Graphical Rotation- "Boosting" the Interpretability of NIR Methods

Charles E. (Chuck) Miller, Neal B. Gallagher,
Jeremy M. Shaver

Eigenvector Research, Inc.

Rasmus Bro

University of Copenhagen, Faculty of Food Science

EIGENVECTOR RESEARCH INCORPORATED

Outline

- Exploratory Modeling in NIR- Why?
- Special NIR Considerations
- "Effect" Spectra (vs. pure component spectra)
- Introduction to Rotation
  - Different types
  - Rotation ***for NIR***
- 3 NIR Case Studies
- Conclusions/Summary

2

Exploratory Modeling in NIR- WHY?

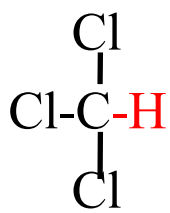
- Compliments deployed models
  - Better-understand the basis of deployed models
  - FDA, others: "What is NIR ***REALLY*** monitoring?..."
- Can enable improved understanding of NIR spectroscopy itself
  - Band assignment based on quantum chemistry cumbersome and inaccurate (anharmonicity)
- NIR as a ***RESEARCH*** tool- WHY NOT?
  - Sampling versatility and non-destructiveness
    - Can study materials unaltered, in their "native" states!

EIGENVECTOR RESEARCH INCORPORATED

3

Exploratory NIR- Considerations

- One "vibrating unit" can generate many bands!
- Bands from different analytes/phenomena typically ***not*** well-resolved!
- **BUT**… model spectra are easy to obtain!
  - More "direct" sampling options!

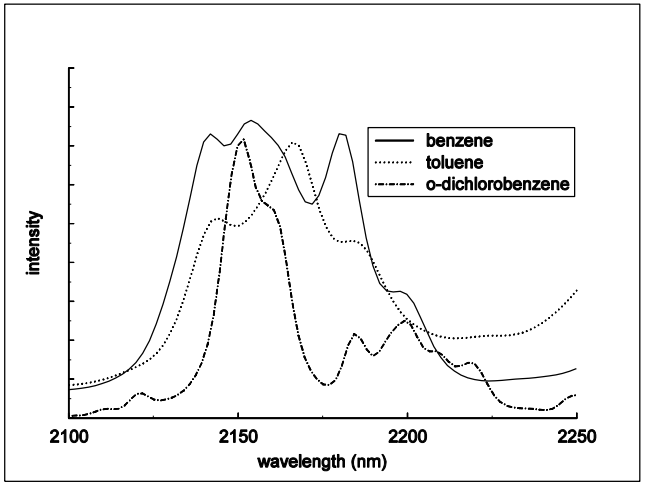


combinations and overtones

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.



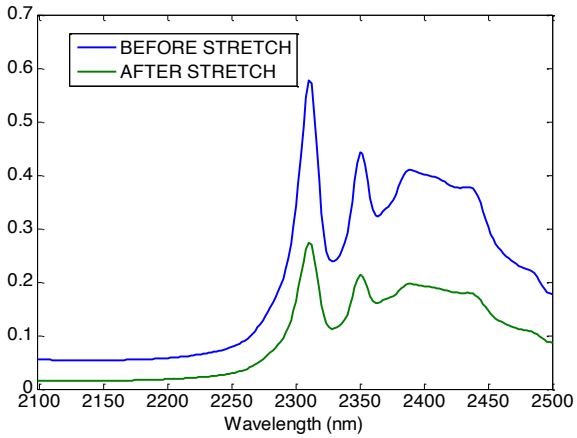
EIGENVECTOR RESEARCH INCORPORATED

4

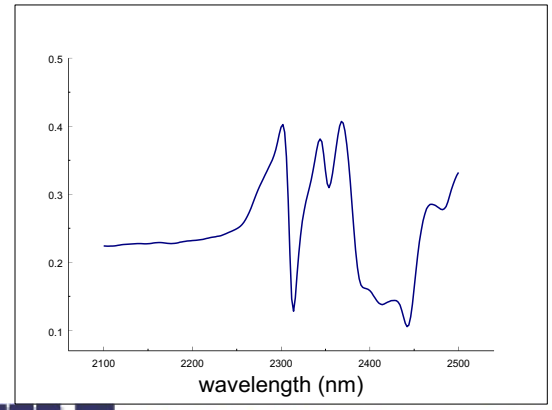
"Effect" Spectrum Concept

- Instead of a pure component spectrum!
- Often- a simple subtraction of two spectra
- Reflects ***differences*** between two states
- Can be used to express ***non-concentration*** properties
- But, they require more thought: ***Positive and Negative peaks*** are meaningful!
- Allows effective interpretation of models built using ***mean-centered*** data!
- Derivative preprocessing- hinders interpretation, adds more peaks!

Polyethylene before and after uniaxial stretching



PE uniaxial stretching ***effect spectrum***



EIGENVECTOR RESEARCH INCORPORATED

5

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
  - Post-Modeling Tools
    - ***Rotation!***

EIGENVECTOR RESEARCH INCORPORATED

6

Introducing…Rotation!

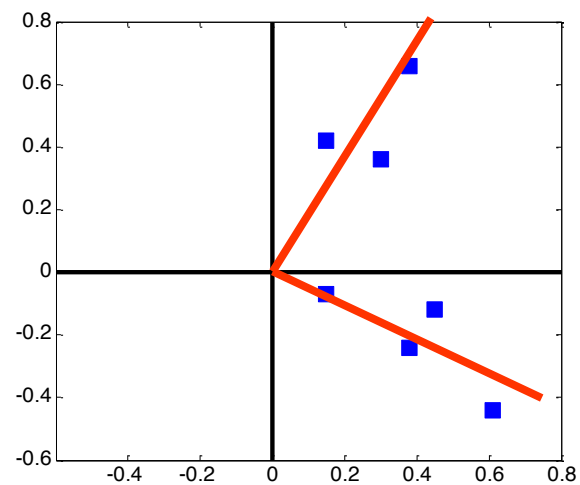
- PCs from PCA constitute only one of ***infinitely many*** possible basis vector sets for your data!
  - Can be ***rotated*** to a more interpretable basis!
- However, must specify a ***target*** for this rotation!
- Thurstone (1947), later authors: "Rotate PCs to attain ***simple structure***"
  -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

$$\mathbf{P} = \begin{bmatrix} .15 & .42 \\ .30 & .36 \\ .38 & .66 \\ .15 & -.07 \\ .38 & -.24 \\ .45 & -.12 \\ .61 & -.44 \end{bmatrix}$$

$$\mathbf{\Theta} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} .88 & .48 \\ -.48 & .88 \end{bmatrix}$$

$$(\Theta = -22.7^{o})$$

$$\mathbf{P_{rot}} = \mathbf{P\Theta} = \begin{bmatrix} -.07 & .44 \\ .09 & .46 \\ .02 & .76 \\ .16 & .01 \\ .45 & -.03 \\ .45 & .11 \\ .75 & -.09 \end{bmatrix}$$



