# NIR Spectroscopy and Microspectroscopy Analysis of Intact Soybean Seeds for Food Applications through Composition Improvements


Tiefeng You* and I.C. Baianu**

Department of Food Science and Human Nutrition
University of Illinois at Urbana-Champaign
*403 Bevier Hall,
** 305 Burnsides Research Laboratory,
1208 W. Pennsylvania Ave. ,URBANA, IL. 61801, USA

# ABSTRACT

**The soybean-derived products are among the most important agricultural products in the USA and the world. Conventional analytical methods for soybean composition analysis are both time consuming and costly. Faster and less expensive methods are required for most practical applications. To improve the accuracy, reliability and sensitivity of NIR, major advancements in instrumentation, as well as, data analysis / calibration methodology are required. Novel NIR instruments, such as DA-NIR and FT-NIR spectrometers developed in recent years have the potential for improving significantly the quantification of soybean composition at a reasonable cost. We present representative calibrations and data for intact soybean composition analysis obtained at the University of Illinois at Urbana with cutting-edge NIR instrumentation that is currently commercially available at reasonable cost.**

# Presentation Outline

- Research objectives, rationale, and significance
- Background and principles of Near Infrared Spectroscopy (NIR)
- Study of NIR calibration models and spectral preprocessing techniques
- Evaluation and selection of state-of-the-art NIR instruments
- Development and validation of NIR calibrations for soybean analysis on the DA7000 and Spectrum One NTS instruments
- Microspectroscopy and chemical imaging of soybean seeds
- Conclusions

# Research Objectives

- Investigate and develop methodologies for NIR spectra pre-processing and data analysis in order to improve the accuracy and reliability of NIR measurements of soybean seed composition.

- Develop and improve calibrations of state-of-the-art NIR instruments for measuring rapidly, accurately, and reproducibly major soybean components that are important for food applications, such as: protein, oil, moisture, and isoflavones.

# Rationale and Significance

- Soybean is one of the most important agricultural products in the world.
- Conventional analytical methods are time consuming and costly. Faster and less expensive methods are required for most practical applications.
- To improve the accuracy, reliability and sensitivity of NIR, major advancements in instrumentation, as well as, data analysis / calibration methodology are required.
- Novel NIR instruments (DA-NIR, FT-NIR spectrometers) developed in recent years have the potential for improving significantly the quantitation of soybean composition at a reasonable price.

# Principles of Near Infrared Spectroscopy

- Energy levels of Molecular Vibrations and Rotations

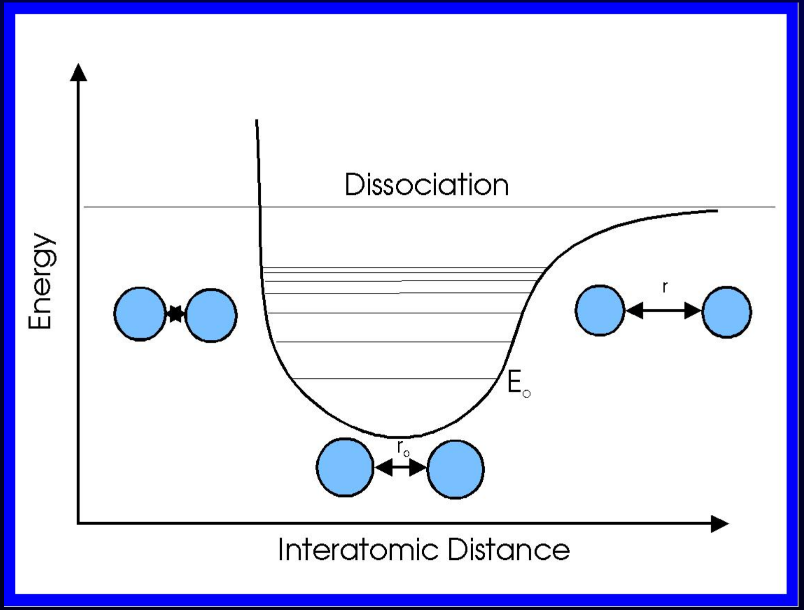
- Transitions between different energy levels generate IR absorption peaks:

  - Fundamentals: Mid-IR, Far-IR (λ >2,500 nm)
  - **Overtones**: in the **NIR** range
  - **Combinations**: in the **NIR** range

- **NIR** wavelengths/frequency ranges: 700 to 2,500 nm / (14,000 to 4,000 $cm^{-1}$)

  - **Overtone bands: 700 to 1,800 nm** (14,000 to 5,500 $cm^{-1}$)

  - **Combination bands: 1,800 to 2,500 nm** (5,500 to 4,000 $cm^{-1}$)



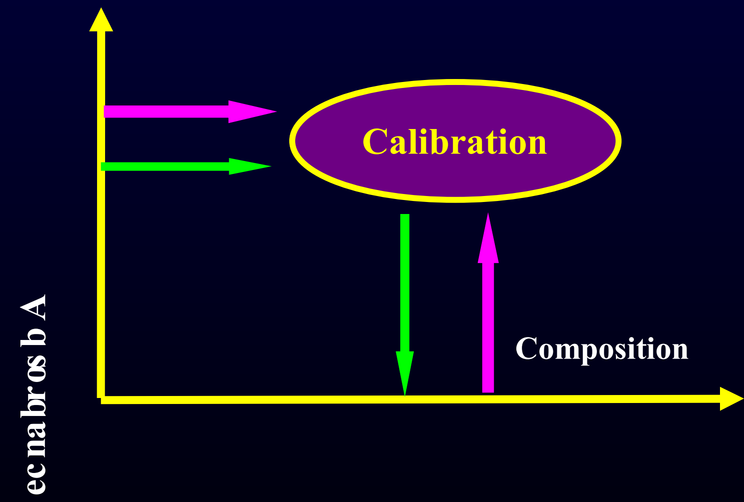
# Quantitative Analysis by NIR

- Principles of Quantitative Analysis:

Lambert-Beer's Law: $A = \varepsilon * L * C$

  - A = True absorbance of sample
  - $\varepsilon$ = Extinction coefficient of analyte
  - L = Light's pathlength in the sample
  - C = Analyte concentration

- Major steps in Quantitative Analysis:
  - Analysis of Standard Samples by primary, or reference, method (s)
  - Collection of NIR spectra of Standard Samples
  - Pre-processing of the NIR spectra of Standards
  - Calibration development with a Selected Calibration Model/Validation based on Lambert-Beer's Law
  - Predictions for unknown samples with the developed calibration



# Data Analysis for NIR Quantitation

- NIR spectra preprocessing
  - Specular reflection / baseline corrections
  - Light scattering correction
- Calibration development
  - PLS
  - PCR
  - MLR
  - CLS

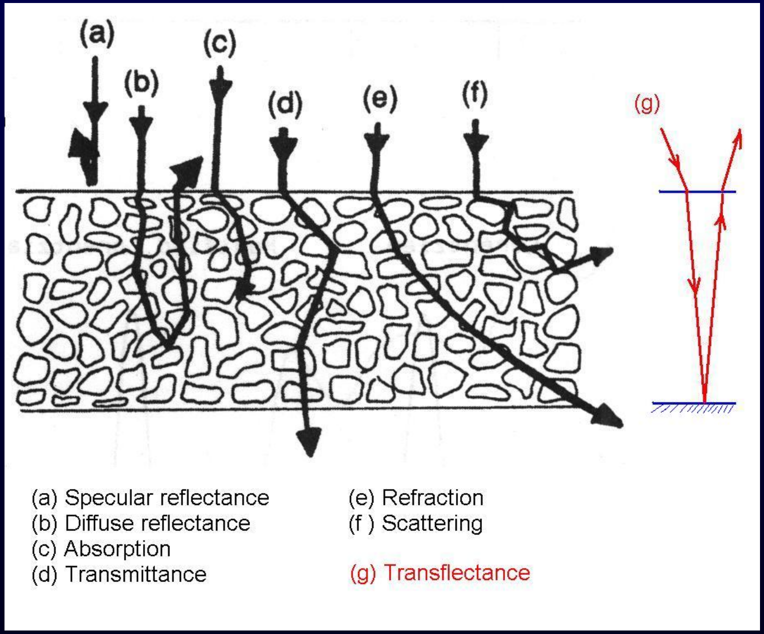
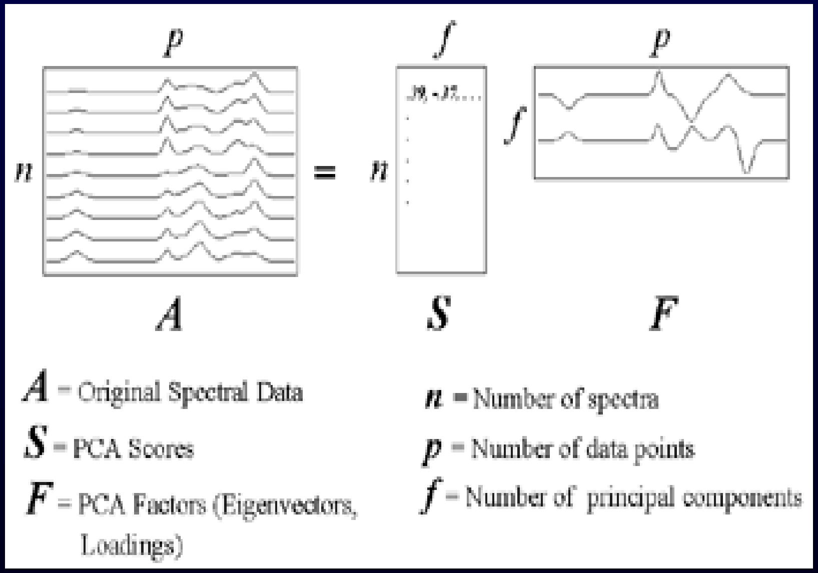


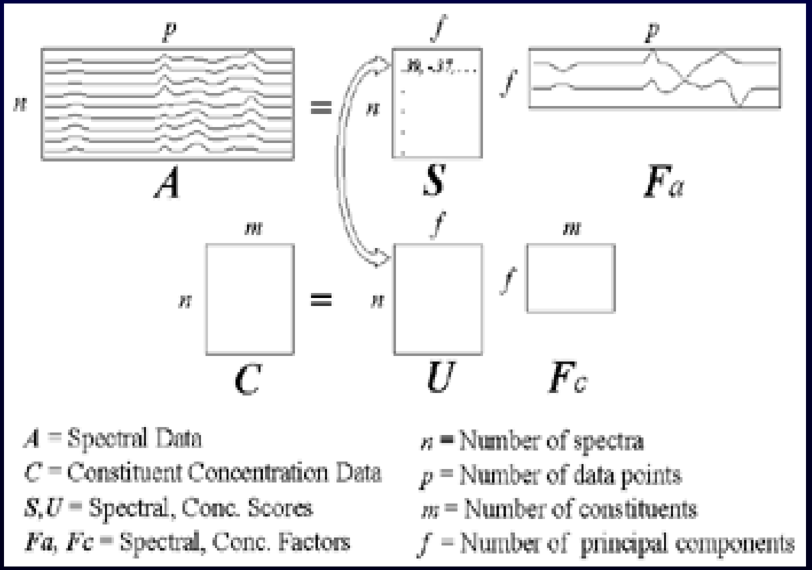

*Diagram based on* Burns, D.A. 2001. Handbook of Near-Infrared Analysis. Marcel Dekker

# Illustration of PCR and PLS

PCR




PLS



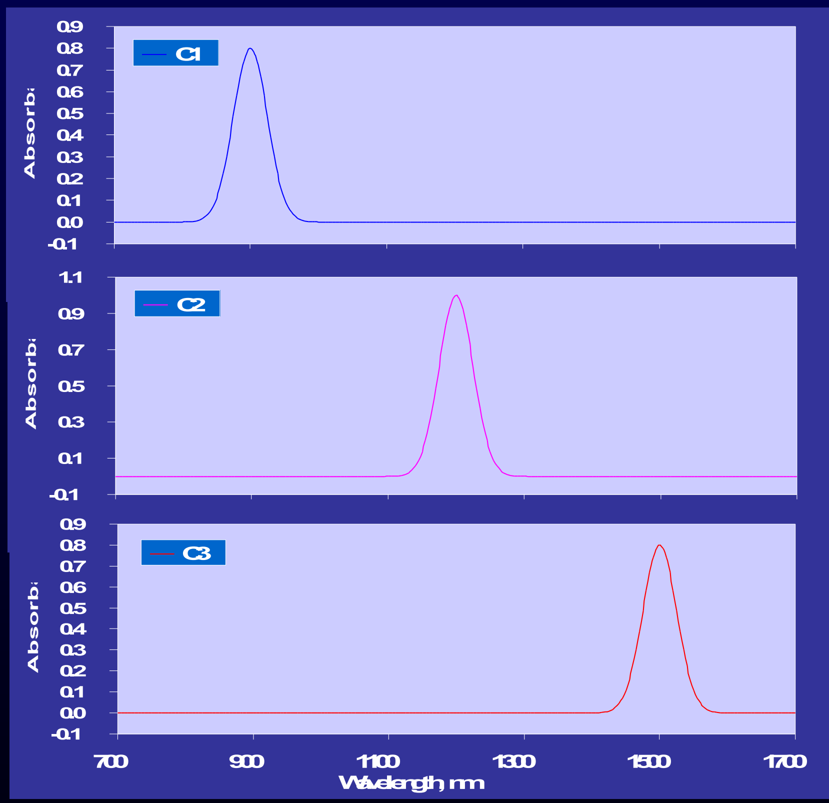

*From* Galactic, 1996. PLSplus/IQ for Grams/32 and Grams/386

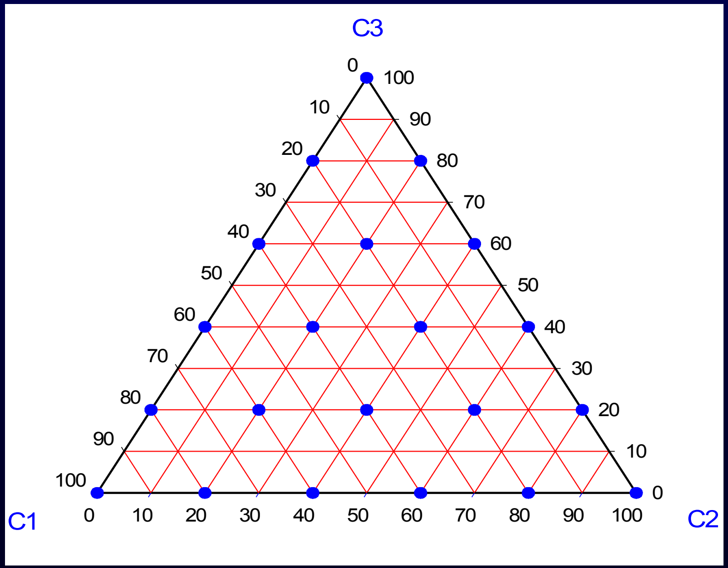
# Section A

## Study of NIR calibration models and spectral pre-processing techniques

- Computer Simulation Study of the PLS-1 Calibration Algorithm
- Light Scattering Corrections of NIR Spectra of Soybeans

