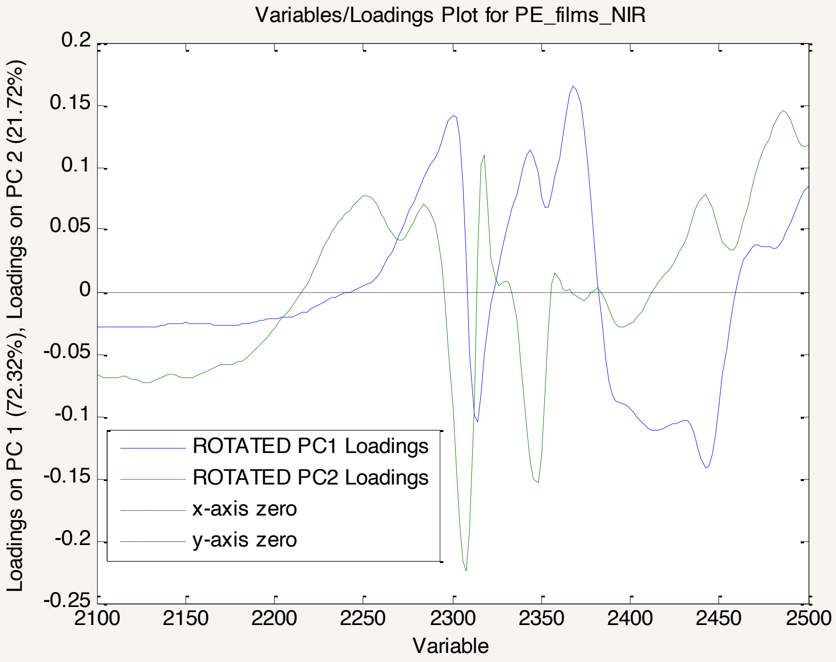
Decluttered

Rotate PCs 1&2 +15°



Decluttered

Resultant (counter-rotated) loadings!

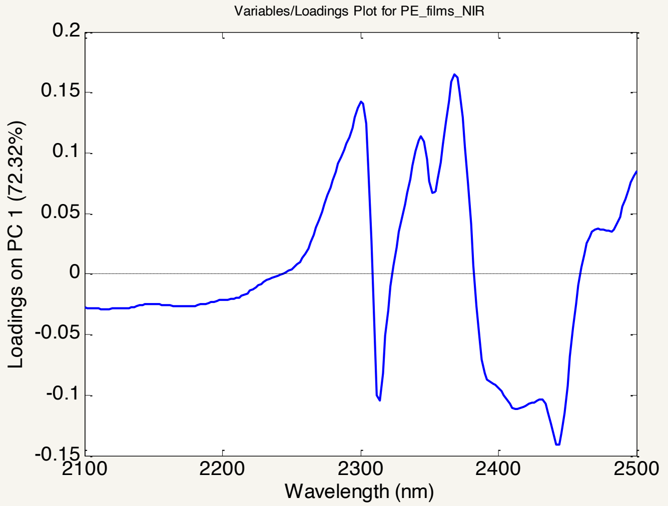


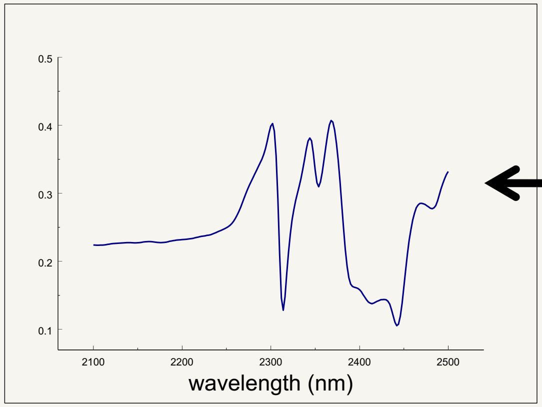
Effects, from counter-rotated loadings

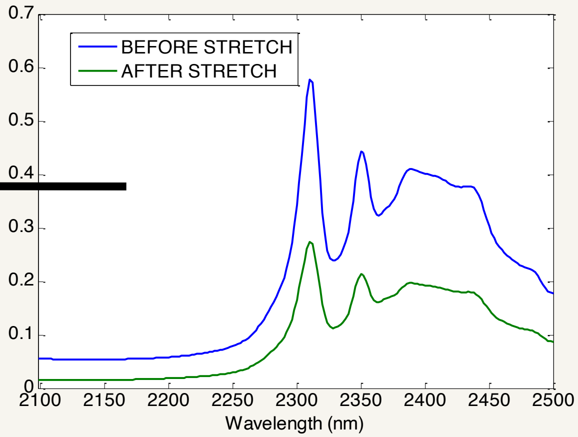
Counter-rotated loading 1 (replicate effect)

Weighted subtraction (stretching effect)

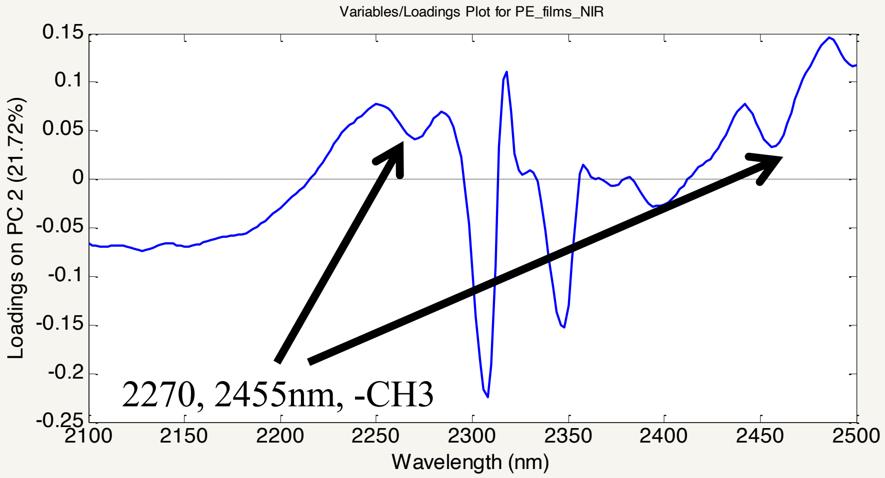
PE film before and after uni-axial stretching