EIGENVECTOR RESEARCH INCORPORATED

7

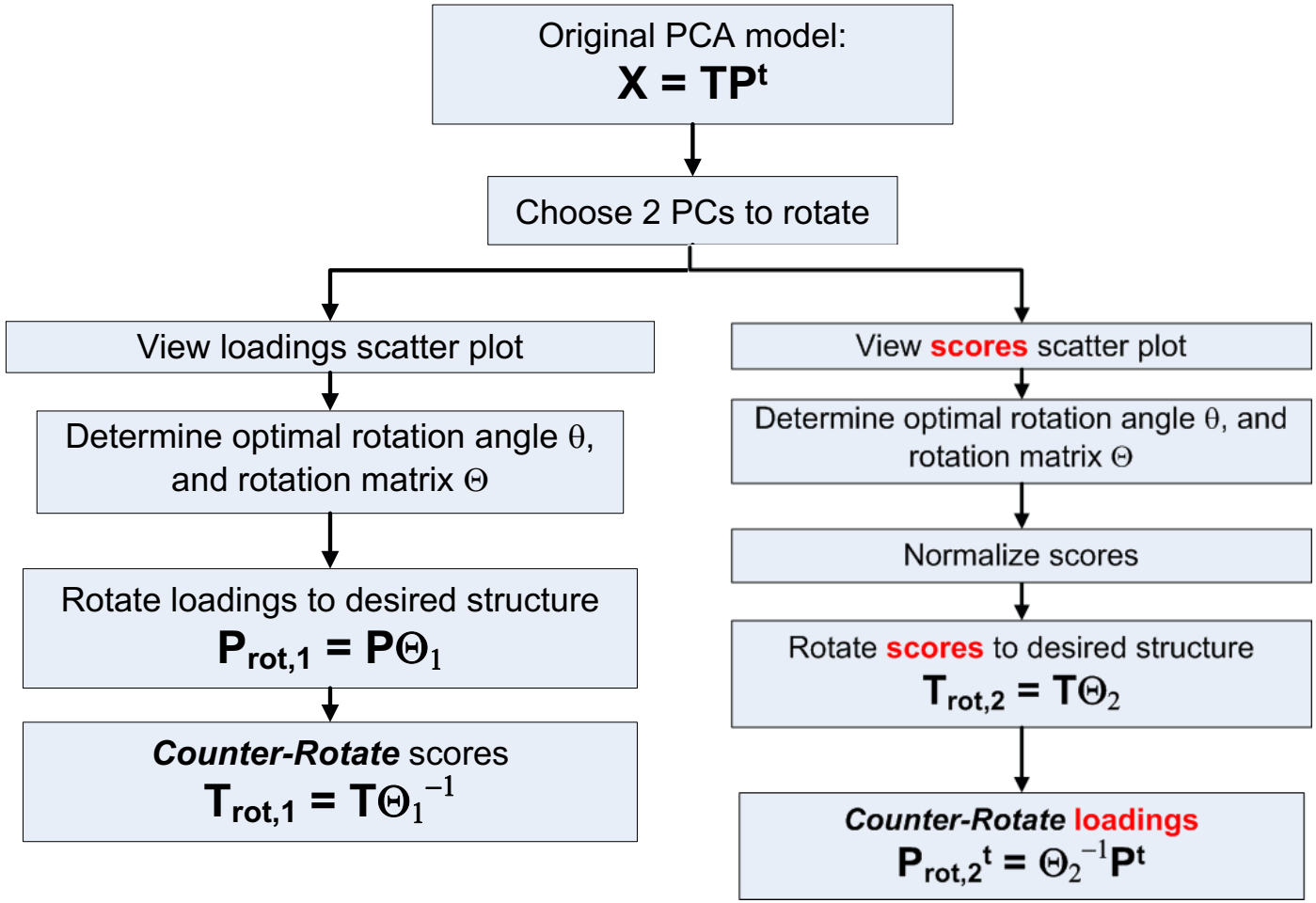
Rotation- Options

- **Typically- two vectors (PCs) at a time**
- ***Orthogonal* Rotation** $\mathbf{\Theta} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$
  - 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

8

Graphical Rotation with PCA- Math



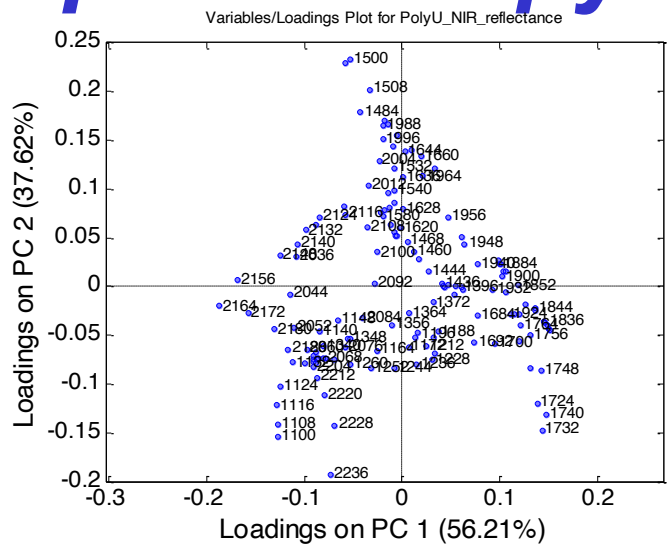
Rotation is NOT just "forcing" what you want: ***must counter-rotate*** *the other vector set in order to explain the original data!*

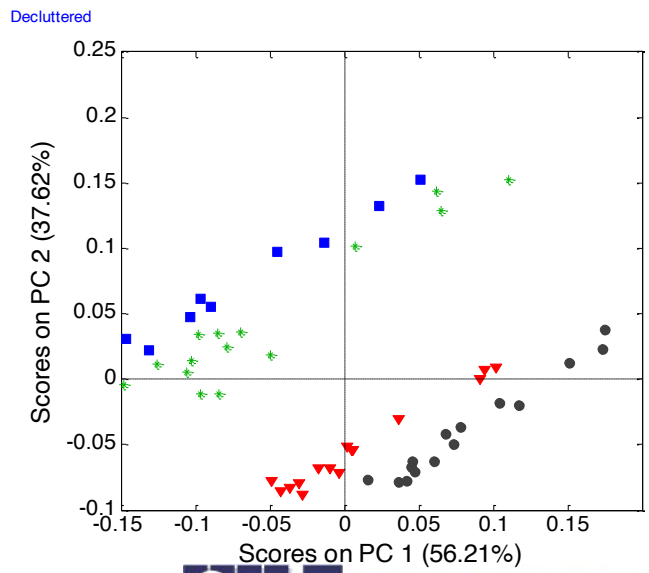
EIGENVECTOR RESEARCH INCORPORATED

9

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***?....
  - Can be done ***WITH sufficient prior knowledge of objects!***



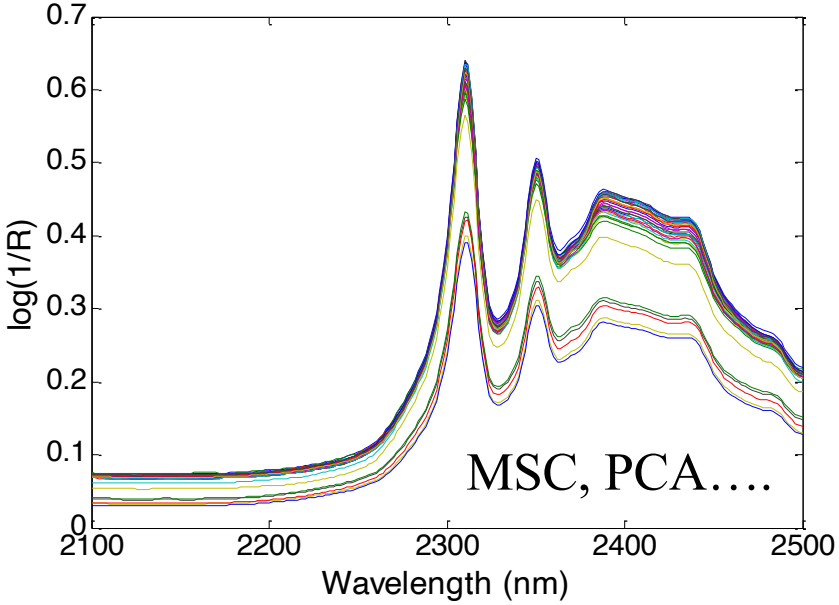


EIGENVECTOR RESEARCH INCORPORATED

10

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)
  - Replicates taken from different parts of extruded film