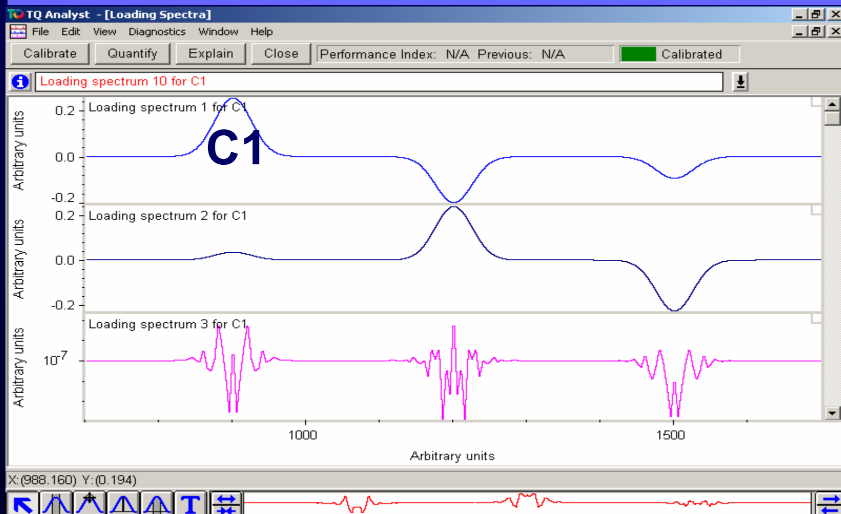
# Computer Simulation Study of the PLS-1 Calibration Algorithm, Using Synthetic Spectra of three Components

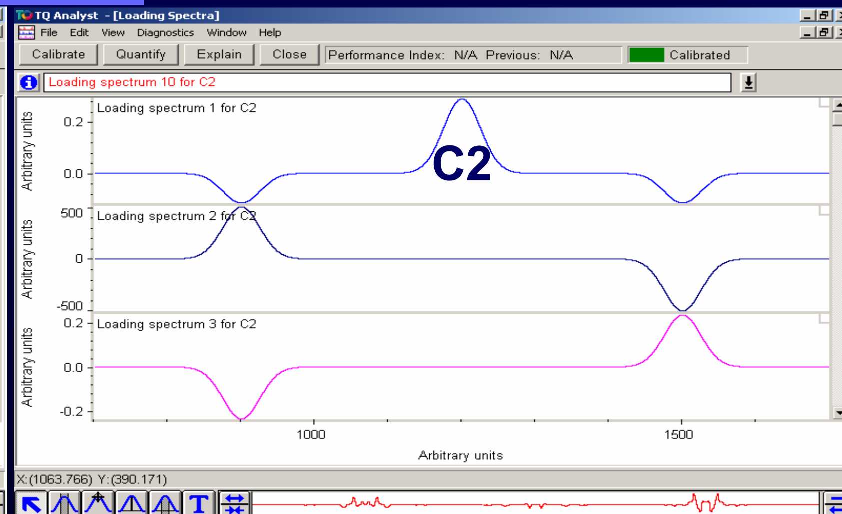


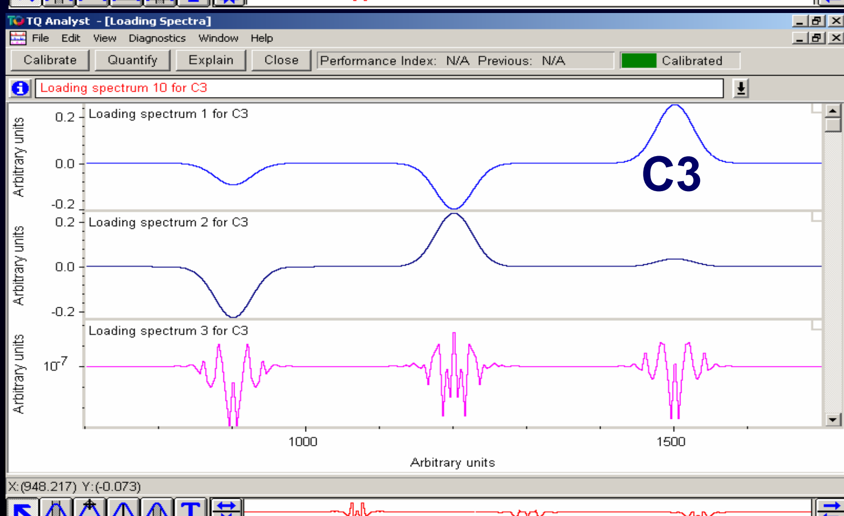




N=21, Calibration Standards

# Loading Vectors and SECV's of Components C1 to C3 for the Simulated PLS-1 Calibration



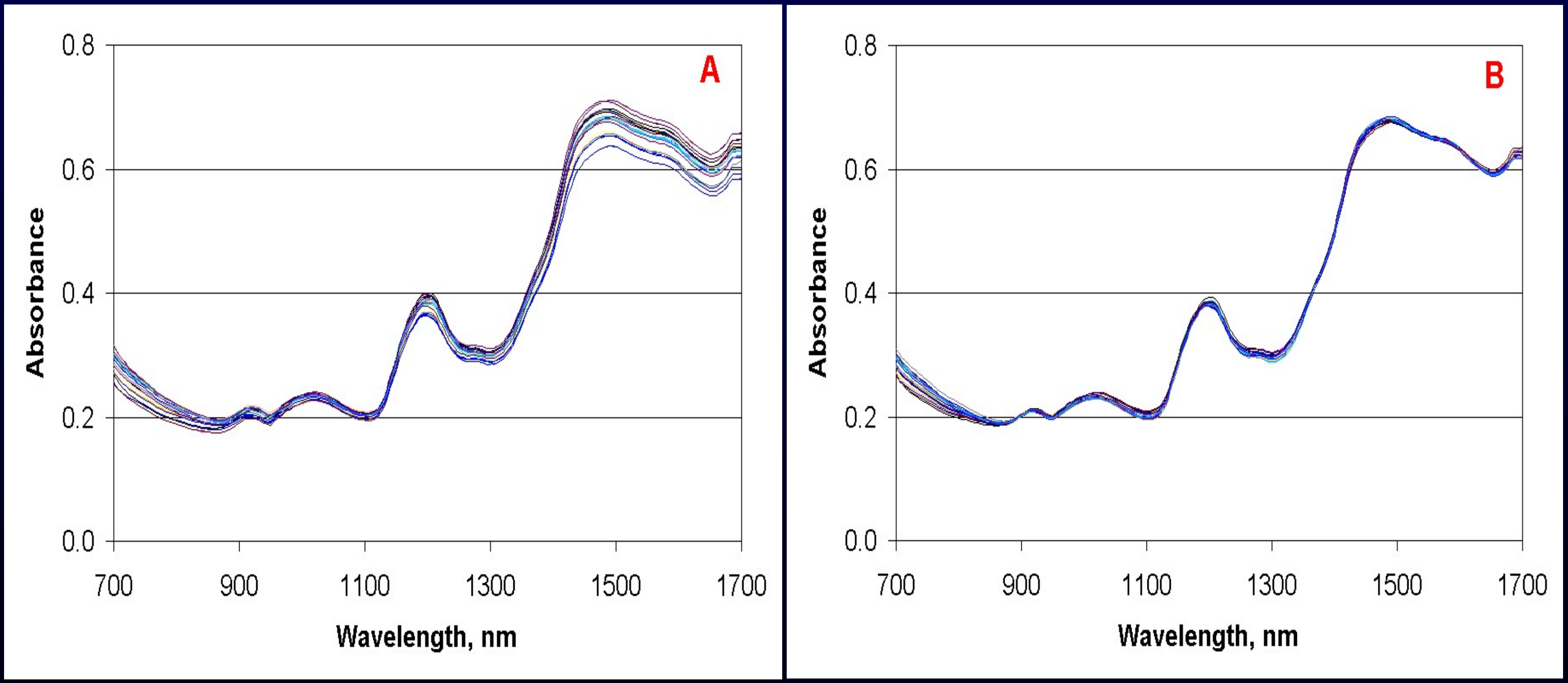

| Components | Number of Factors | R | SECV |
|---|---|---|---|
| C1 | 2 | 100% | $9 \times 10^{-6}$ |
| C2 | 2 | 100% | $1 \times 10^{-6}$ |
| C3 | 2 | 100% | $2 \times 10^{-6}$ |

R: Correlation coefficient
SECV: Standard Error of Cross Validation

# Light Scattering Corrections of NIR Spectra of Soybeans

- Spectra variations between soybean samples can be caused by chemical composition differences, specular reflection, and light scattering effects.
- Effects of light scattering and specular reflection on NIR spectra of soybean needs to be corrected.
- Current methods for spectra pre-processing / corrections of light scattering and specular reflection effects are in need of further improvements

# DA-NIR Spectra of Bulk Soybean Samples, before (A) and after (B) MSC (700-1700nm)



# Effect of Light Scattering on NIR Analysis Accuracy (DA7000 Calibration)

| Procedures | Number of Factors | SECV |
|---|---|---|
| None | 19 | 0.58 |
| MSC-VIS/NIR | 14 | 0.76 |
| MSC-NIR | 9 | 0.56 |
| MSC-NIR+SG1 | 6 | 0.52 |

SECV: Standard Error of Cross Validation

# Section B

Evaluation and Selection of State-of-the-Art NIR Instruments:

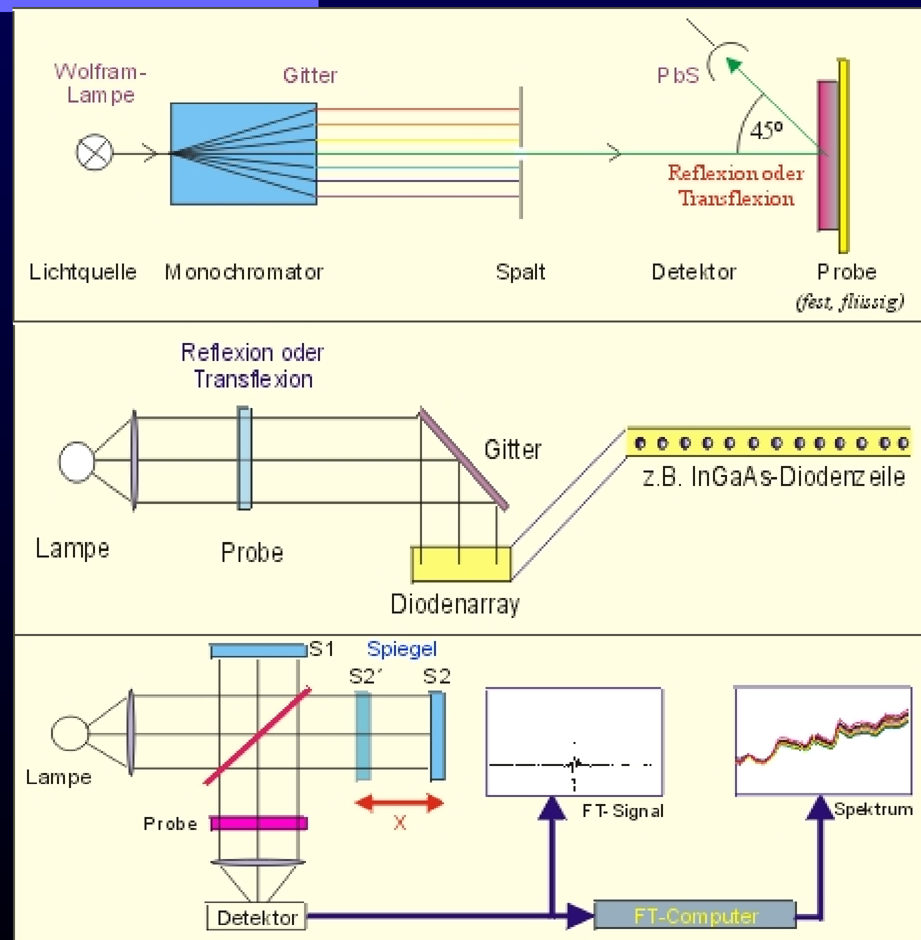
- Diode Array (DA-NIR)
  - Perten’s DA7000 Dual-Diode Array Instrument
- Fourier Transform (FT-NIR)
  - Perkin-Elmer‘s Spectrum One NTS
  - Nicolet Antaris and Nexus-670
  - Bruker Vector-22 and Vector-33

# NIR Instrumentation Techniques

Dispersive

Diode Array

Fourier Transform




http://www.pirnet.bafz.de/enirintro.htm

# Fourier Transform (FT-NIR) Instruments

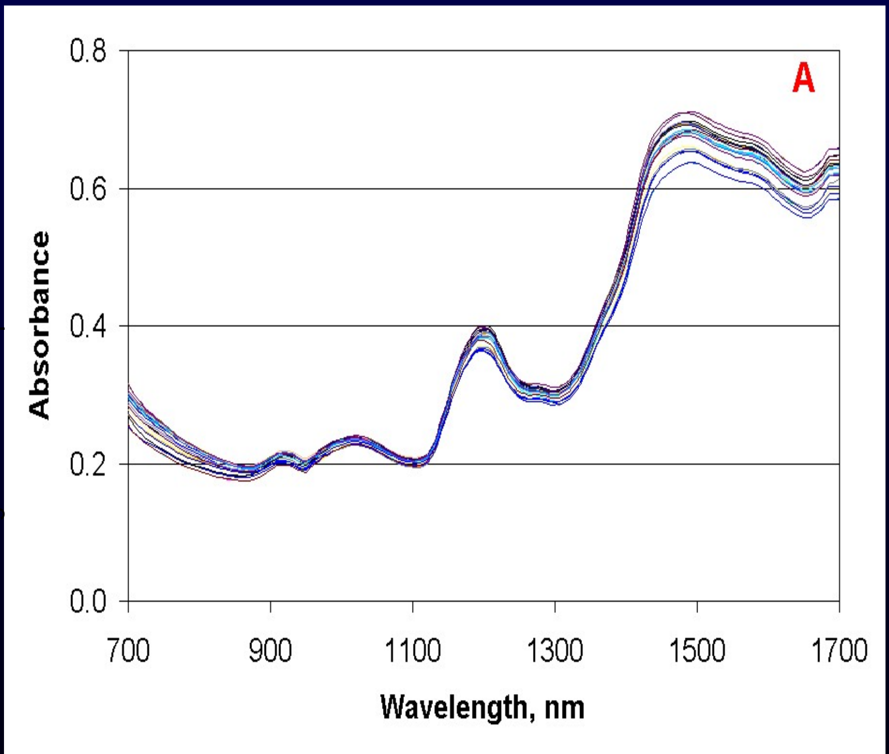
- Perkin-Elmer's Spectrum One NTS
  - Advantages: Resolution, flexibility, sensitivity towards IR range, AVI
  - Spectral Range: 12000 ~ 4000 $cm^{-1}$