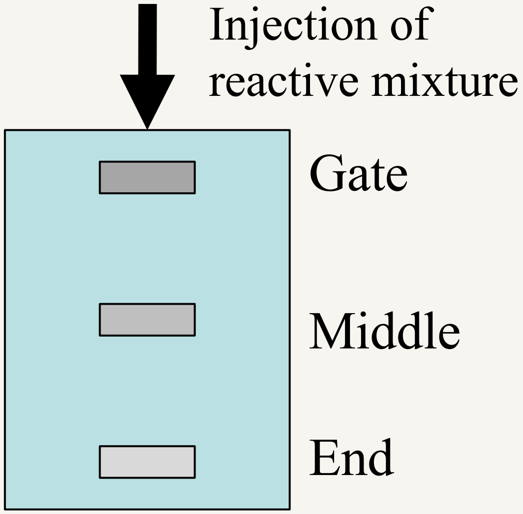
Counter-rotated loading 2 (100% HDPE discriminator)



2270, 2455nm, -CH3

Case Study - Polyurethanes

- Polyurethanes produced via reaction-injection-molding
  - Sheets produced at 4 nominal compositions
    - 42.5, 46, 51 & 55 % hard block
  - 3 sample strips cut from each sheet at different locations
  - NIR diffuse reflectance
    - Technicon I/A 500

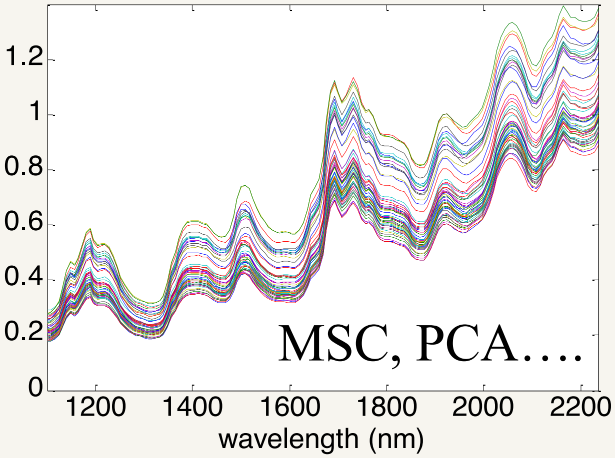


Injection of reactive mixture

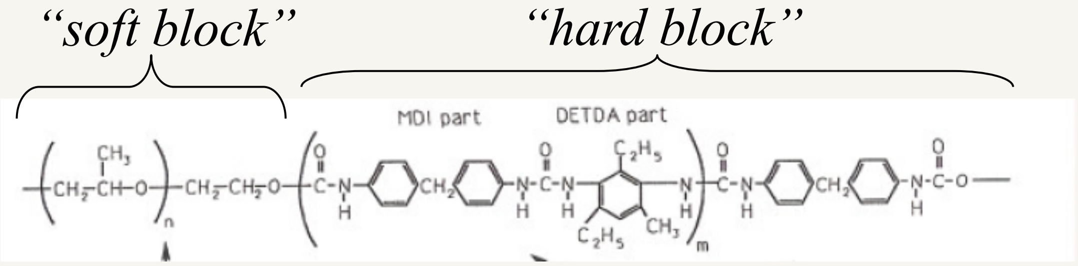
Gate

Middle

End



MSC, PCA….