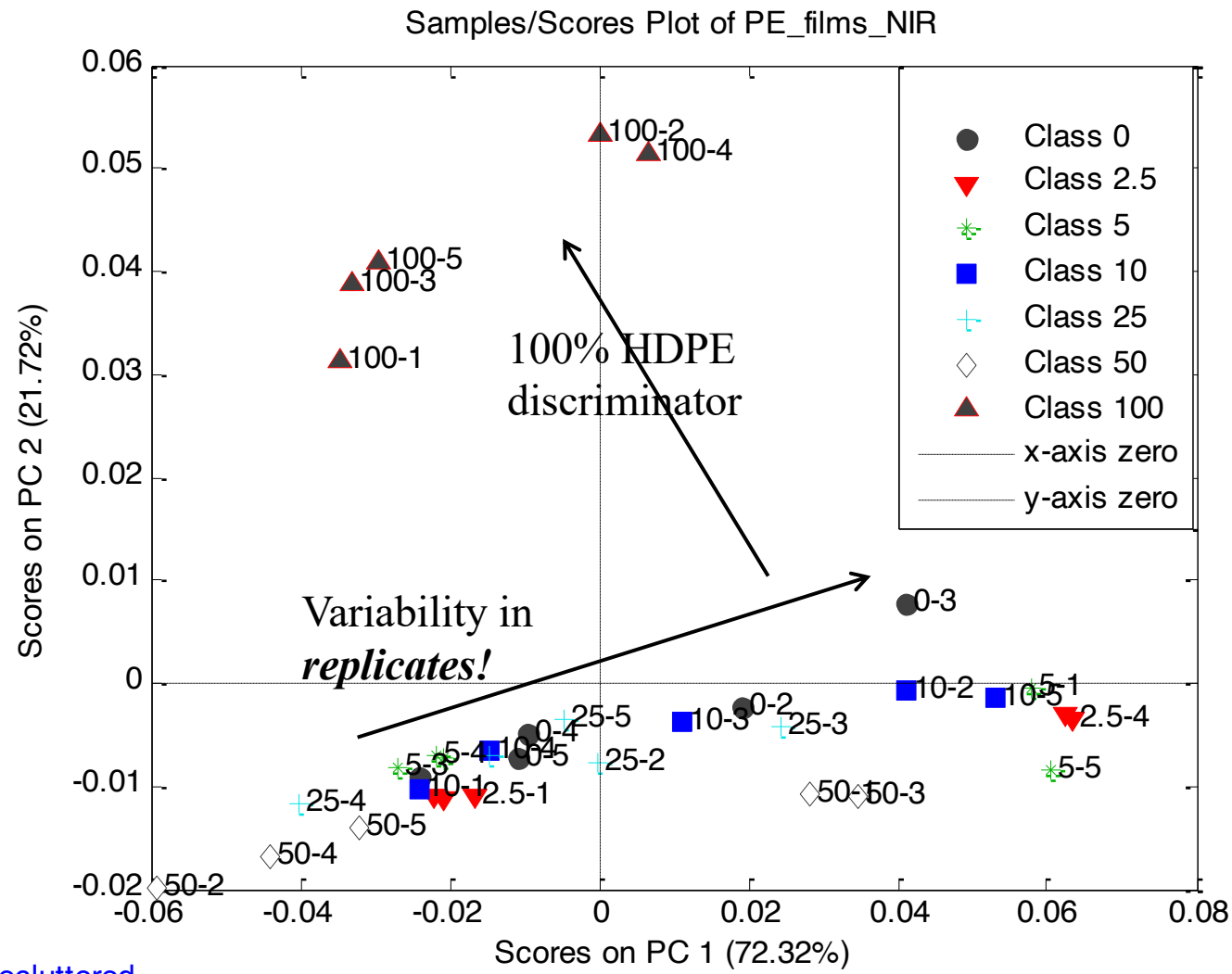


Ref: C. Miller, Appl. Spectrosc., 47(2), 222 (1993).

EIGENVECTOR RESEARCH INCORPORATED

11

Polyethylene- PC 1 & 2 scores



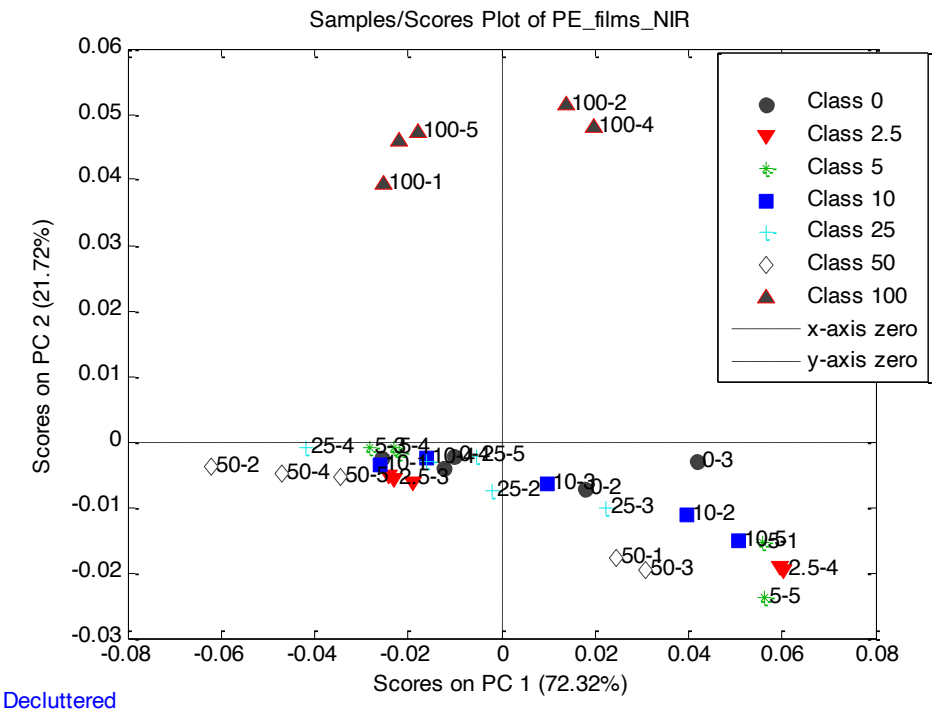
Decluttered

EIGENVECTOR RESEARCH INCORPORATED

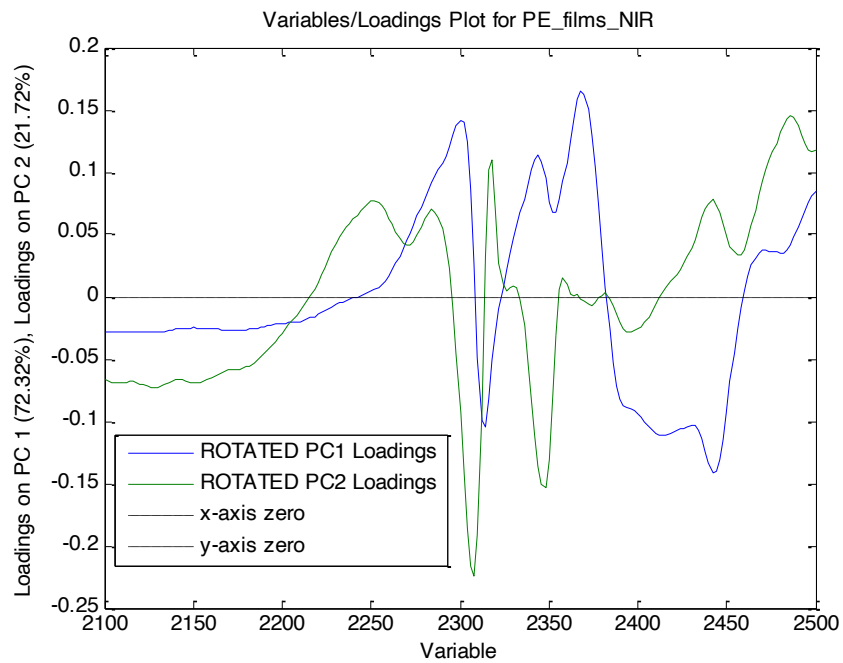
12

Rotate PCs 1&2 +15°

Resultant (counter-rotated) loadings!