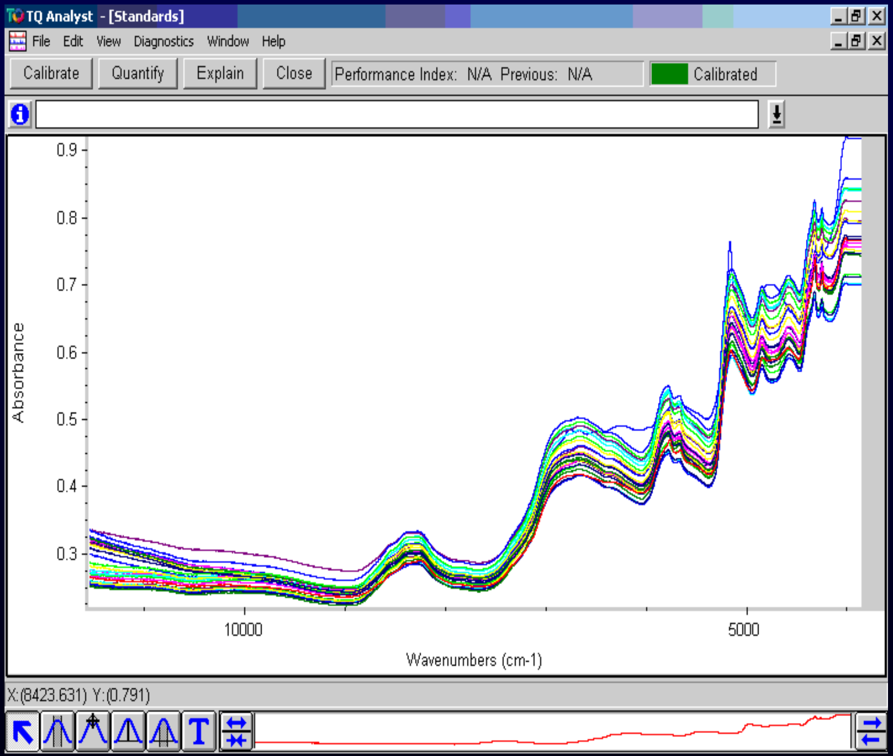
- Nicolet's Antaris FT-NIR Spectrometer
  - Detector: InGaAs
  - Spectral Range: 12000 ~ 4000 $cm^{-1}$

# Overlay Plot of NIR Spectra of Soybeans

## Left: DA-7000, Right: Spectrum One NTS

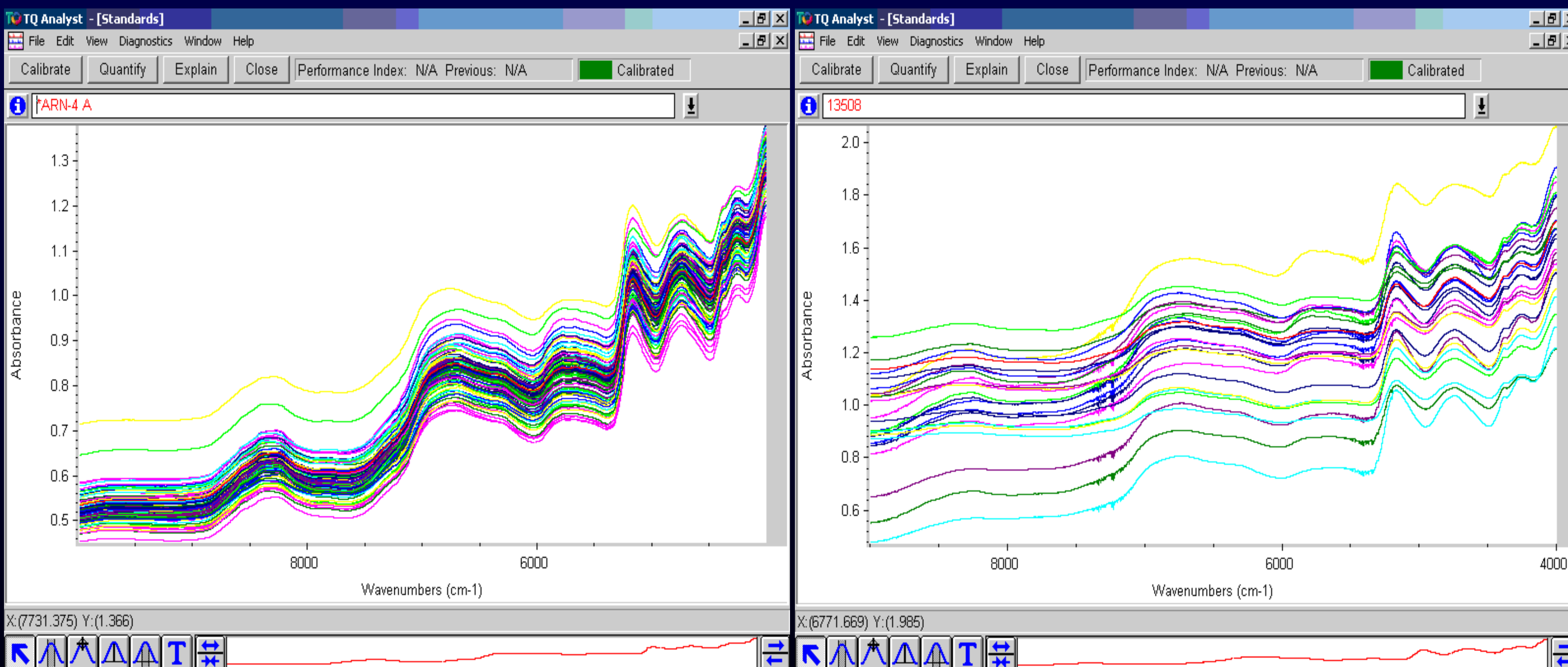




# Overlay Plot of NIR Spectra of Soybeans

## Left: Antaris, Right: Nexus-670



# Section C

Development and Validation of NIR Calibrations for Soybean Analysis on the DA7000 DA-NIR and SpectrumOne NTS FT-NIR Instrument

# Calibrations Developed on Perten's DA7000 for Bulk Soybean Analysis

| Component | Number of Factors | R | SECV |
|---|---|---|---|
| Protein | 8 | 99.9% | 0.13 |
| Oil | 7 | 99.9% | 0.07 |
| Moisture | 8 | 99.7% | 0.04 |

R: Correlation Coefficient
SECV: Standard Error of Cross Validation

# Calibrations Developed on Perkin-Elmer's Spectrum One NTS for Bulk Soybean Analysis

| Component | Number of Factors | R | SECV |
|---|---|---|---|
| Protein | 13 | 99.9% | 0.26 |
| Oil | 15 | 99.9% | 0.13 |
| Moisture | 15 | 99.9% | 0.17 |

R: Correlation Coefficient
SECV: Standard Error of Cross Validation

# Calibrations Developed on Perkin-Elmer's Spectrum One NTS for Single Soybean Analysis

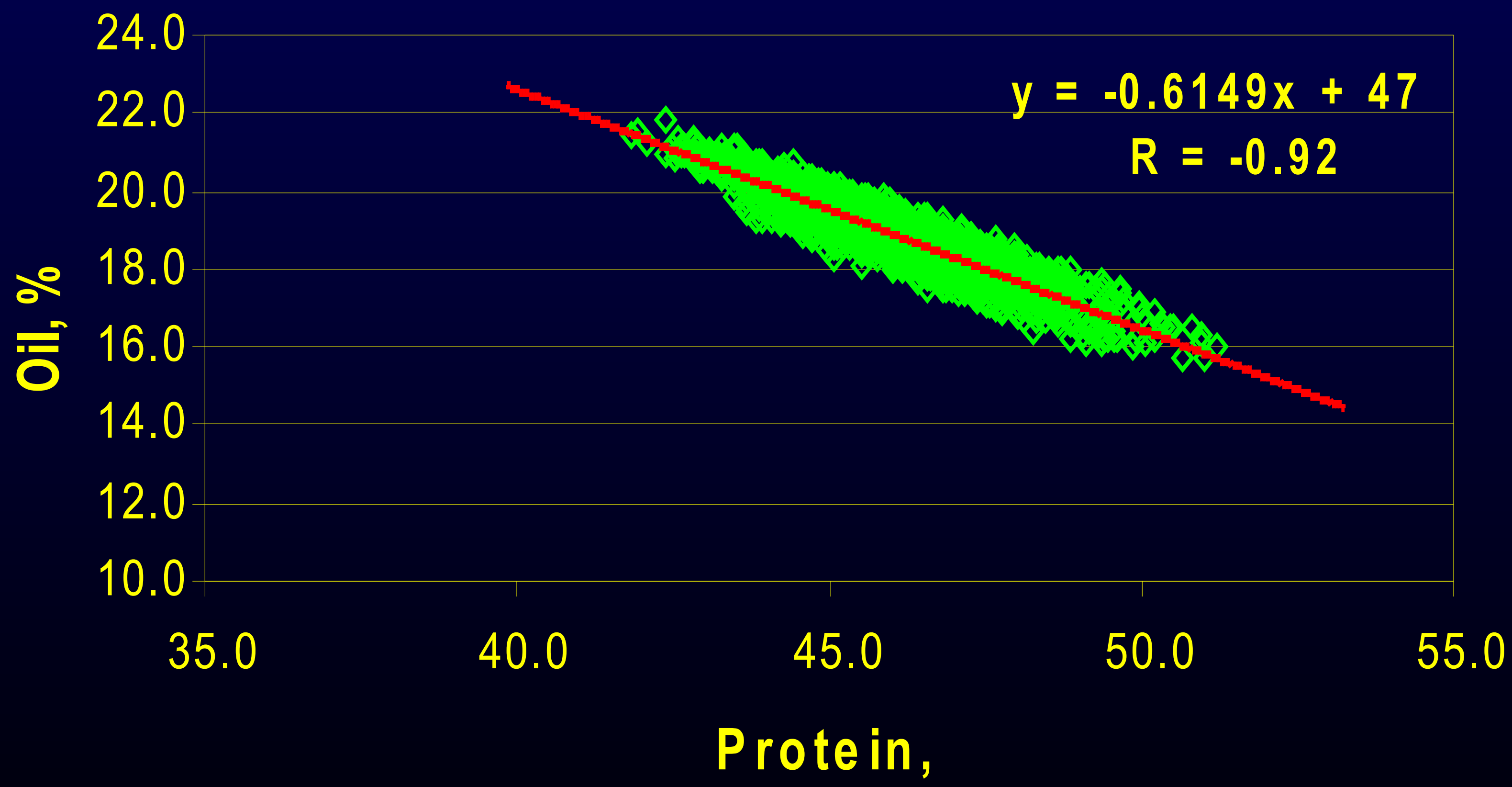
Single Intact Seeds

| Component | Number of Factors | R | SECV |
|---|---|---|---|
| Protein | 14 | 99.4% | 0.31 |
| Oil | 14 | 99.3% | 0.15 |
| Moisture | 12 | 99.5% | 0.29 |

Single Half-Seeds

| Component | Number of Factors | R | SECV |
|---|---|---|---|
| Isoflavones | 12 | 99.1% | 0.01 |
| Protein | 12 | 99.7% | 0.27 |
| Oil | 12 | 99.6% | 0.10 |
| Moisture | 12 | 99.5% | 0.22 |