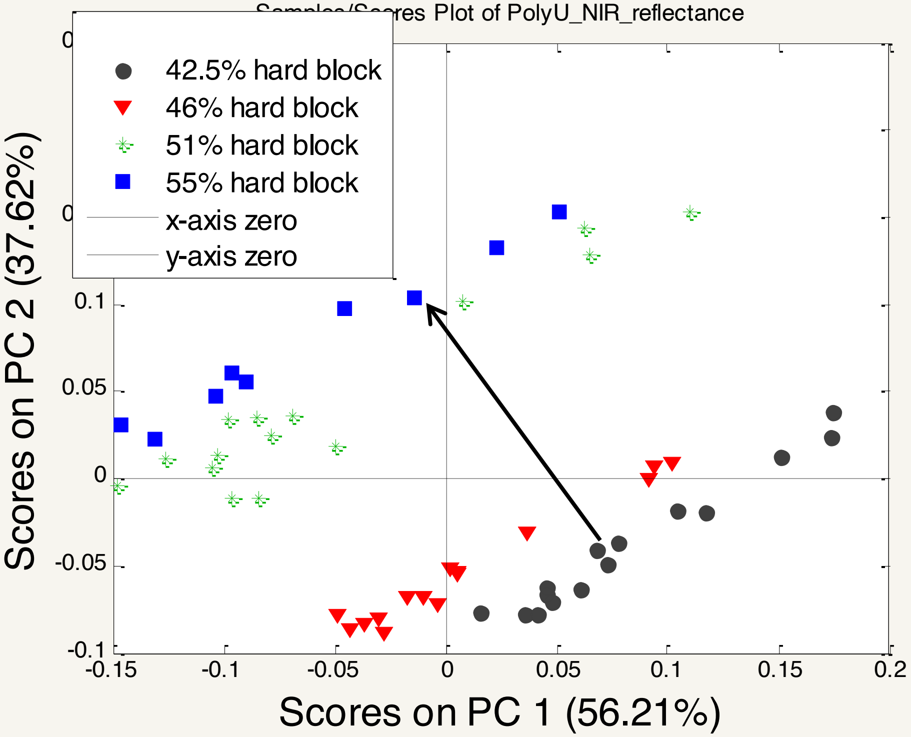


C. Miller, B. Eichinger, J. Appl. Poly Sci., 42, 2169 (1991).

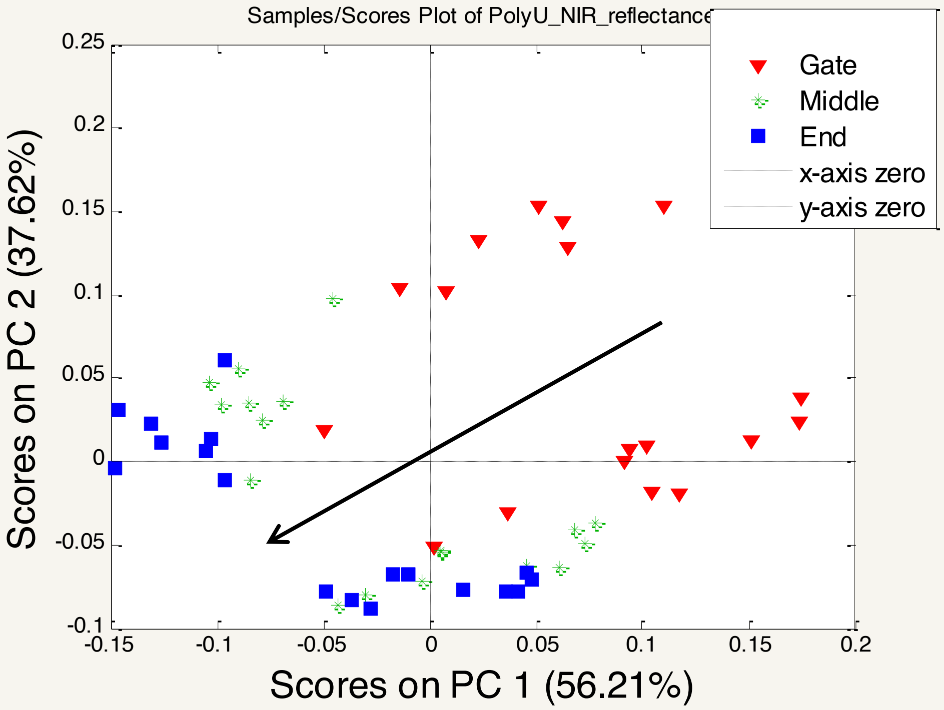
EIGENVECTOR RESEARCH INCORPORATED

PCA scores 1 & 2

By hard block %



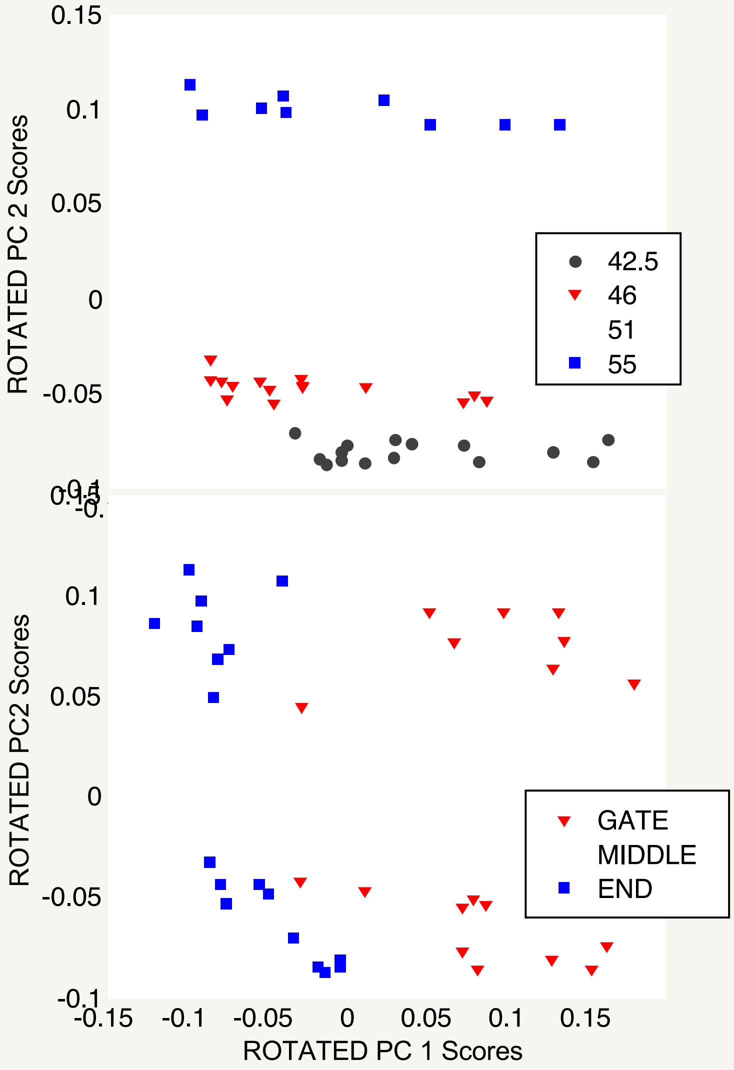
By sheet position



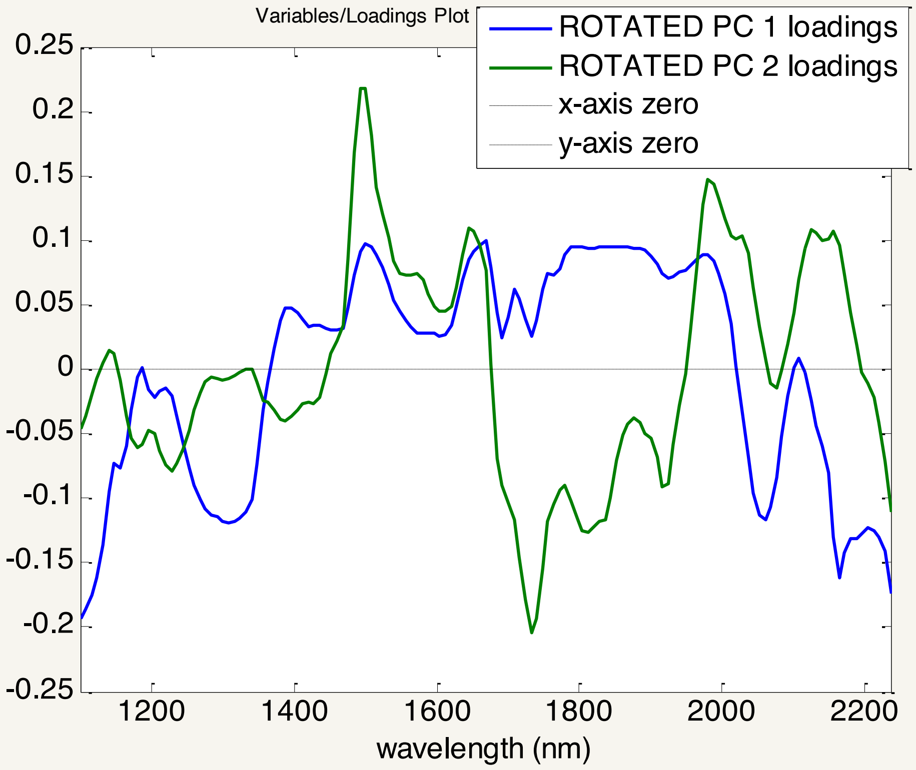
"% hard block" and "sheet position" effects appear to be orthogonal (in this 2 PC-space)