Decluttered

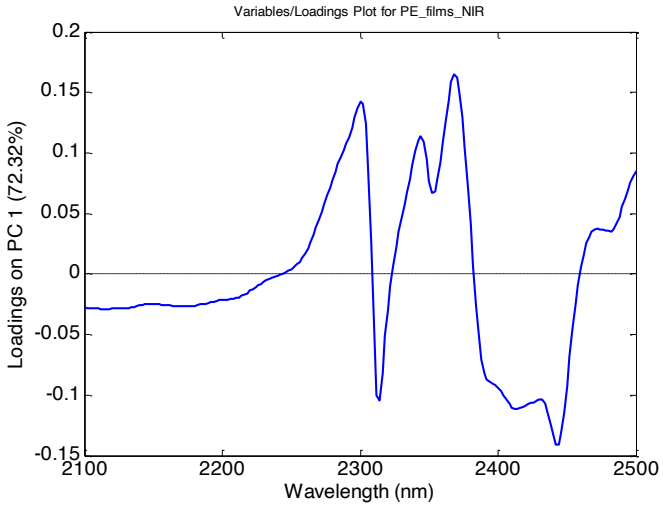


EIGENVECTOR RESEARCH INCORPORATED

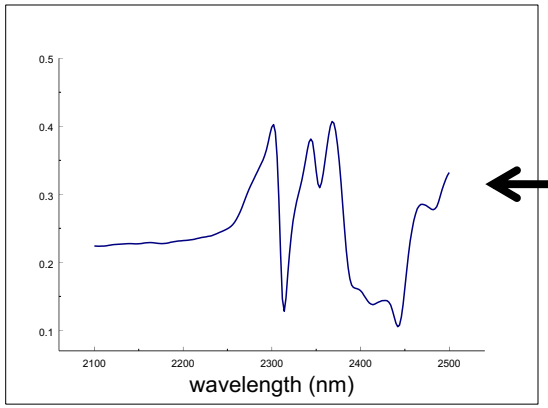
13

Effects, from counter-rotated loadings

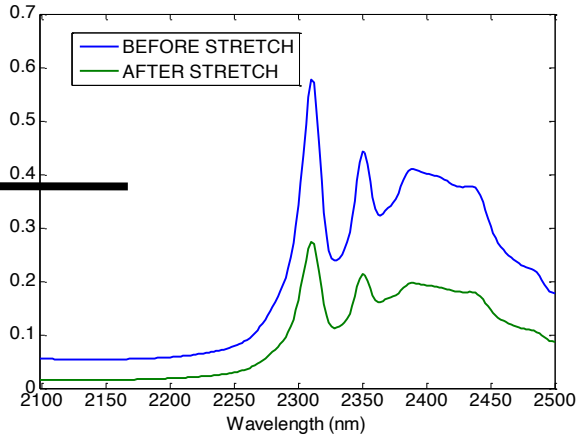
Counter-rotated loading 1 (replicate effect)



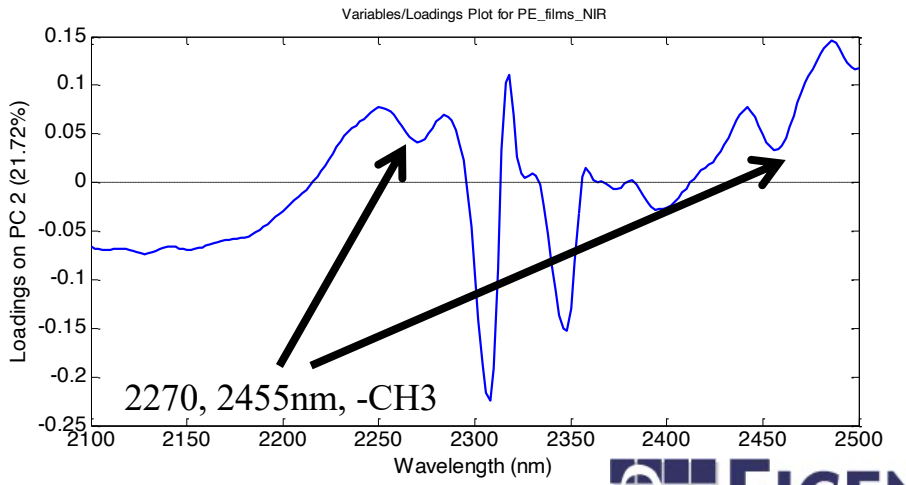
Weighted subtraction (stretching effect)



PE film before and after uni-axial stretching



Counter-rotated loading 2 (100% HDPE discriminator)