# Protein-Oil Inverse Correlation for 4500 MapIII Developmental Line Soybean Samples

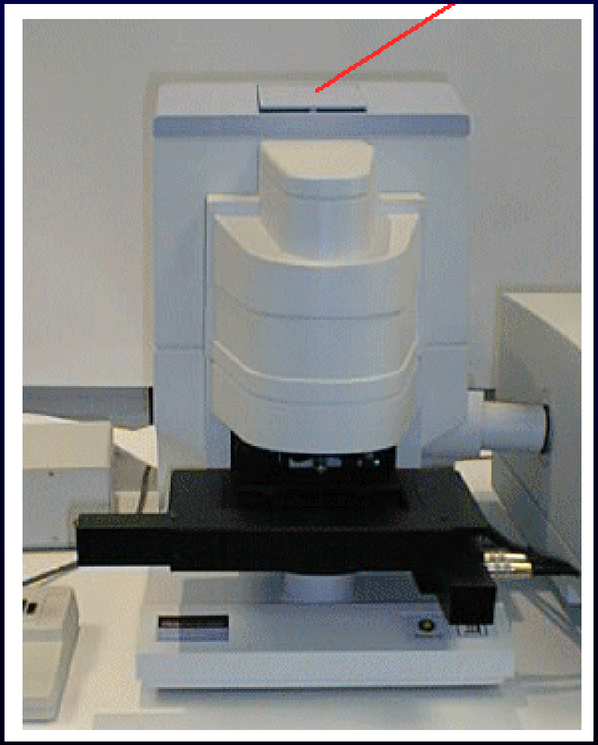


# Section D

Microspectroscopy and Chemical Imaging of Soybean Seeds

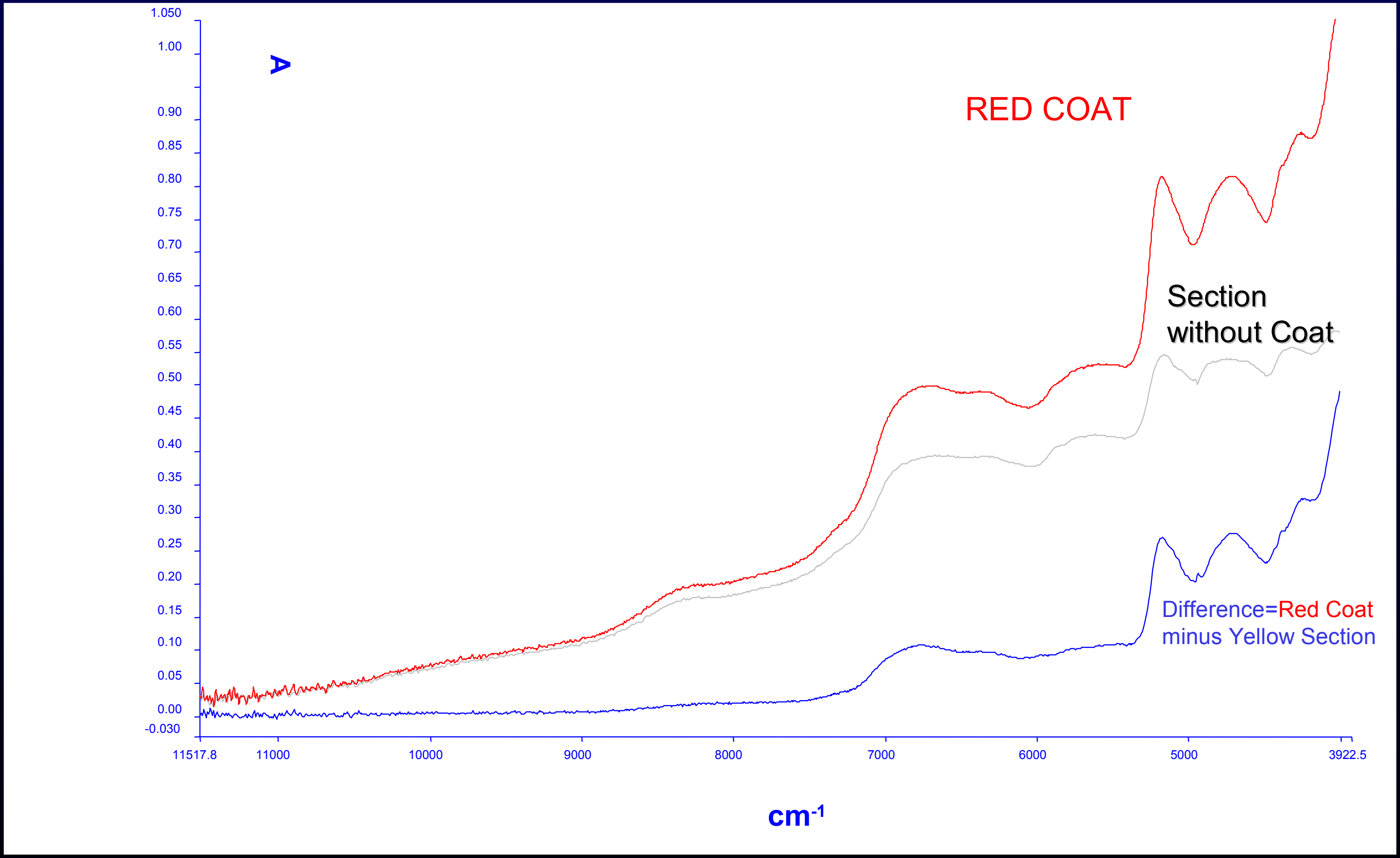
# FT-IR/NIR Microspectrometers Employed in this Study

PerkinElmer NIR AutoImage

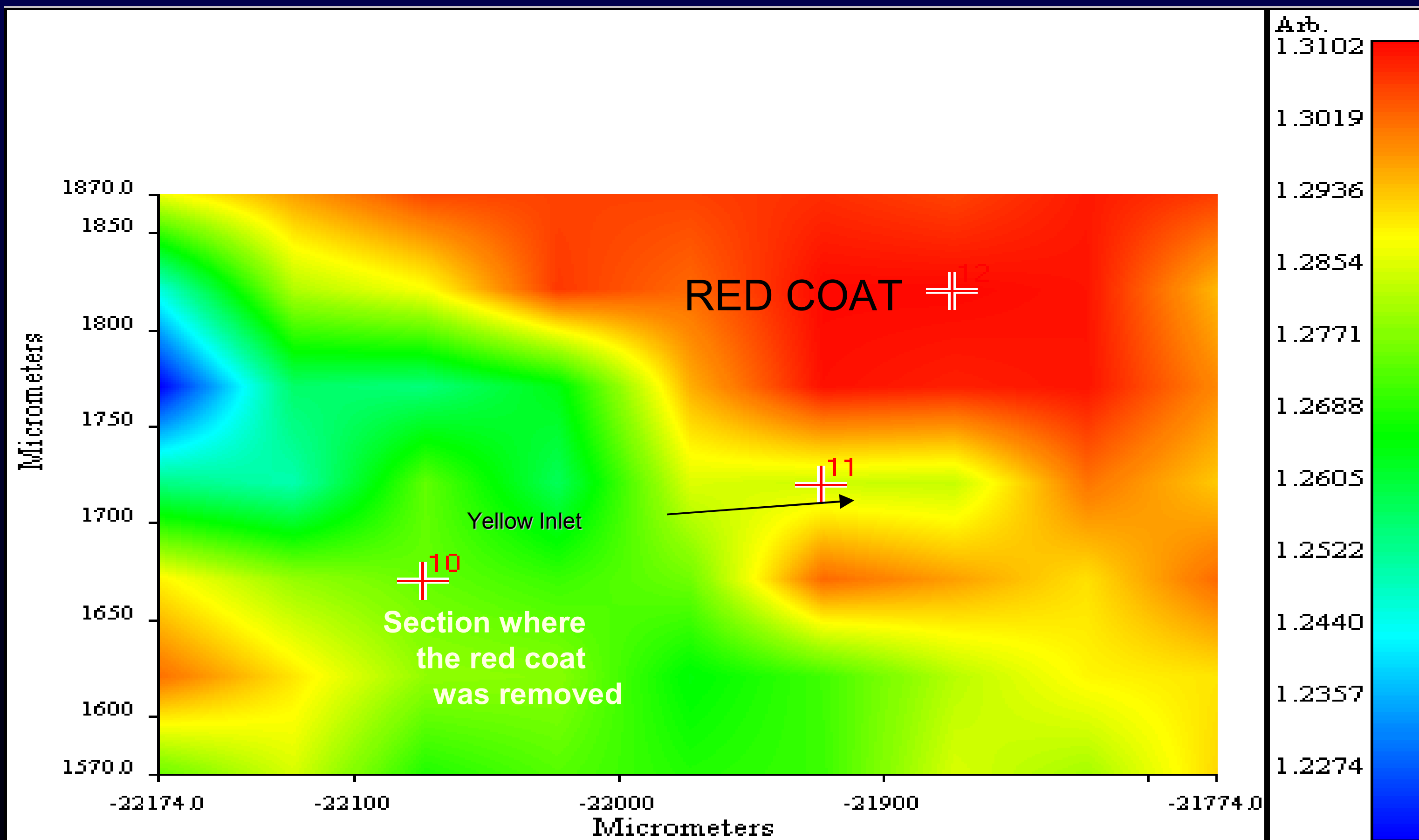


PerkinElmer Spotlight300

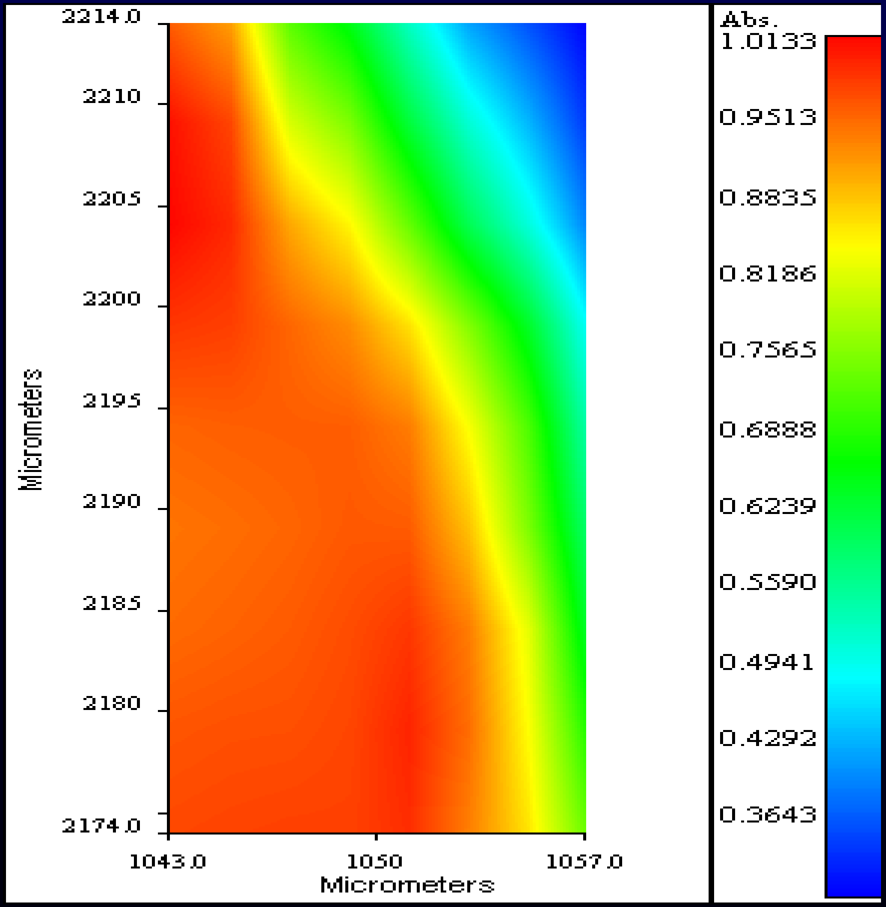
# FT-NIR Chemical Imaging by Difference Spectroscopy (CIDS): Red Coat Spectra Minus the Yellow, Under-Coat Section Spectra



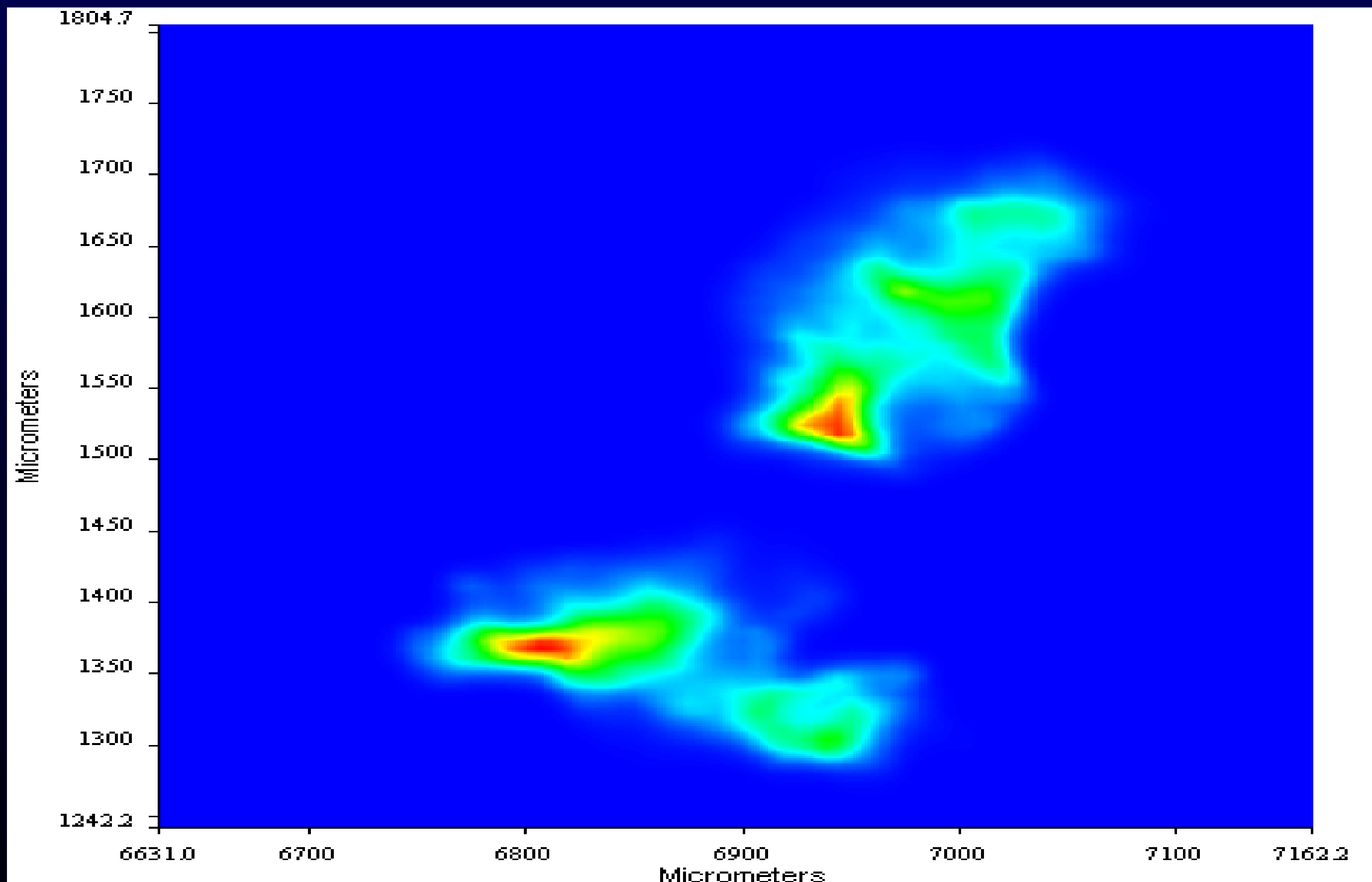
# FT-NIR Reflectance Chemical Image of a Red Coat Bean (*Vigna angularis*) and a Section without the Red Coat (left, below diagonal).