EIGENVECTOR RESEARCH INCORPORATED

Rotate Scores 1&2 +37°



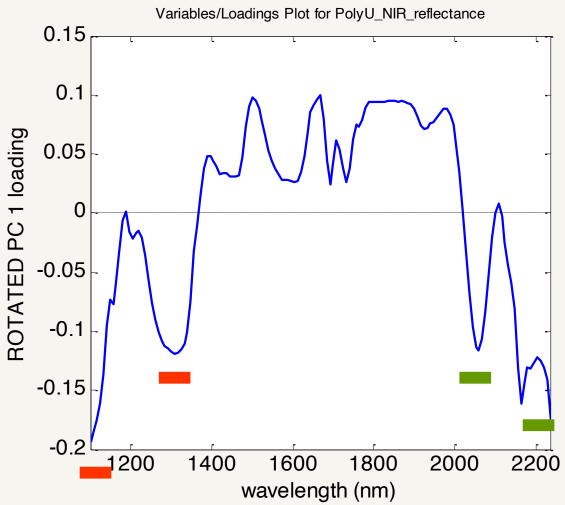
Rotated loadings

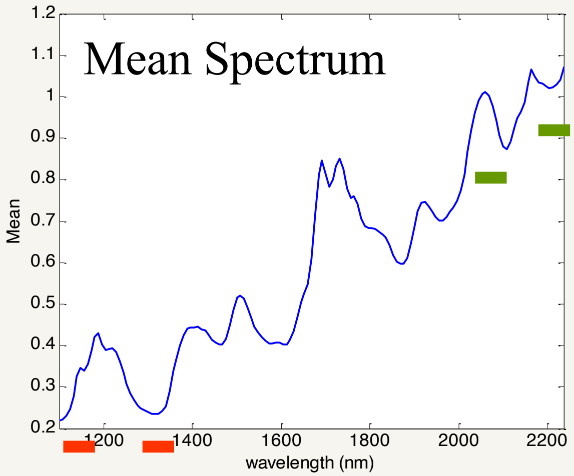


EIGENVECTOR RESEARCH INCORPORATED

Effects: Rotated Loadings

Rotated Loading 1 (sheet position effect)





Mean Spectrum

Effect:

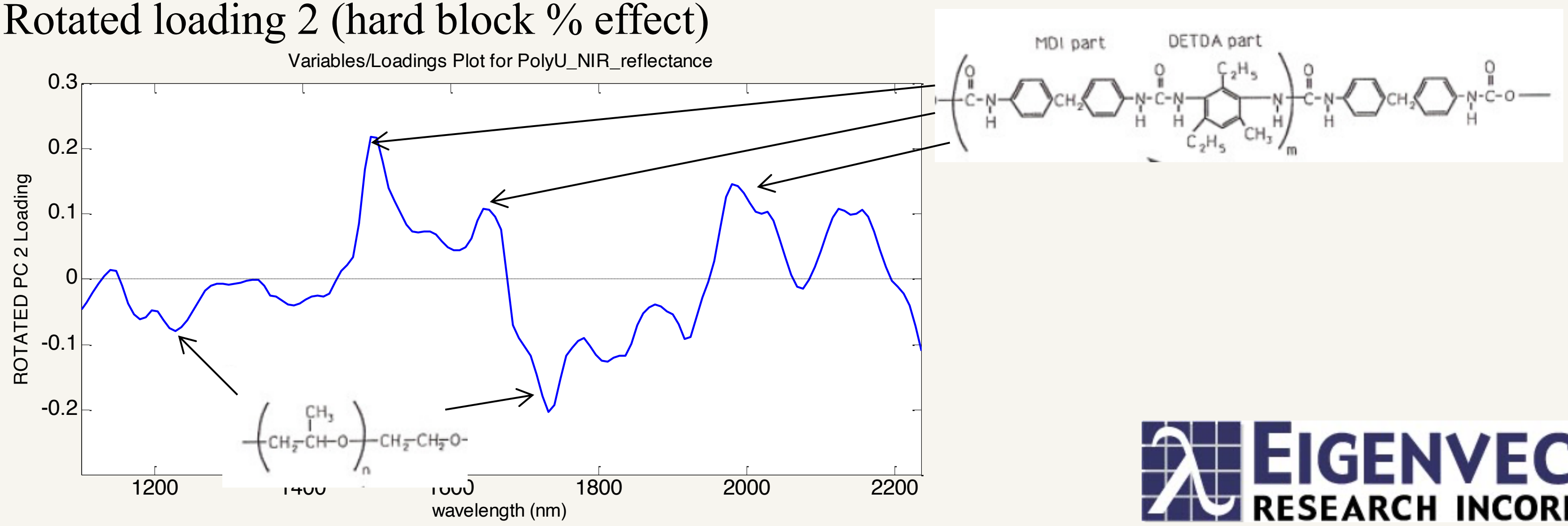
as score on PC 1 increases

a) valleys deeper at shorter wavelengths, and

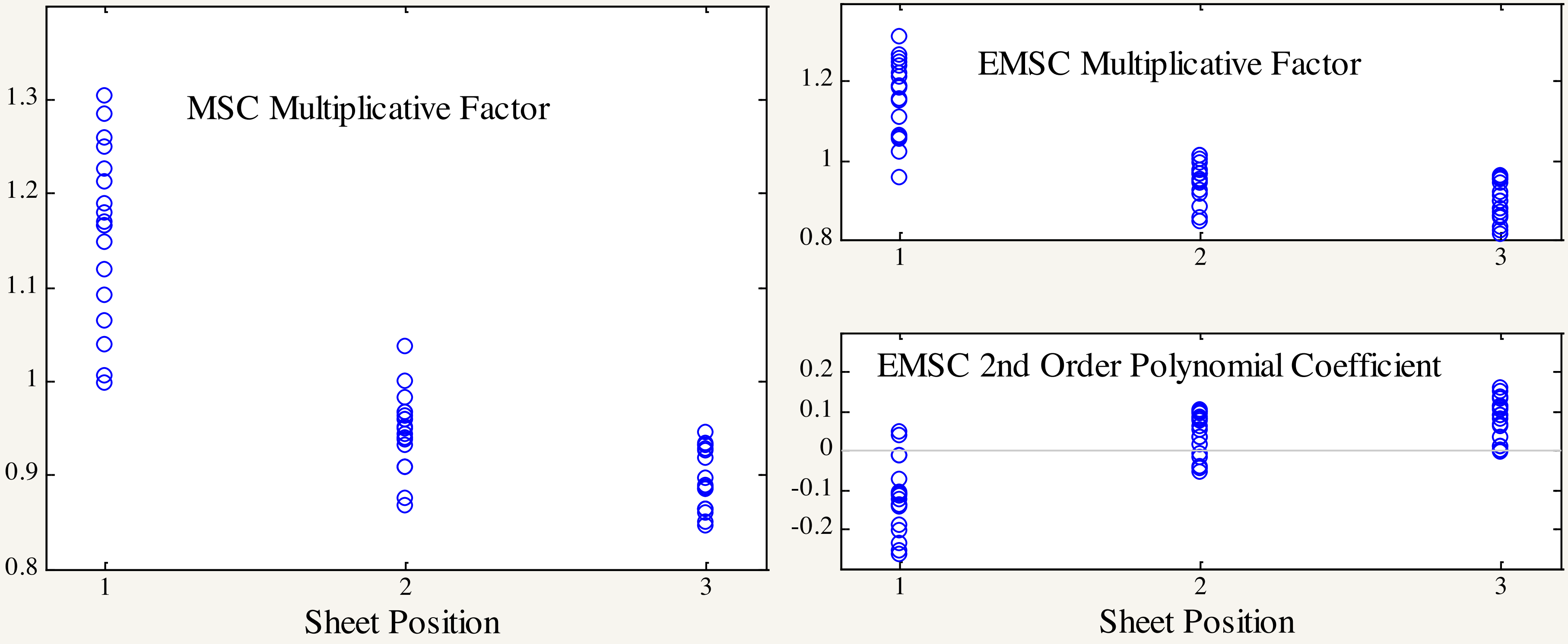
b) peaks shorter at longer wavelengths --

remember MSC!

Rotated loading 2 (hard block % effect)



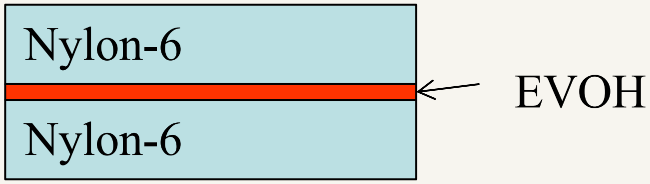
Loadings 1: Multiplicative

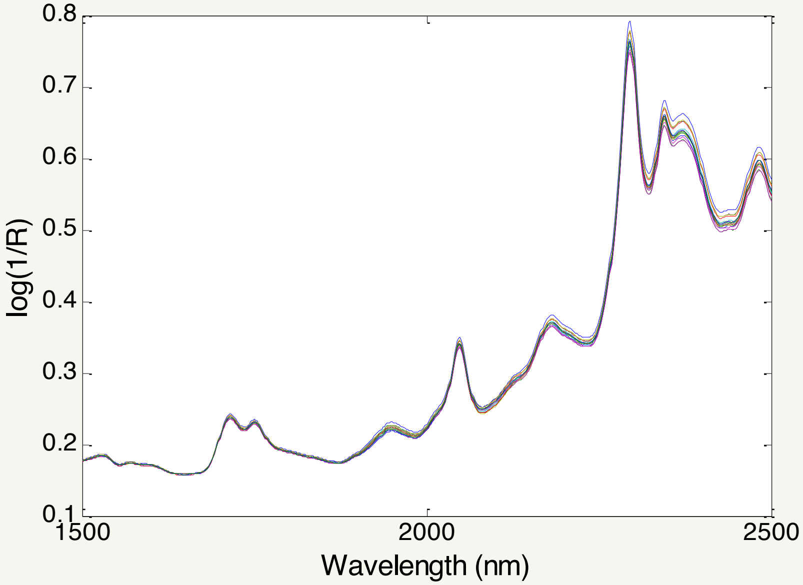


EIGENVECTOR RESEARCH INCORPORATED

Case Study – Nylon/EVOH/Nylon Films

- Used as moisture barrier in food packaging
  - Nylon-6/EVOH/Nylon-6 “sandwich”
  - 16 film samples cut from different regions of same extruded sheet
- Measured with NIR and polarized light microscopy
  - EVOH content varies from 3-8%