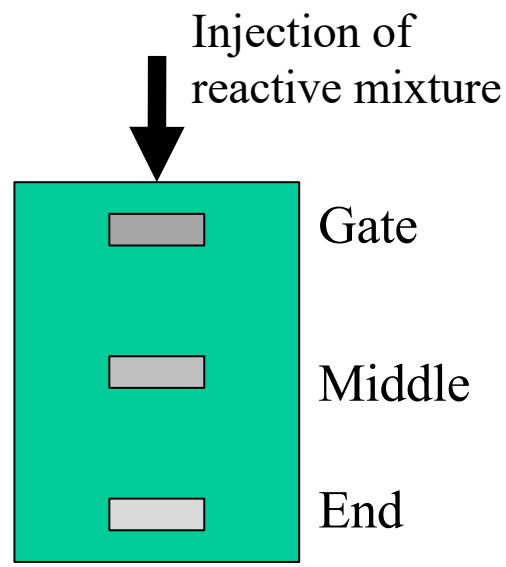


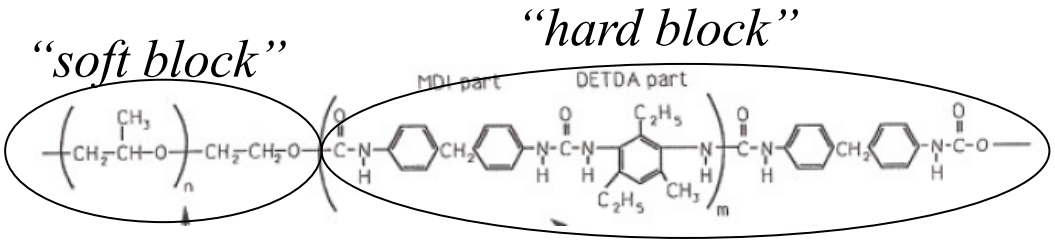
EIGENVECTOR RESEARCH INCORPORATED

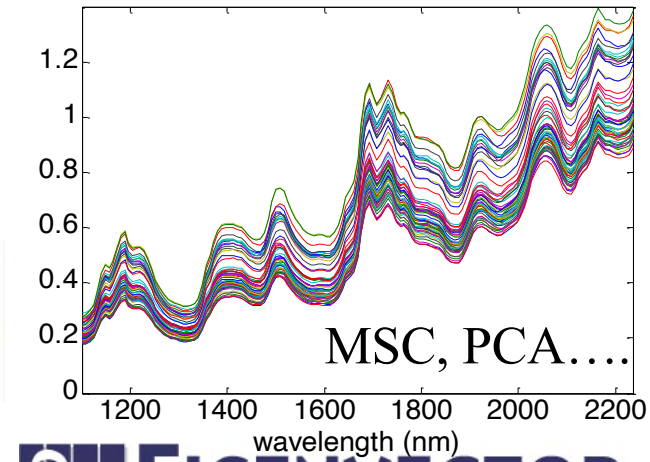
14

Case Study 2- Polyurethanes

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







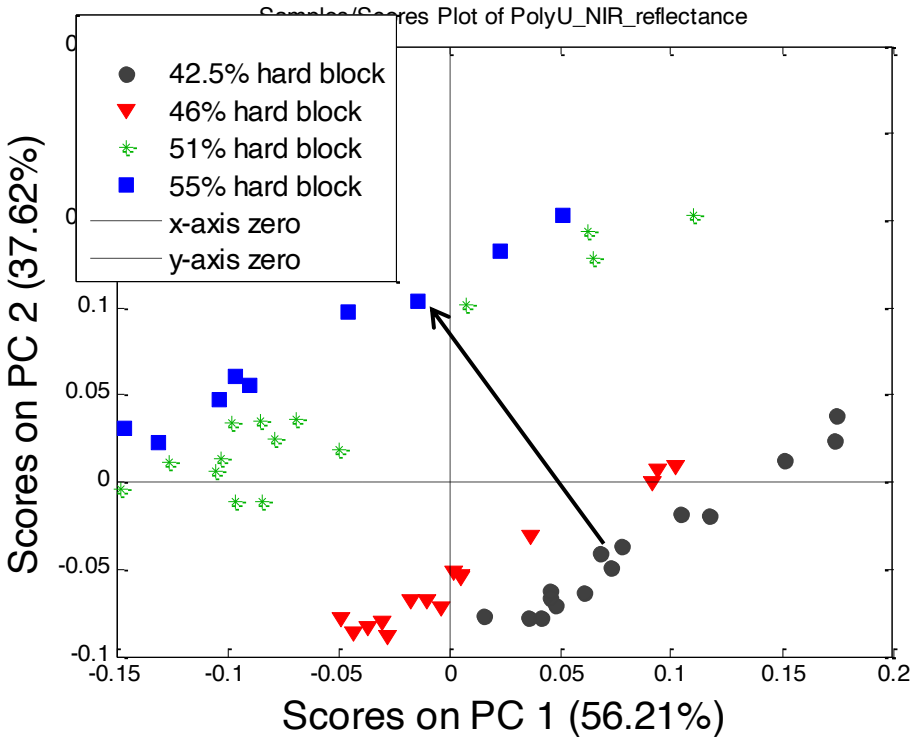
Ref: C. Miller, B. Eichinger, J. Appl. Poly Sci., 42, 2169 (1991).

EIGENVECTOR RESEARCH INCORPORATED

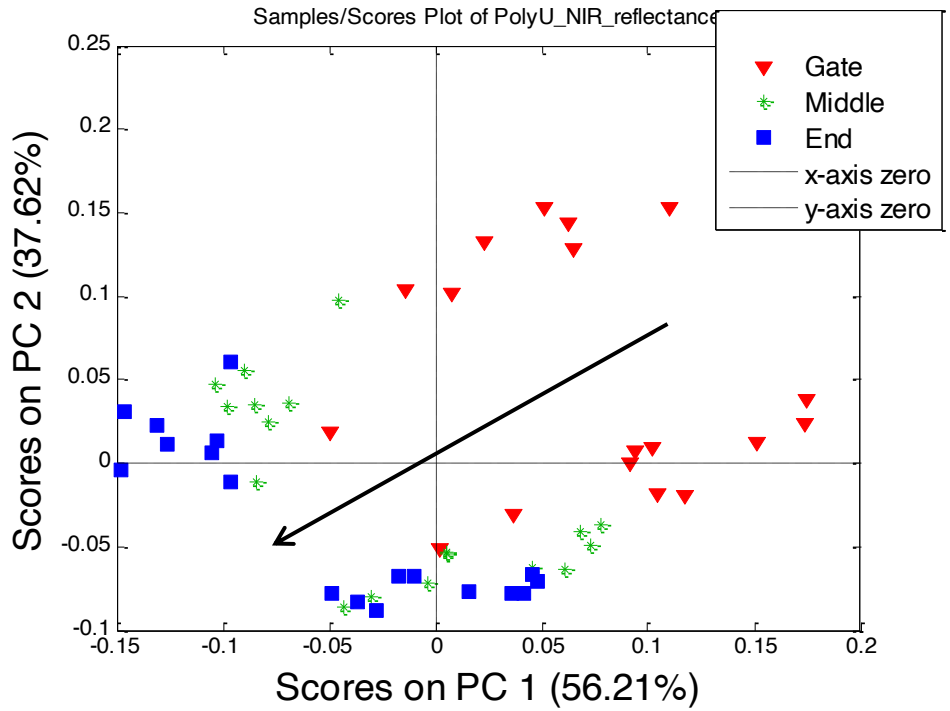
15

PCA scores 1 & 2

By hard block %



By sheet position

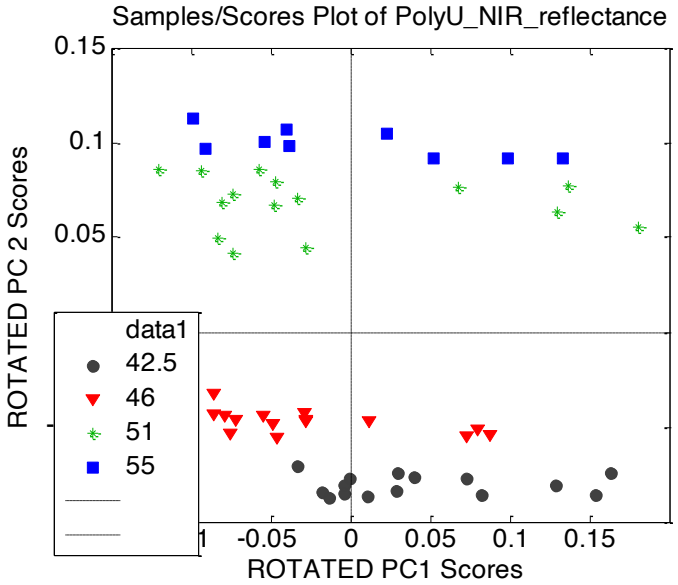


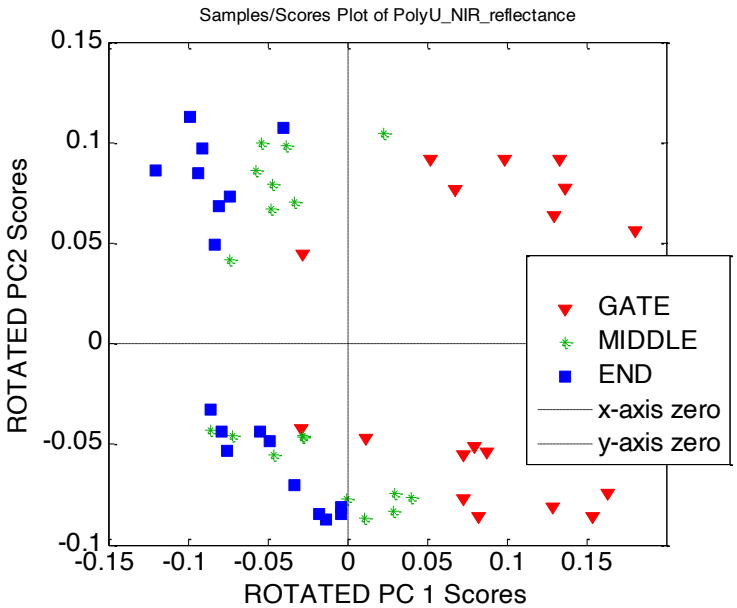
Interestingly, these two pure effects appear to be orthogonal (in this 2 PC-space)