# FT-NIR Chemical Image of Oil Distribution in a Mature Soybean Embryo Section



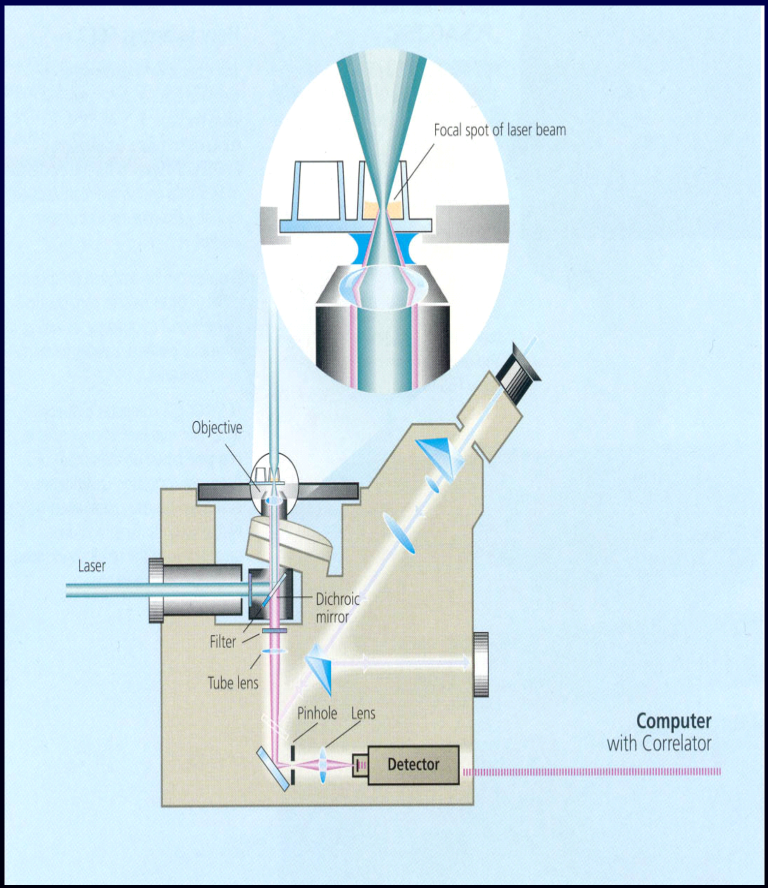
# FT-IR Chemical Image of Protein Distribution in a Mature Soybean Embryo Section

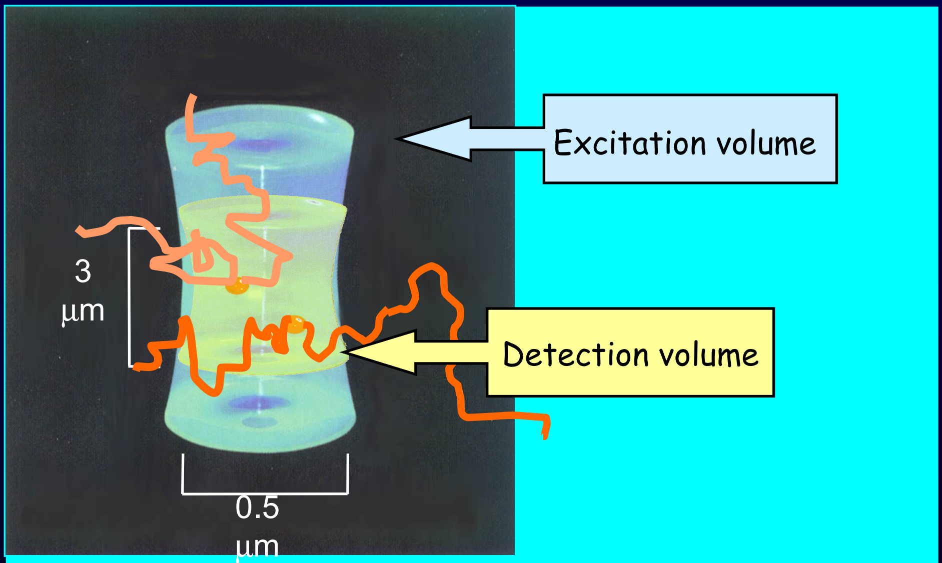


# The FCS Alba™ Spectrometer- Microscope System manufactured by ISS Inc.

# Fluorescence Correlation Spectroscopy (FCS) Principles

## Excitation and detection of fluorescence



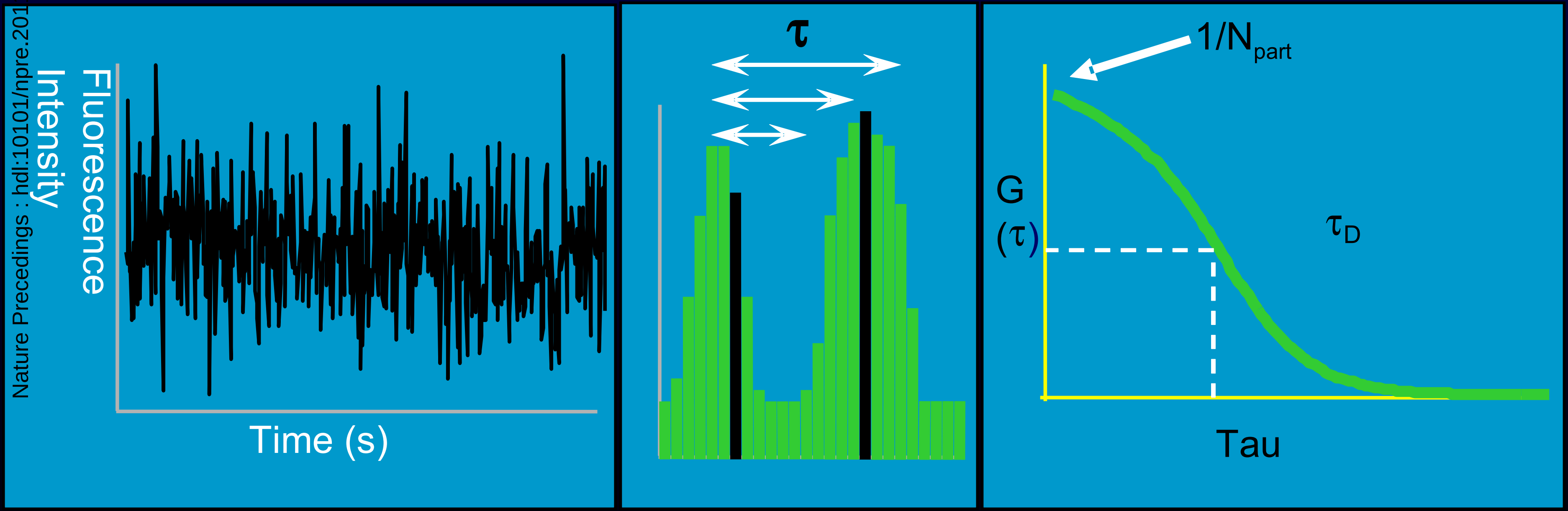




Volume: < 1 fL
Concentration: ~ 1 nM

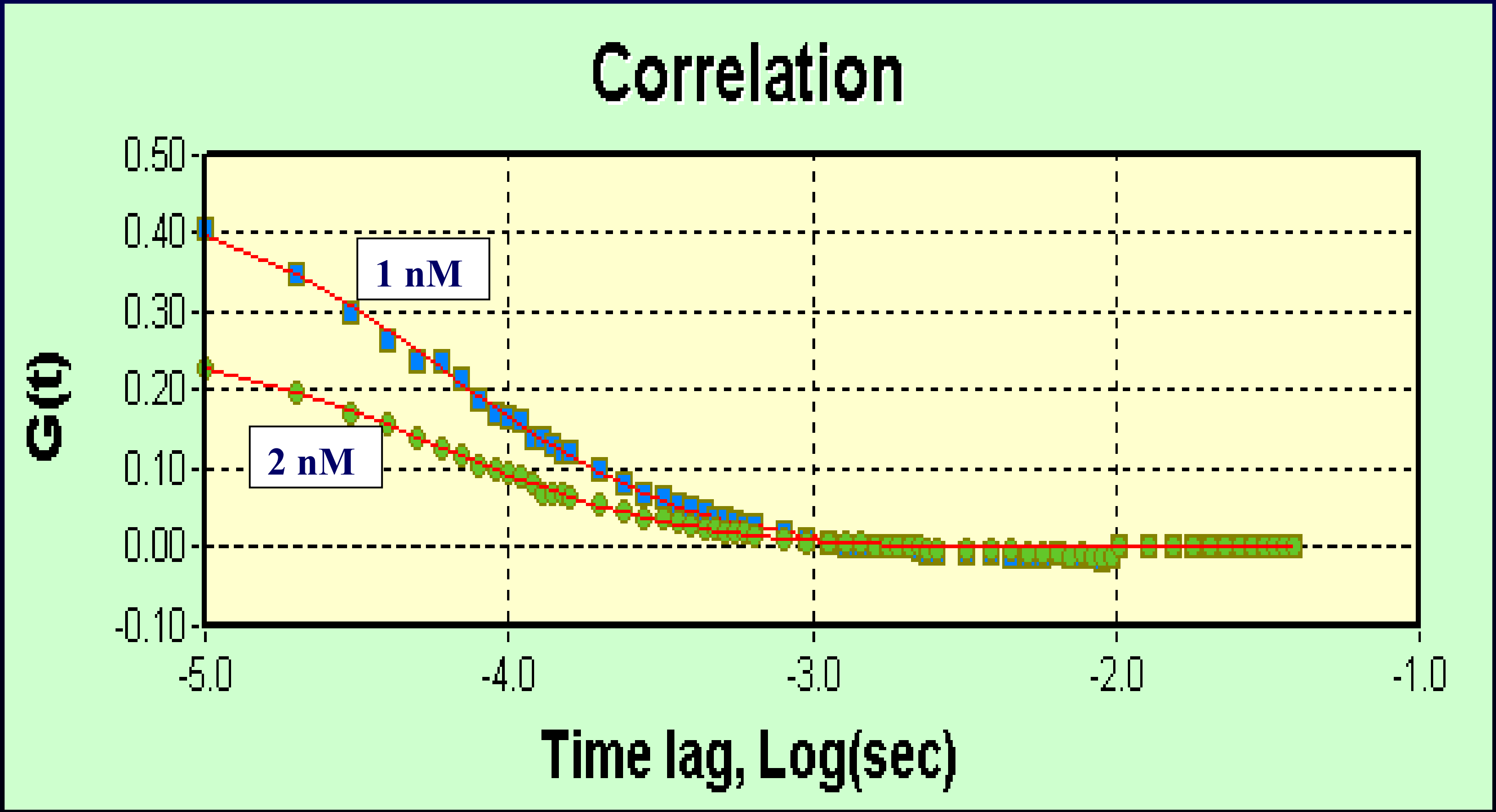
# FCS Principles:

## Autocorrelation of the fluorescence fluctuations

$$G(\tau) = \frac{I(t) * I(t+\tau)}{<I>^2} = 1 + \frac{<\delta I(t) * \delta I(t+\tau)>}{<I>^2}$$



# Auto-Correlation of 1 and 2 nM Solutions of the Fluorescent Dye Rhodamine 110G



# Potential Applications of Microspectroscopy

- Single molecule detection
- Live cell study
- Medical research
- Food biotechnologies
- Genomics

# Conclusions

- Novel spectroscopic and microspectroscopic techniques were developed and evaluated for the analysis and characterization of important chemical components of soybean.
- Computer simulation studies demonstrated the reliability of the NIR calibration algorithms.
- Effective corrections for light scattering and baseline effects could be achieved through optimization of Multiplicative Scattering Correction (MSC) and spectra derivatives methods.

# Conclusions

- Reliable, rapid, and cost effective composition analysis methodologies were developed on the state-of-the-art NIR instruments.
- Such methodologies can result in huge savings of costs for rapid composition analysis.

# Conclusions

- Novel-design instrumentation for FT-IR/NIR Microspectroscopy/ Chemical Mapping is now capable of automated visualization of composition distribution and developmental changes in single soybean seeds and embryos.

- The NIR Microspectroscopy analysis precision is now in the *picogram* range. This would allow researchers to speed up the genetic selection of soybean lines with improved protein and oil composition.

- Fluorescence Correlation Spectroscopy (FCS), furthermore, provides submicron spatial resolution, and sensitivity to single molecules level, thus offering the most exciting opportunities, for example, to monitoring DNA hybridization kinetics, ligand-receptor interactions.

# Acknowledgements


- Professor Randall L. Nelson
- Professor William Artz
- Professor Hao Feng
- Professor Manabu Nakamura

- Dr. Jun Guo
- Mr. Valentin Prisecaru
- Ms. Doina M. Costescu
- Ms. Hsiao-Chen Lin

- Renessen Co., C-FAR, ISPOB and USDA
- Perkin-Elmer Instrument Co.
- Perten Instrument Co.
- Nicolet Co.
- Bruker Co.

# Reference Methods for Soybean Composition Analysis

- Total Protein
  - Kjeldahl (AOCS method Ac 4-91)
  - Lowry / Biuret, $^{13}C$ NMR
- Total Oil
  - Solvent extraction (AOCS method Ac 3-44)
  - $^{13}$CNMR: AOCS-approved analyses
- Moisture
  - Oven drying (AOCS method Ac 2-41)
  - NMR, and Isopiestic method: AOCS-approved
- Isoflavones
  - Extraction first, followed by HPLC Analysis

# Notations

- Standard Error of Cross Validation (SECV) and Standard Error of Prediction (SEP):

$$SECV = SEP = \sqrt{\frac{\sum_{i=1}^{n} (Y_{P(i)} - Y_{R(i)})^2}{n}}$$

- Coefficient of Determination ($R^2$):

$$R^2 = \frac{\sum_{i=1}^{n} (Y_{P(i)} - \overline{Y_R})^2}{\sum_{i=1}^{n} (Y_{R(i)} - \overline{Y_R})^2}$$

$Y_P$: Prediction value; $Y_R$: Reference value

# Current Status of NIR Applications

- NIR is widely applied in agricultural, pharmaceutical, biomedical, and material science.
- There are several NIR reports of calibrations for soybean protein and oil analysis, but they have several problems and limitations, such as:
  - Most commercial NIR calibrations for protein and oil analysis are only in a narrow range and therefore have limited use for genetic selection and breeding studies
  - There are no published NIR reports on single soybean analysis
  - There is no published NIR work on isoflavone analysis of soybeans
- Advances in both instrumentation and data analysis are needed.

# NIR Instruments

- A wide variety of NIR instruments available:
  - Reflectance, transmittance, transflectance
  - Dispersive, Fourier Transform (FT)
  - Filters, grating
  - Moving grating, static grating (Diode Array)

- Accessories
  - Integrating spheres
  - UpIR
  - ATR
  - Etc.

- Important factors affecting selection of NIR instruments: speed, resolution, sensitivity, reproducibility, spectra range, costs, sample requirements, etc.