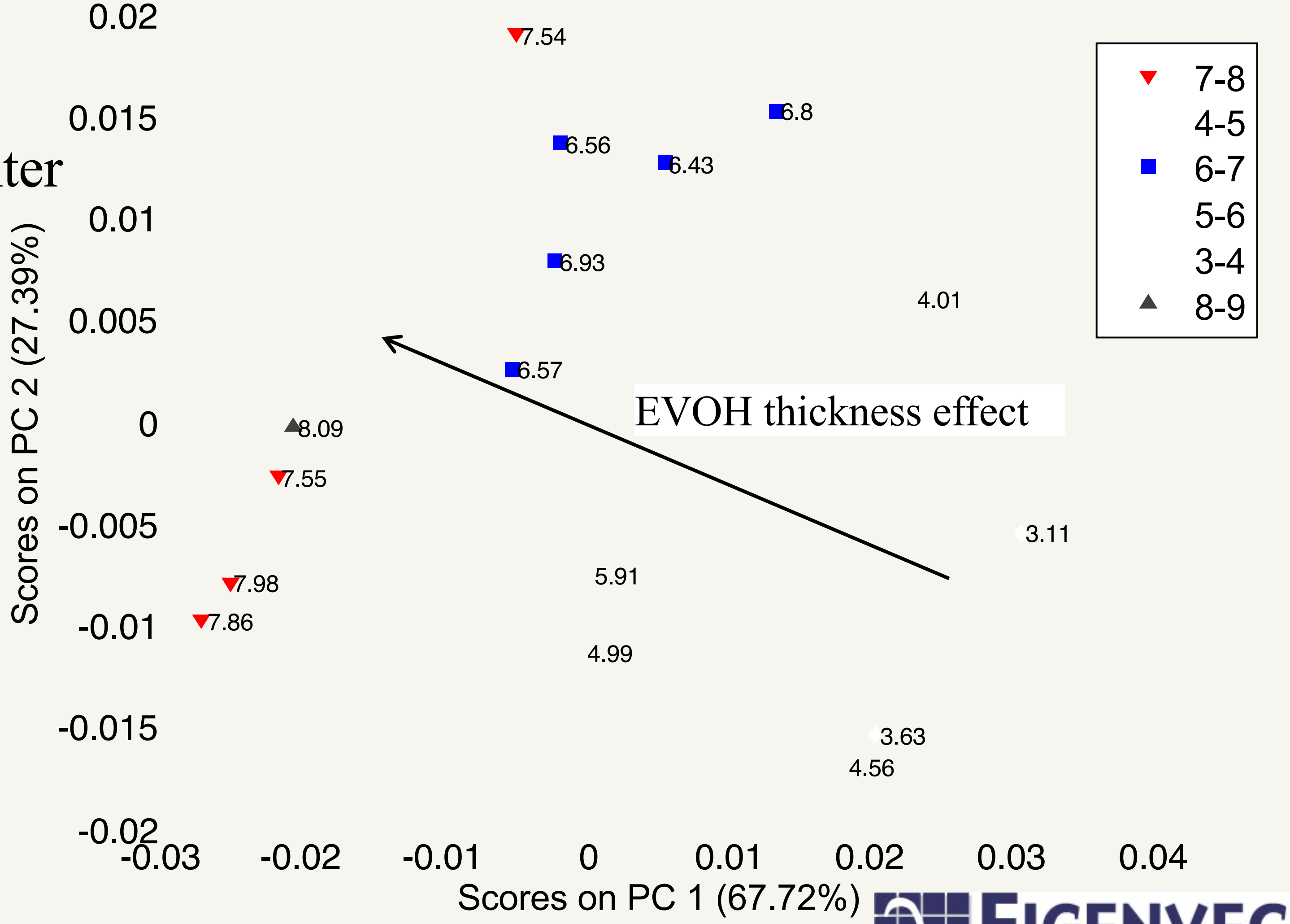




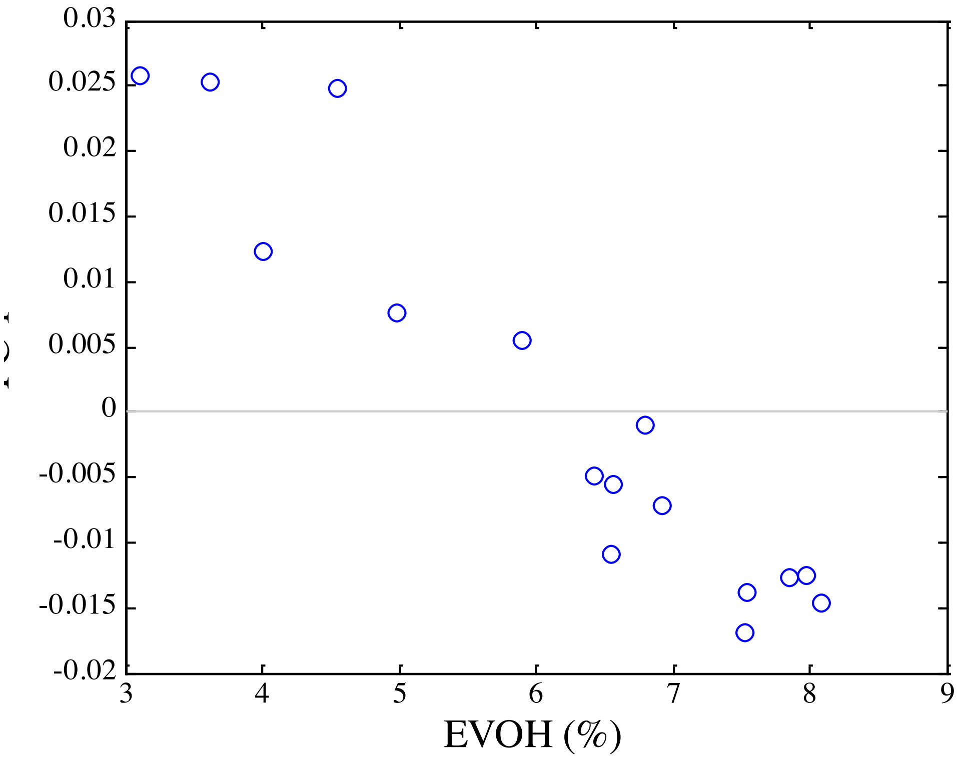
C. Miller S. Svendsen, T. Næs, Appl. Spectrosc., 47(3), 346 (1993).