EIGENVECTOR RESEARCH INCORPORATED

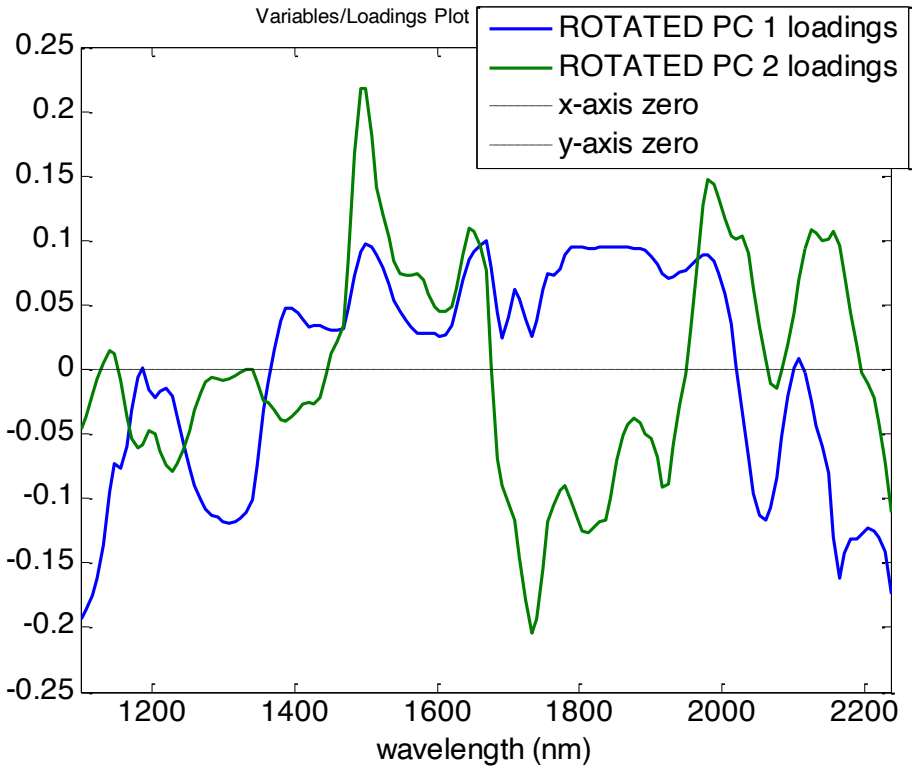
16

Rotate PCs 1&2 +37°





Resultant (counter-rotated) loadings!

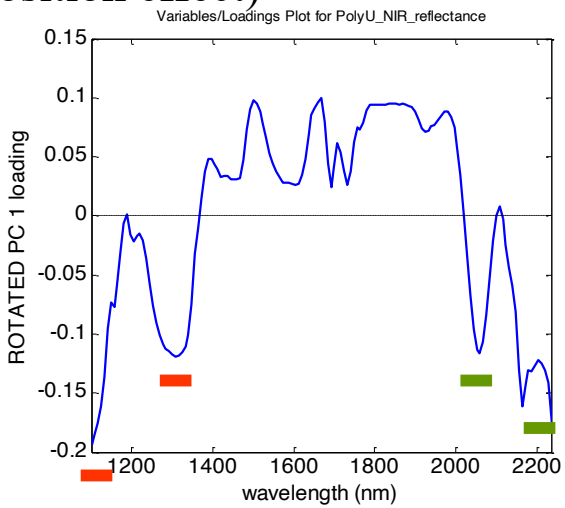


EIGENVECTOR RESEARCH INCORPORATED

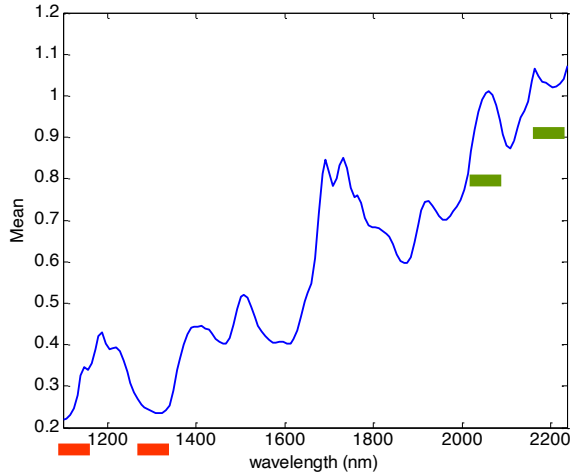
17

Effects, from counter-rotated loadings

Counter-rotated loading 1 (sheet position effect)

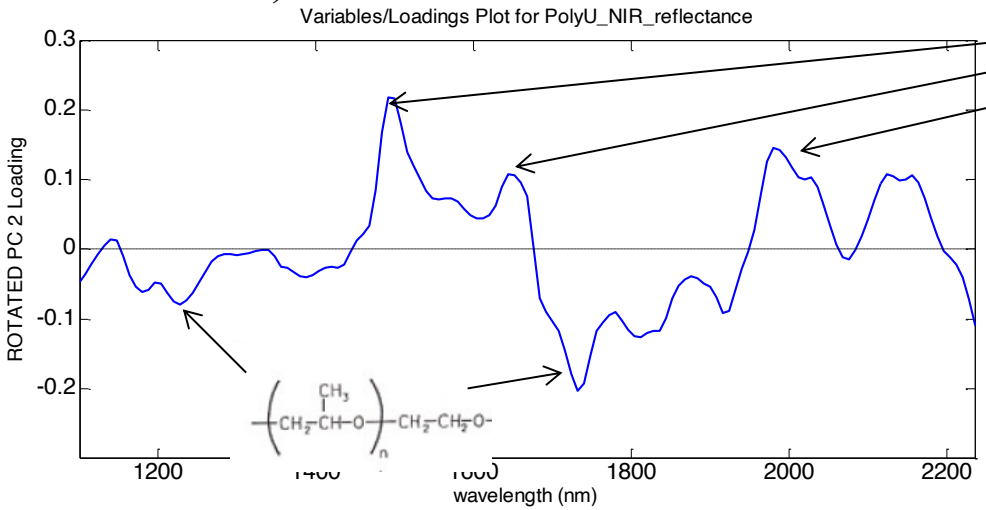


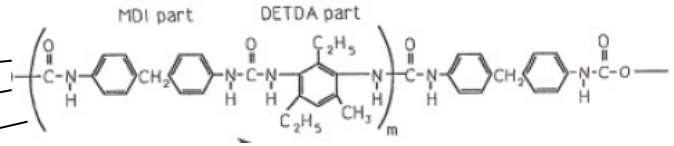
MEAN SPECTRUM



Effect: as score on PC 1 increases; valleys deeper at shorter wavelengths, and peaks shorter at longer wavelengths (remember- MSC!)

Counter-rotated loading 2 (hard block % effect)



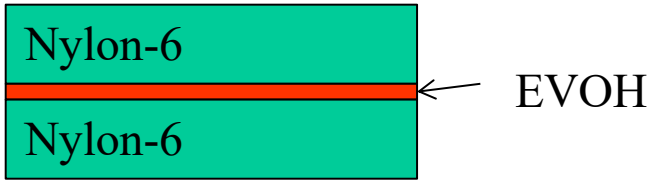


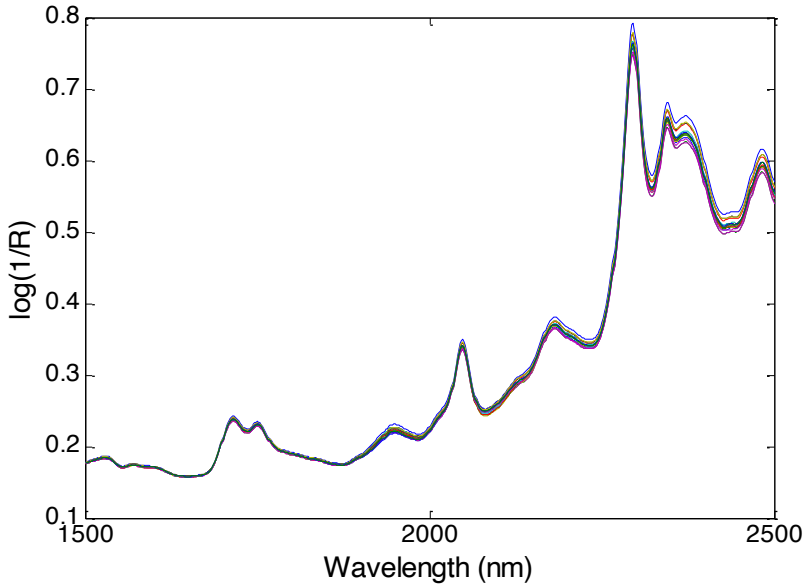
EIGENVECTOR RESEARCH INCORPORATED

18

Case study 3- 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
- NIR spectra obtained
- Subsequently analyzed by polarized light microscopy
  - EVOH content varies from 3-8%