# Diode Array (DA-NIR) Instruments

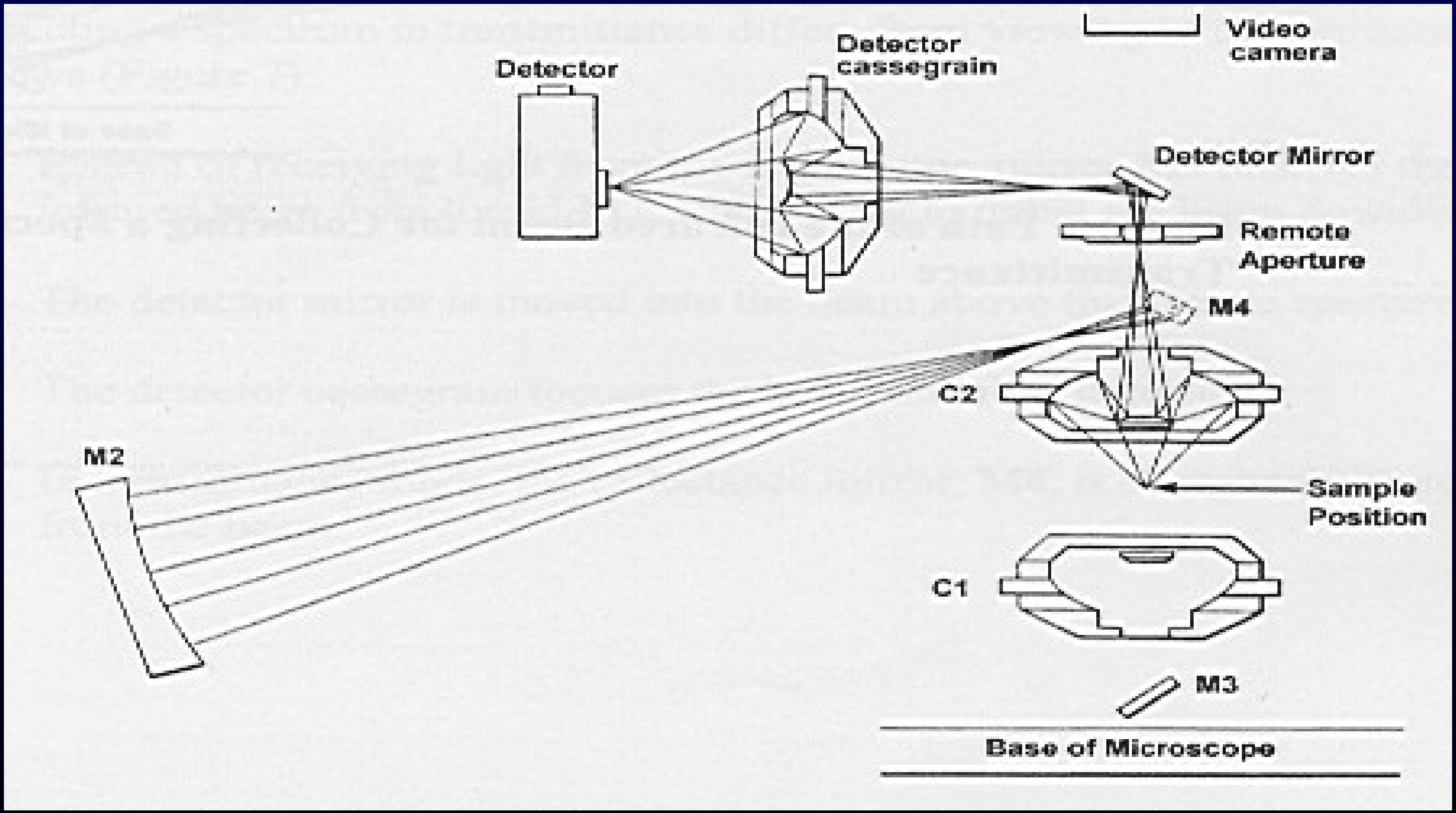
- Perten's DA7000 Dual-Diode Array Instrument
  - Advantages: Speed, sensitivity towards visible range
  - Disadvantages: Limited resolution, low thermal stability, spectral range

# Fourier Transform (FT-NIR) Instruments, ctd

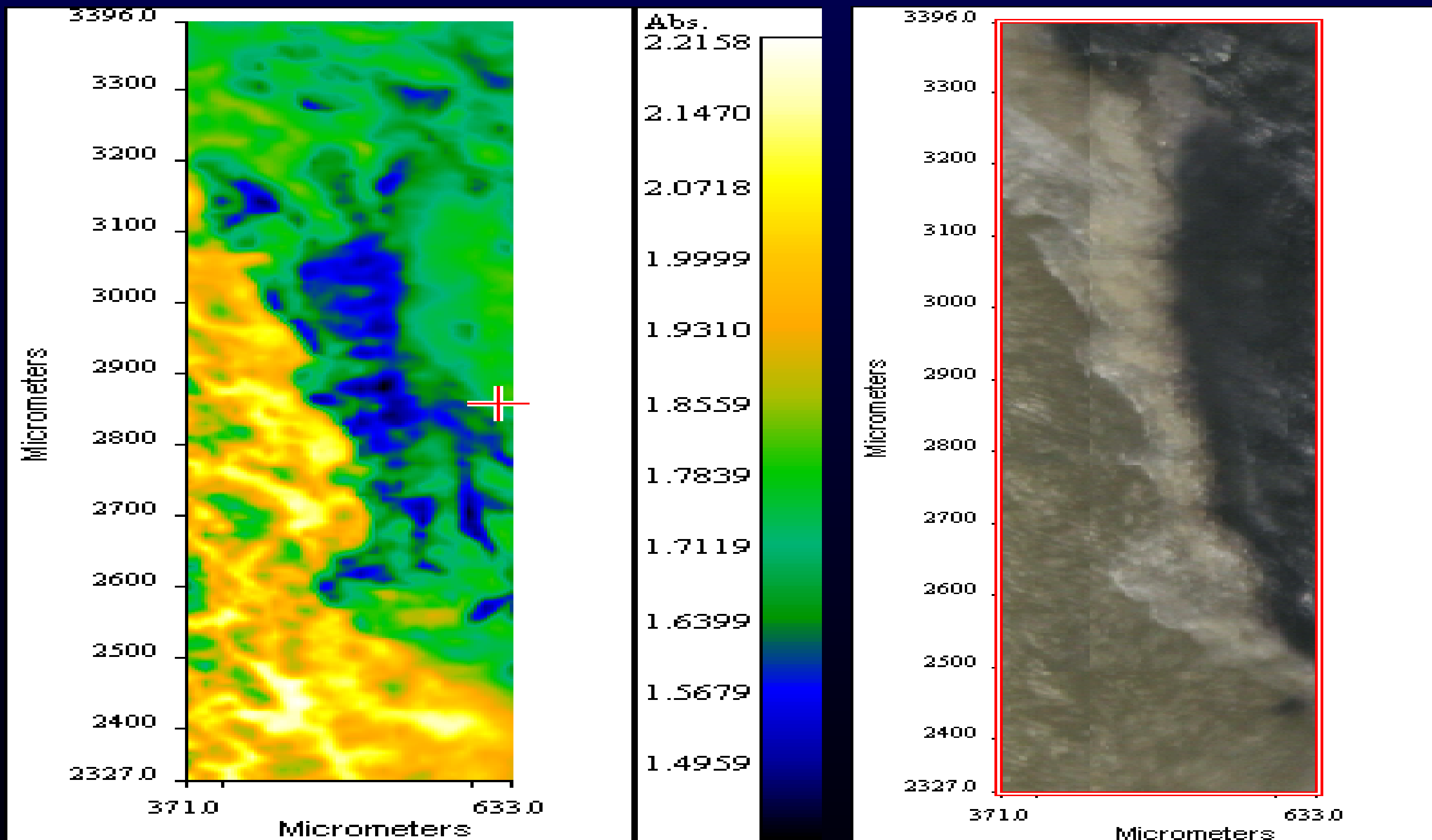
- Bruker's Vector-22 FT-NIR Spectrometer
  - Detector: PbSe
  - Spectral Range: 10000-4000 $cm^{-1}$

- Bruker's Vector-33 FT-IR/NIR Spectrometer
  - Detector: PbSe
  - Spectral Range:10000- 4000 $cm^{-1}$

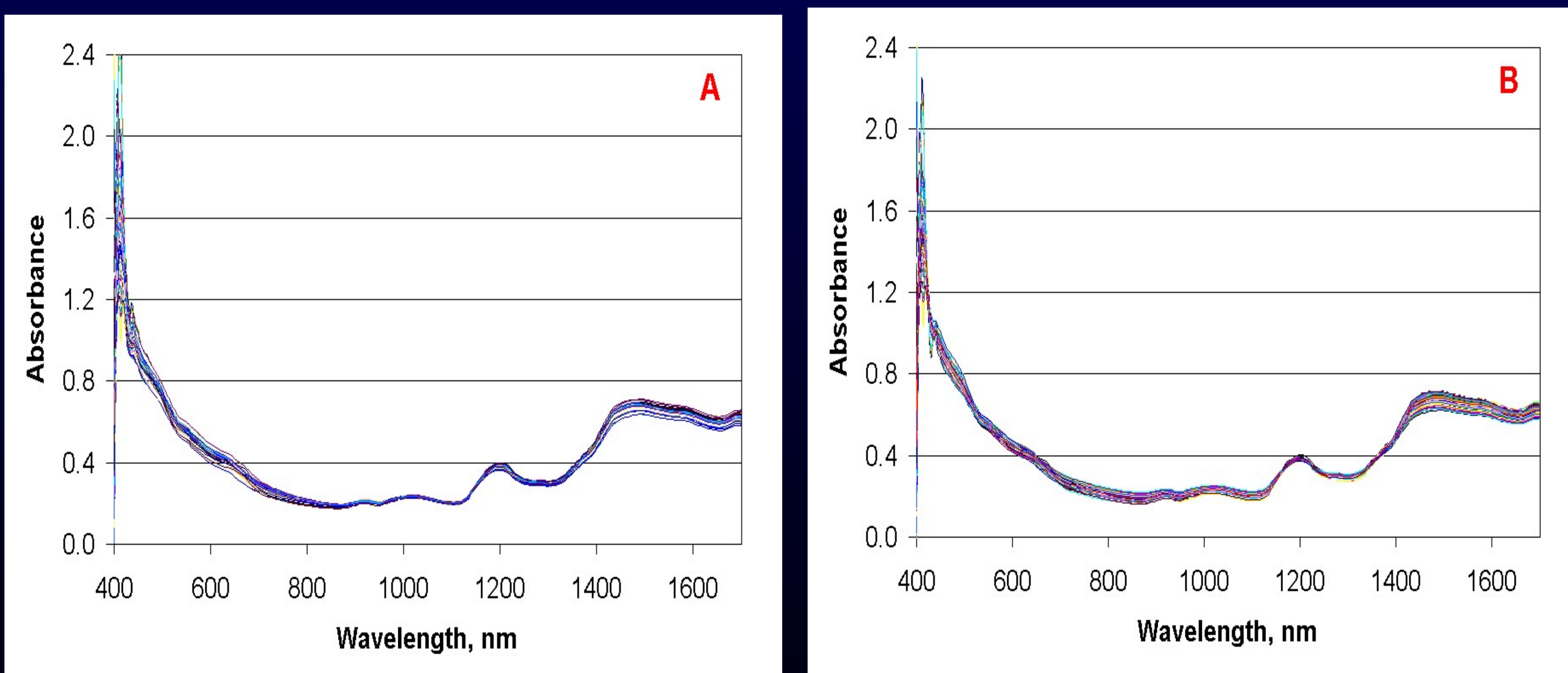
# Simplified Diagram for NIR Microspectroscopy (Perkin-Elmer NIR AutoImage System)



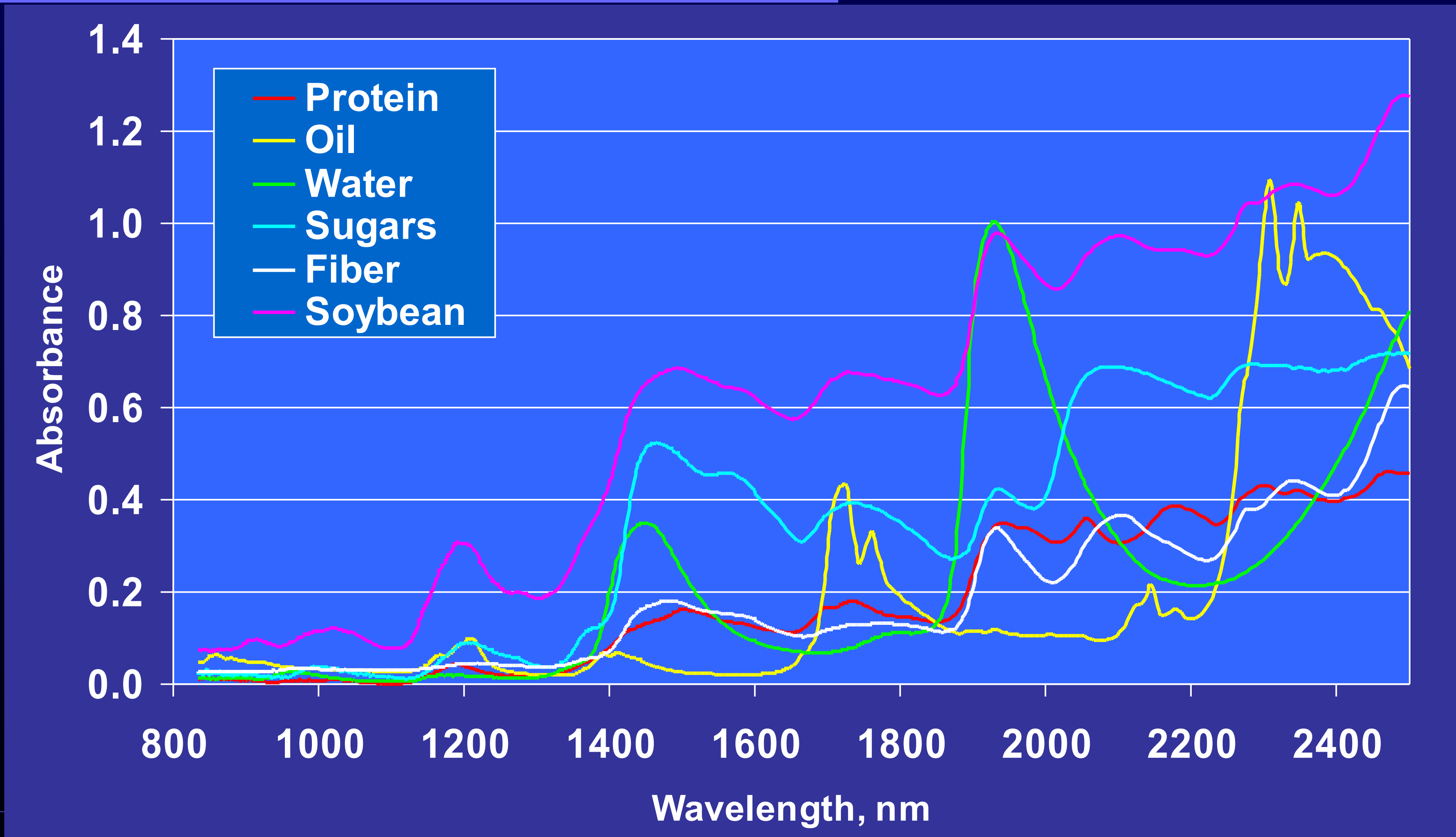
# FT-IR Chemical Image (Left) and Visible Light Micrograph (Right) of a Black Coat Soybean with Part of the Coat Removed