PCA scores 1 & 2

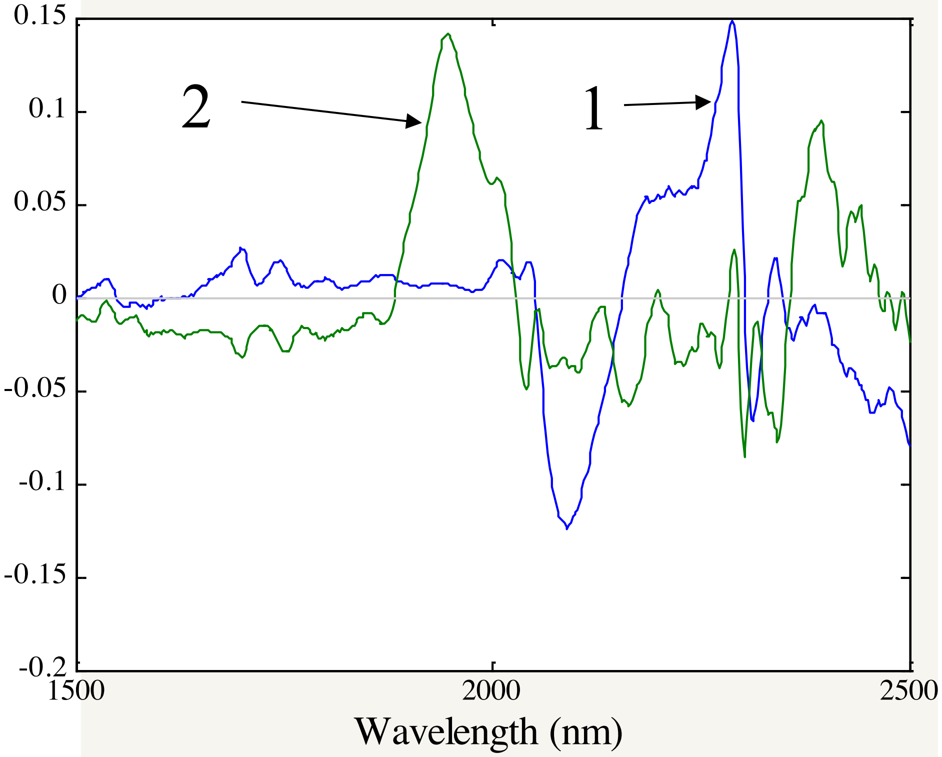
MSC
Mean-center



Rotate PCs 1 & 2: 45°



Rotated loadings



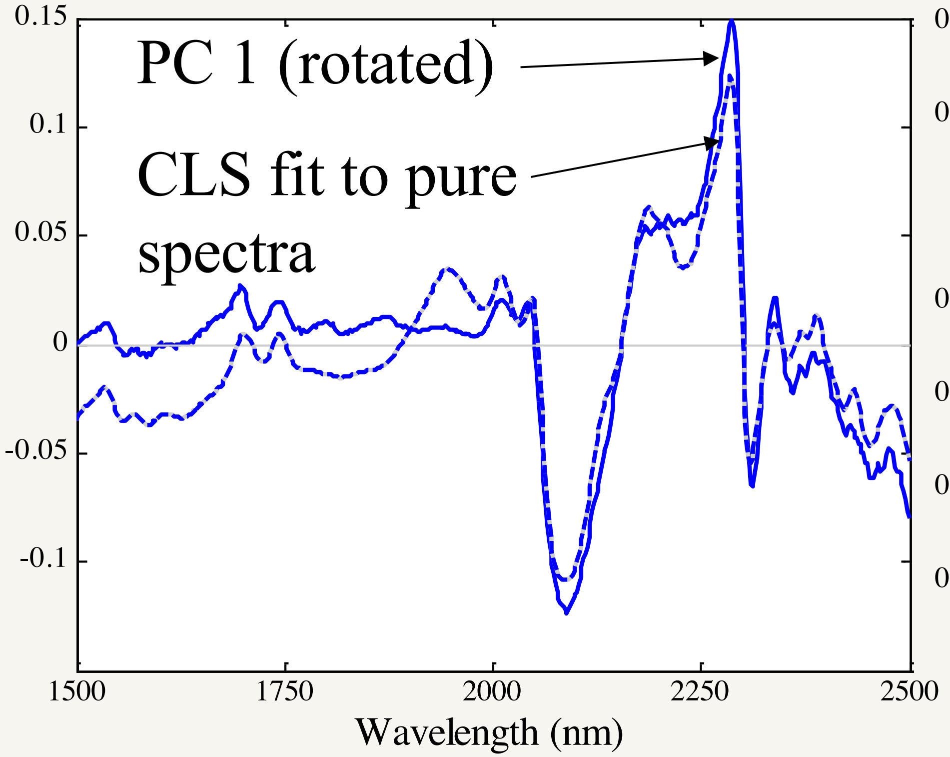
EIGENVECTOR RESEARCH INCORPORATED

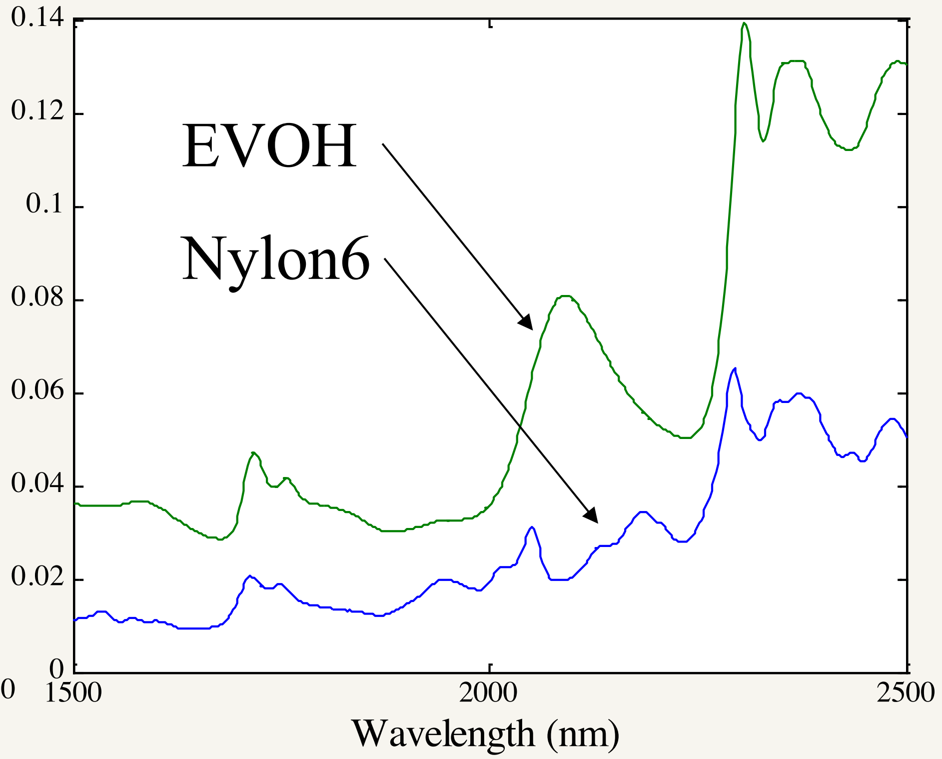
Effects: Rotated Loadings

[Nylon6 – EVOH]
(composition effect)