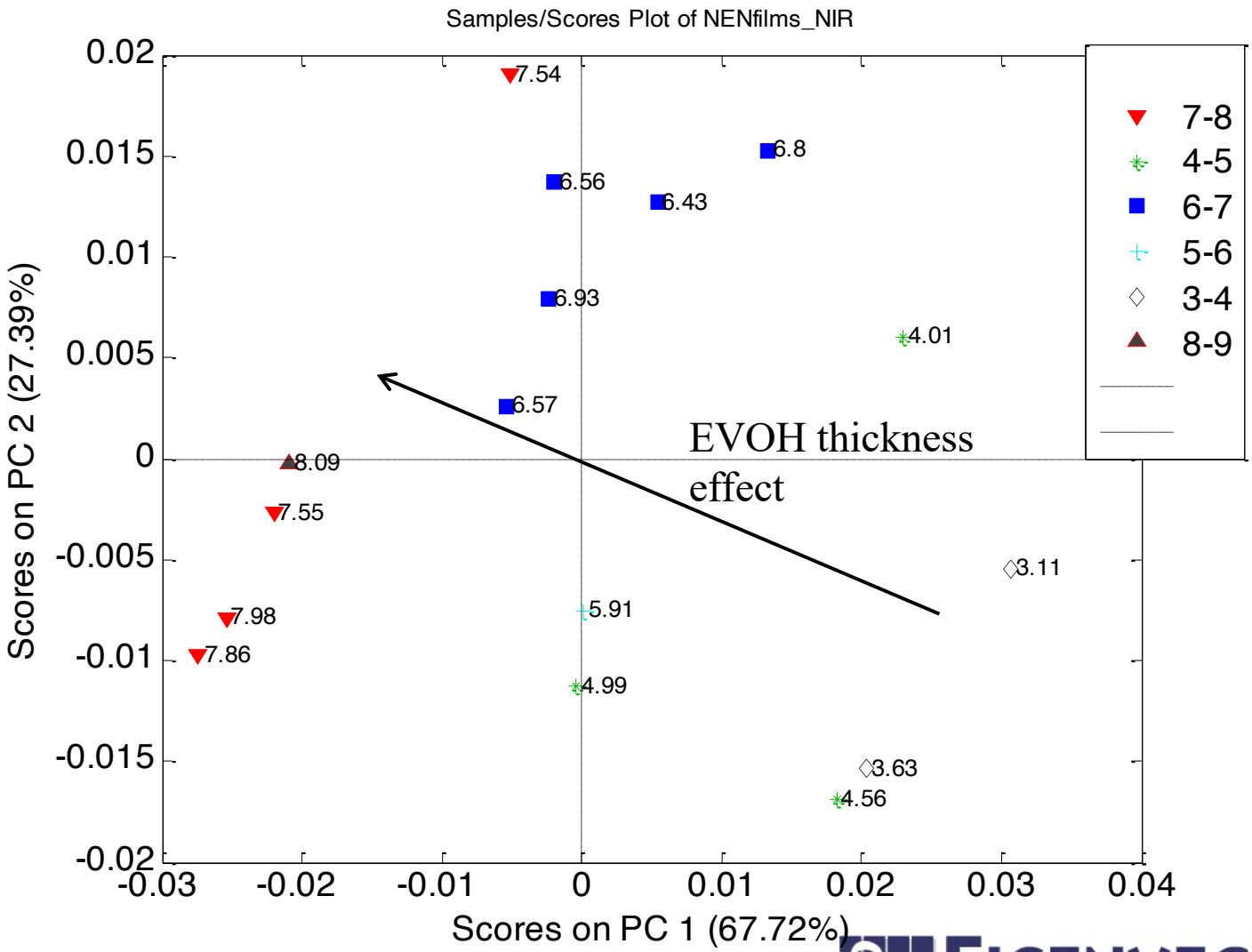


Ref: C. Miller S. Svendsen, T. Næs, Appl. Spectrosc., 47(3), 346 (1993).