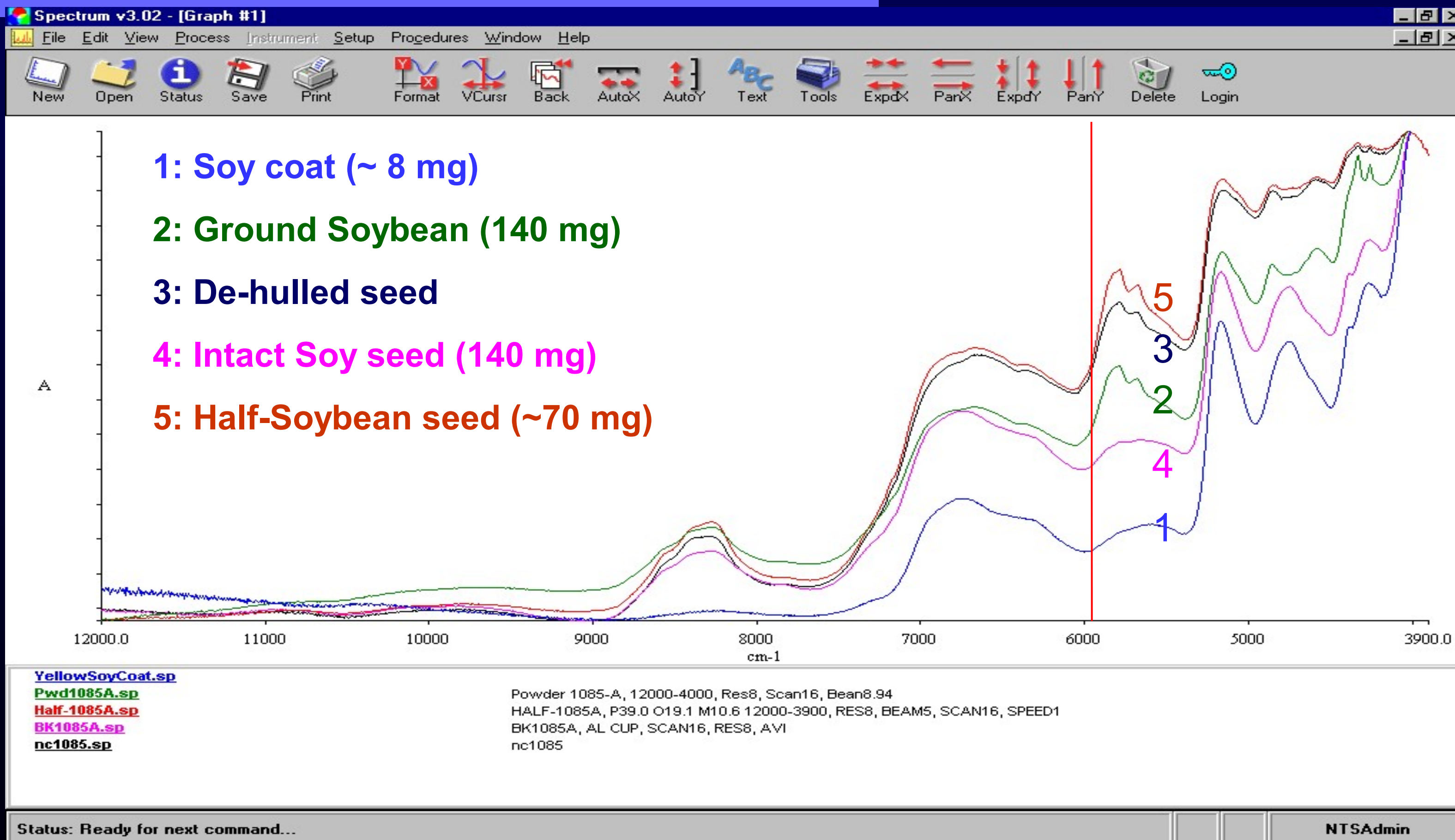
# DA-NIR Spectra of Bulk Soybean Samples, before (A) and after (B) MSC (400-1700nm)



# FT-NIR Spectra of Major Soybean Components Obtained with SpectrumOne NTS



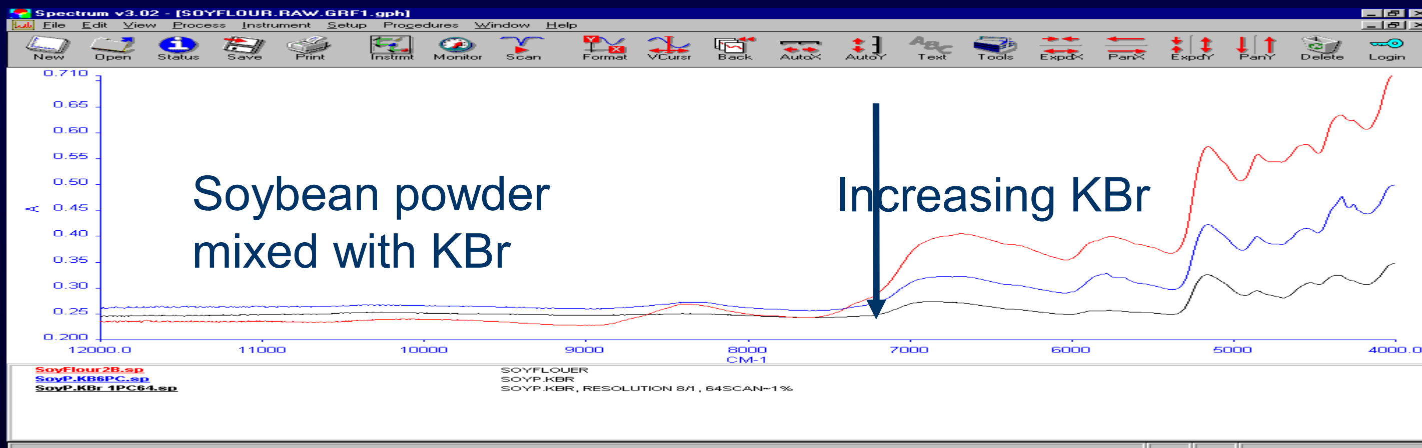
# FT-NIR Spectra of Intact, Single Soybean Seeds (4), Half-Seeds (5), De-hulled Soybean Seeds(3), and Coats(1), Compared with those of Ground Soybean Seeds (2)

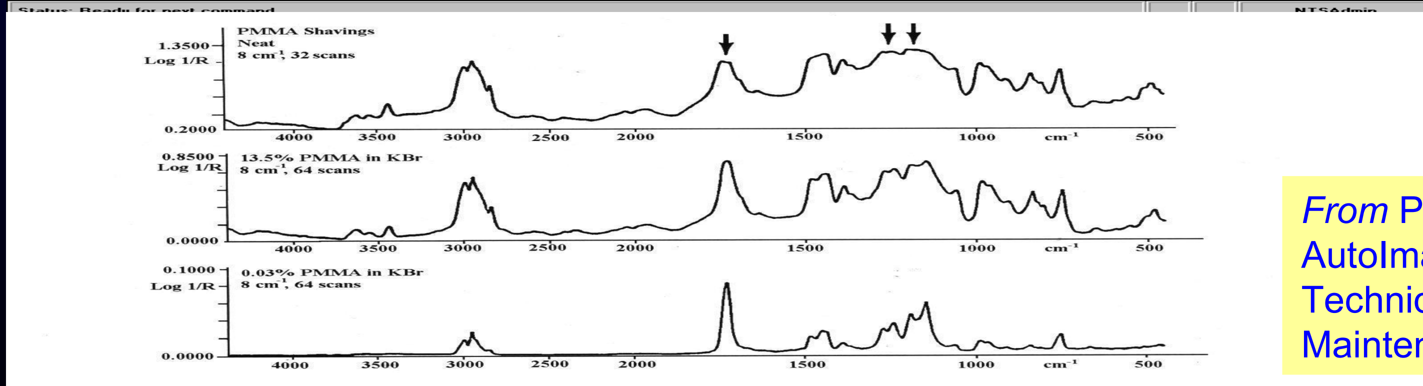



**Note:** Spectra are autoscaled, the red vertical line indicates the limit of the DA-7000 Diode Array spectra, and also the effects on the Oil NIR band shapes/intensity at ~5800 and 4200 $cm^{-1}$

# NIR spectra pre-processing

Effects of Specular Reflectance on FT-NIR Spectra of Soybeans and FT-IR Spectra of PMMA Polymers with/ without KBr

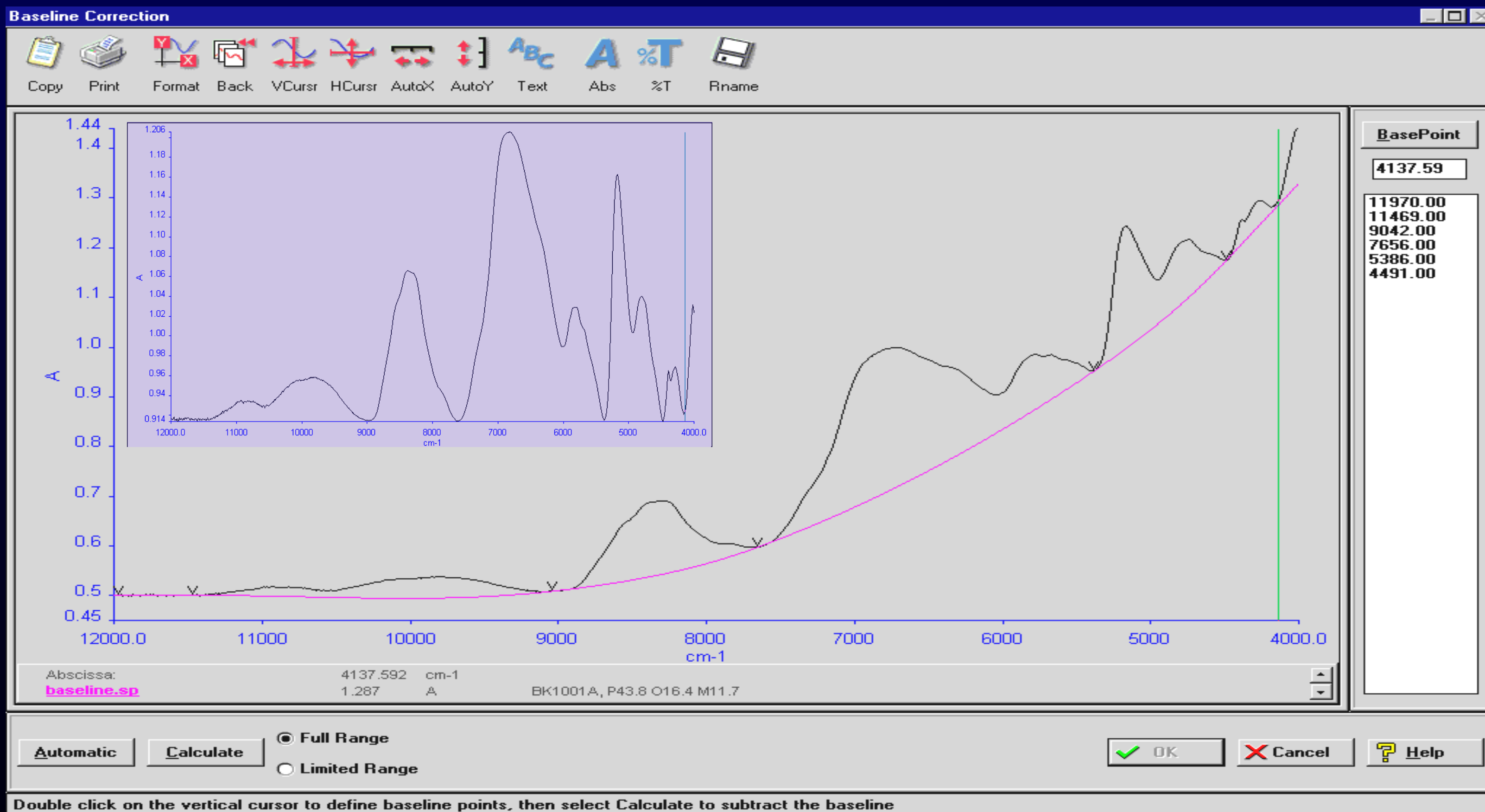



Mid-IR Absorption Spectra of PMMA polymer powder collected in the Diffuse Reflectance mode with an AI, Perkin-Elmer Mid-IR microscope at different levels of added KBr Powder

*From* Perkin-Elmer: AutoImage Microscope Techniques and Maintenance Guide

# Illustration of an Interactive Baseline Correction* of FT-NIR Soybean Spectrum




*Note: Only PE Spectrum program supports it

# Cubic Spline Functions

- Given a set of points $\{(x_1,y_1), (x_2, y_2), \ldots(x_n, y_n)\}$, generate a cubic function between every two adjacent point:

$$x(t) = a_x t^3 + b_x t^2 + c_x t + d_x$$

$$y(t) = a_y t3 + b_y t^2 + c_y t + d_y$$

- In order to solve the 4n-4 variables, setup constraints:
  - Zero order continuity condition: curve pass through points
  - First order continuity condition: 1st derivative continuous
  - Second order continuity condition: 2nd derivative continuous
  - Additional constraints at two end points: Natural cubic spline, Hermite spline, Cardinal spline, etc.