Rotated Loading 1
decrease of EVOH thickness effect

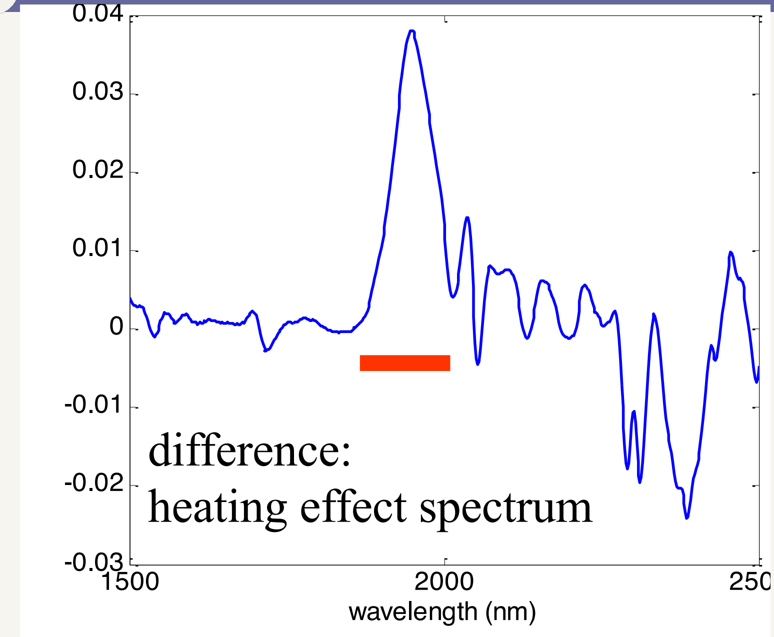
Nylon6 and EVOH pure component spectra

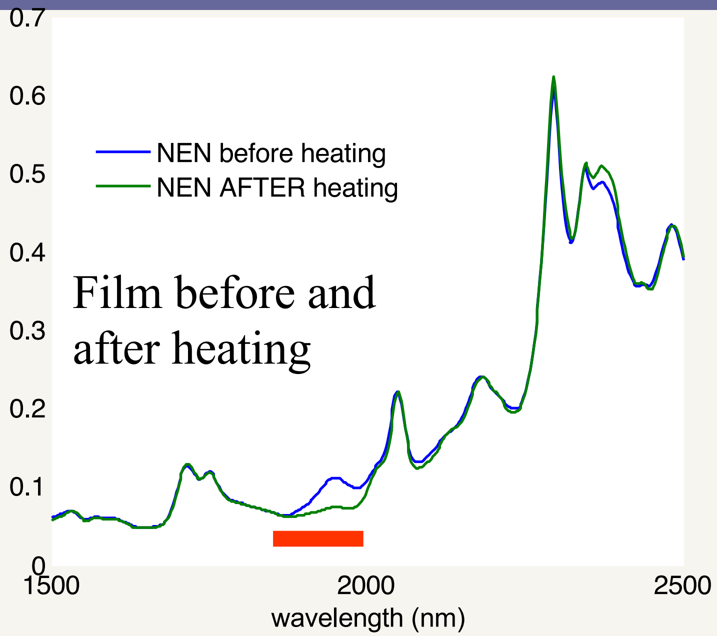


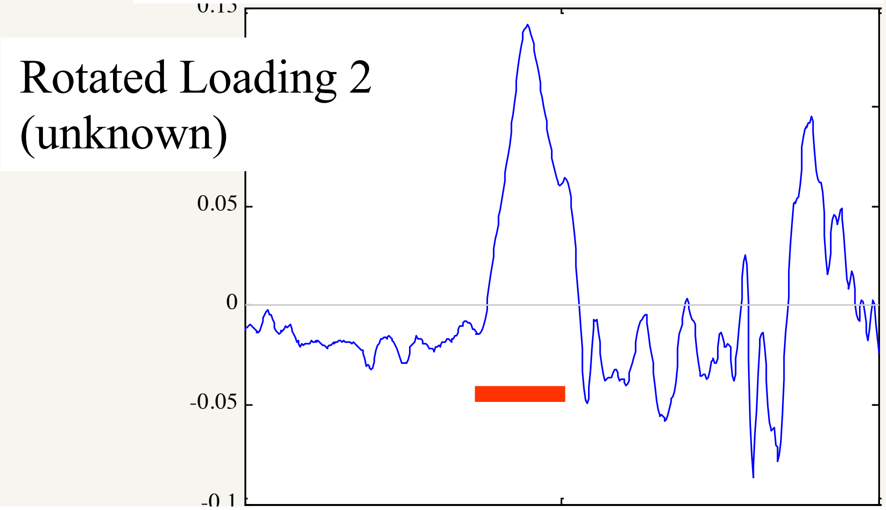


c = [6.61 -2.97]

Effects: Rotated Loadings


difference:
heating effect spectrum


Film before and
after heating


Rotated Loading 2
(unknown)

EIGENVECTOR
RESEARCH INCORPORATED

Summary 1 of 2

- Interpreting loadings from NIR is useful
  - can rotate to more interpretable spectra
  - can interpret chemical and
  - physical “effect” spectra
- Rotation
  - some forms are very simple
  - useful alternative for exploratory NIR
- In NIR
  - it is expected that rotation is more useful when target is based on structure in ***scores***, not loadings

EIGENVECTOR
RESEARCH INCORPORATED

Summary 2 of 2

- Can be extended to *other factor based analysis*
- Rotation concept is a good framework for development of “purity” methods
  - Oblique geometrical methods for self-modeling mixture analysis
    - e.g., SIMPLSMA, DISTSLCT
  - Good initial guess for least-squares based multivariate curve resolution