EIGENVECTOR RESEARCH INCORPORATED

20

PCA scores 1 & 2

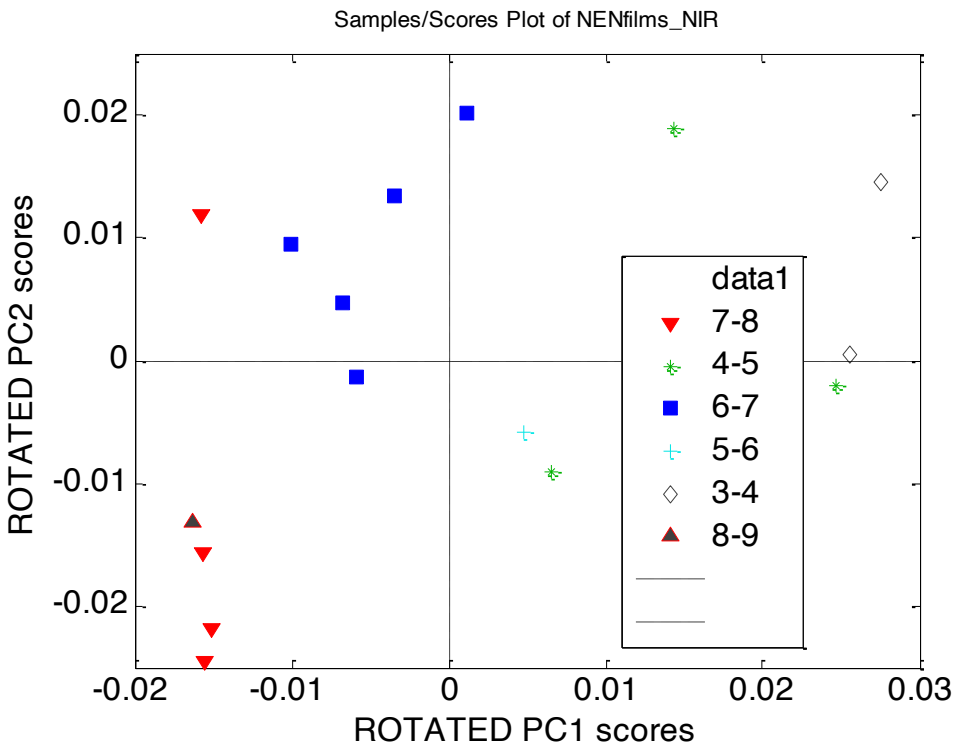


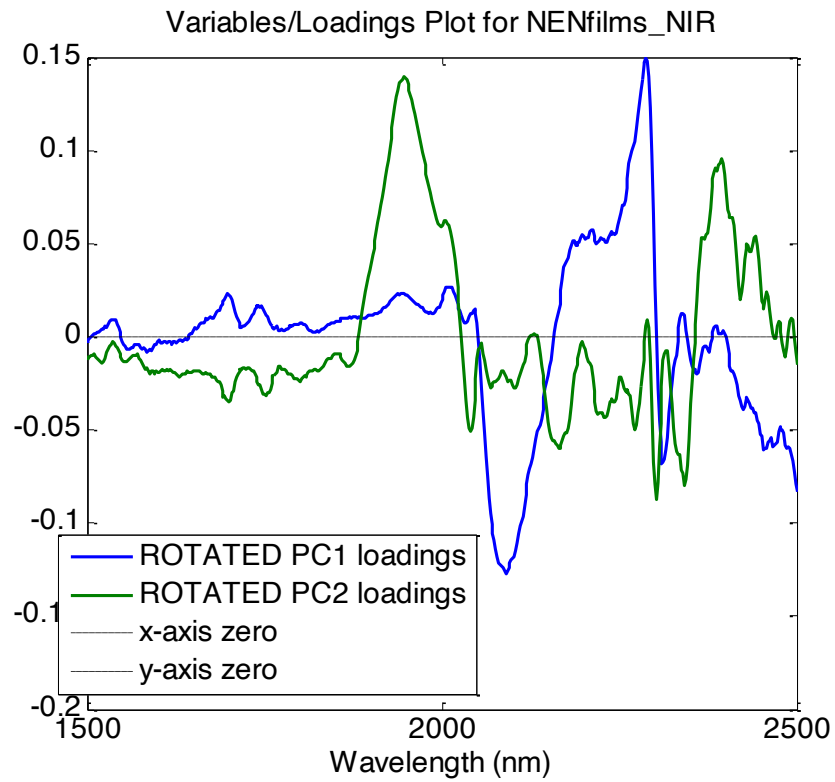
EIGENVECTOR RESEARCH INCORPORATED

21

Rotate PCs 1&2 -38º

Resultant (counter-rotated) loadings!



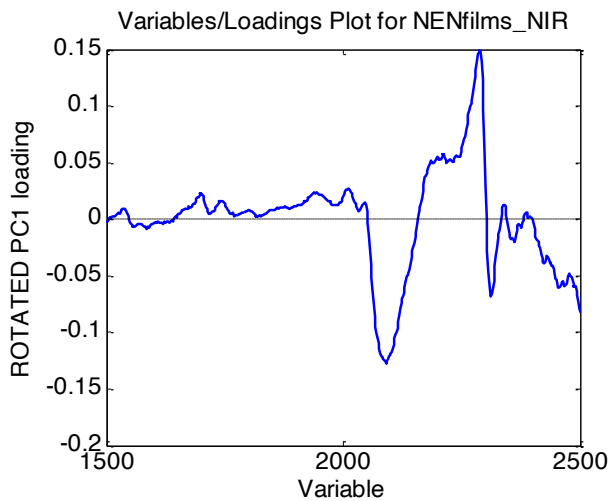


EIGENVECTOR RESEARCH INCORPORATED

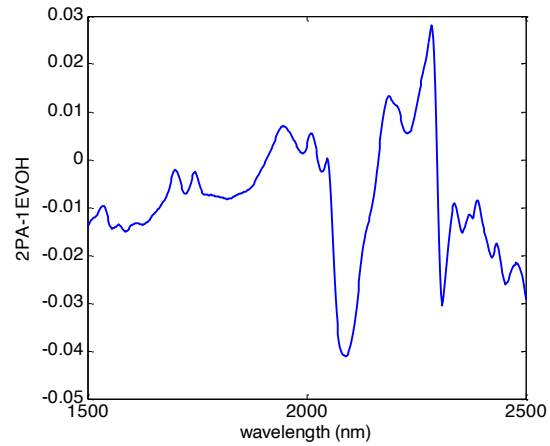
22

Effects, from counter-rotated loadings

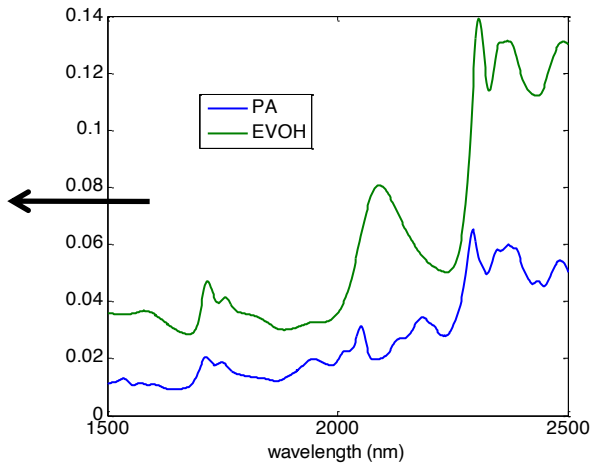
Counter-rotated loading 1 (***inverse*** of EVOH thickness effect)



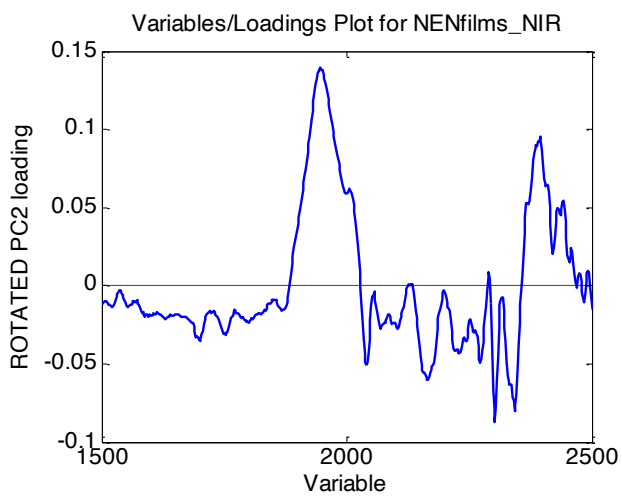
[Nylon6 – EVOH] (composition effect)



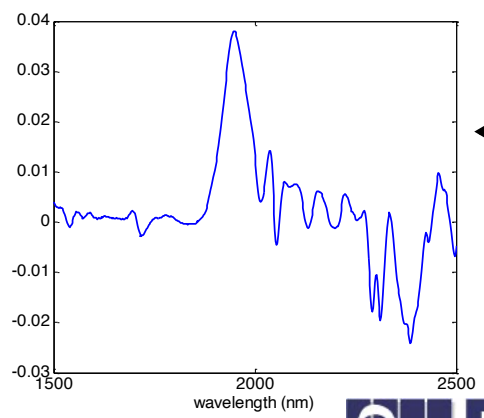
Nylon6 and EVOH pure component spectra



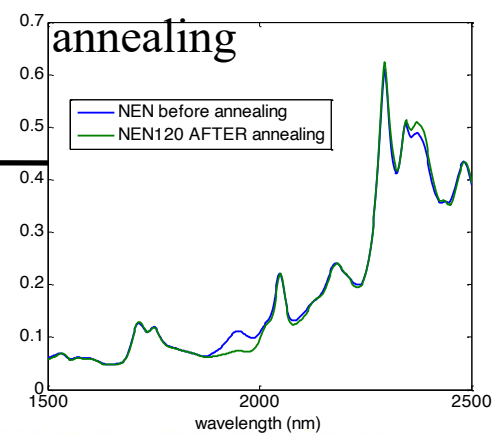
Counter-rotated loading 2 (unknown variation)



Annealing effect spectrum



Film before and after annealing



EIGENVECTOR RESEARCH INCORPORATED

23

Summary

- Interpreting loadings from NIR is ***not*** hopeless!
  - But must accept "effect" spectra,
  - Get background information, auxiliary data!
  - Rotate to a more interpretable structure
- Rotation: a very simple, time-tested alternative for exploratory NIR
- In NIR, though, rotation is more useful when target is based on structure in ***scores***, not loadings
- Can be extended to ***multi-block*** PCA, PLS
- Rotation concept is a good framework for development of "purity" methods (MCR….)

EIGENVECTOR RESEARCH INCORPORATED

24

TEST

PLS Latent Variables can be considered a

_____

_____

of the PCA Principal Components generated from the same data set.

ORTHOGONAL ROTATION!

25