# MSC Coefficients vs Fat% of Meat

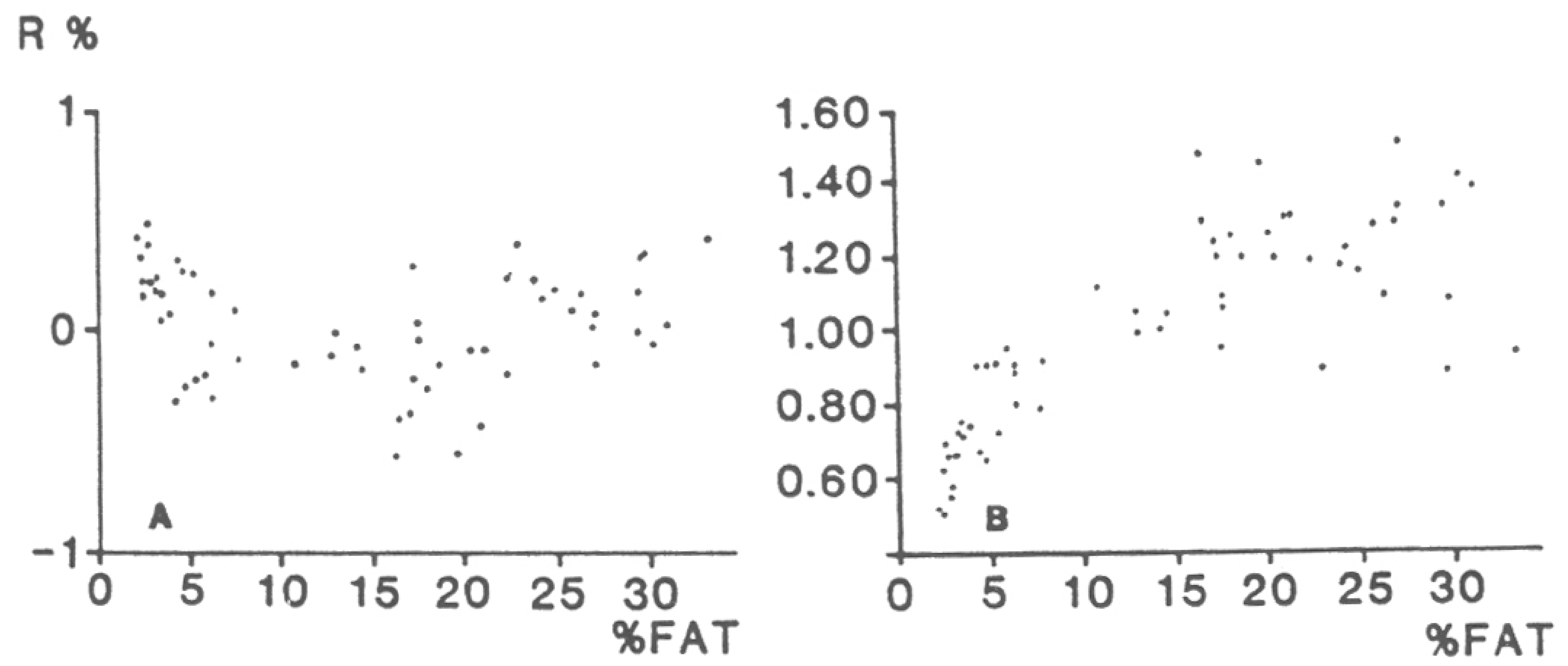
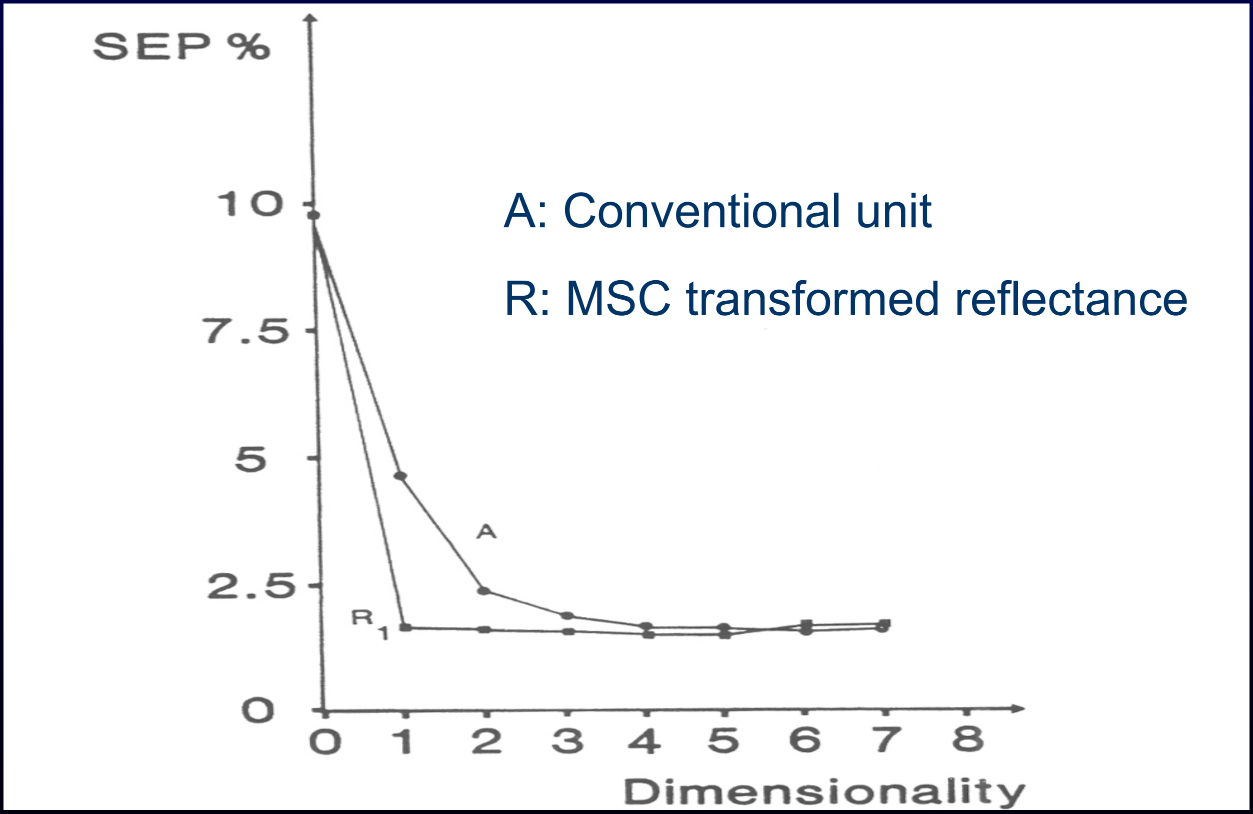



FIG. 10. MSC-correction coefficients $a_i$ (additive) (A) and $b_i$ (multiplicative) (B) for reflectance ($R$) data from the sample set 3, plotted against fat percentage.

# Effect of MSC on Accuracy of Fat Analysis in Meat



# Effect of Light Scattering on NIR Analysis Accuracy (DA7000 Calibration)

| | Number of Factors | | $R^2$ | | SECV | |
|---|---|---|---|---|---|---|
| Component | NoMSC | MSC | NoMSC | MSC | NoMSC | MSC |
| Protein | 12 | 7 | 0.980 | 0.995 | 0.77 | 0.45 |
| Oil | 8 | 7 | 0.970 | 0.989 | 0.49 | 0.30 |
| Moisture | 12 | 7 | 0.950 | 0.988 | 0.60 | 0.23 |

$R^2$: Coefficient of determination
SECV: Standard Error of Cross Validation

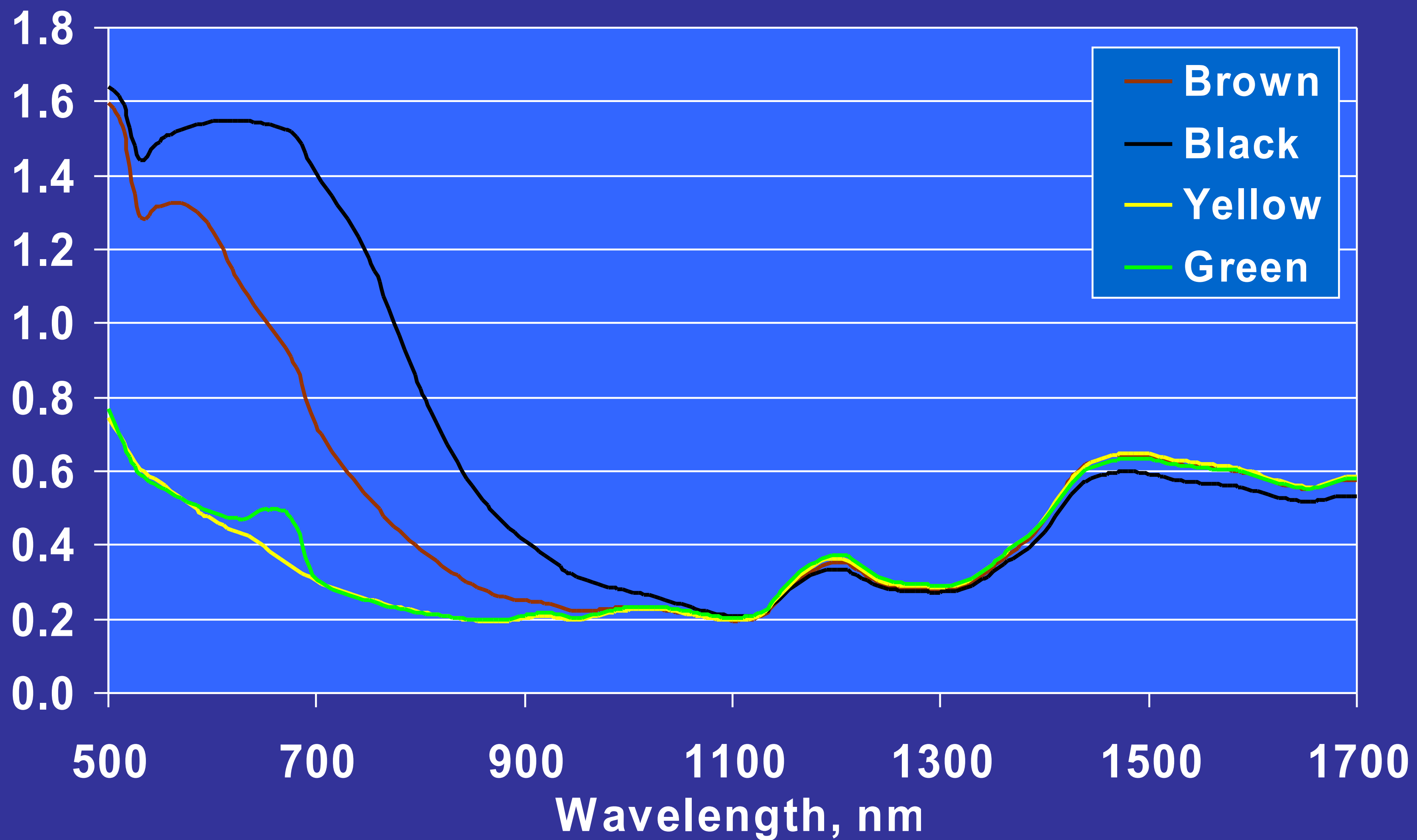
# Calibration of Perten's DA7000 Instrument for Single Seed Soybean Analysis

| Components | Number of Factors | $R^2$ | SECV |
|---|---|---|---|
| Protein | 8 | 97.4% | 1.1 |
| Oil | 7 | 98.0% | 0.5 |
| Moisture | 8 | 98.7% | 0.3 |

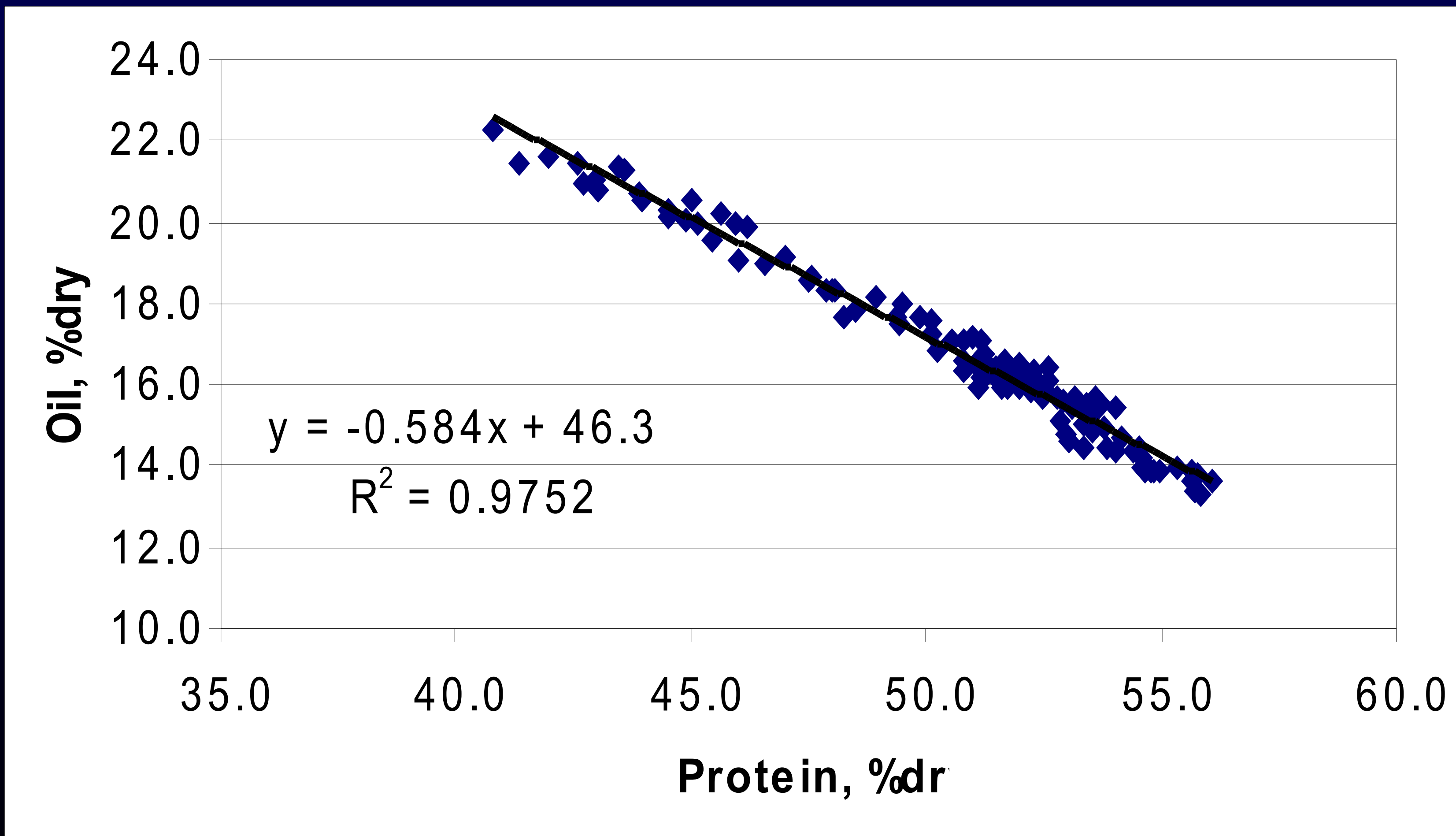
$R^2$: Coefficient of determination

SECV: Standard Error of Cross Validation

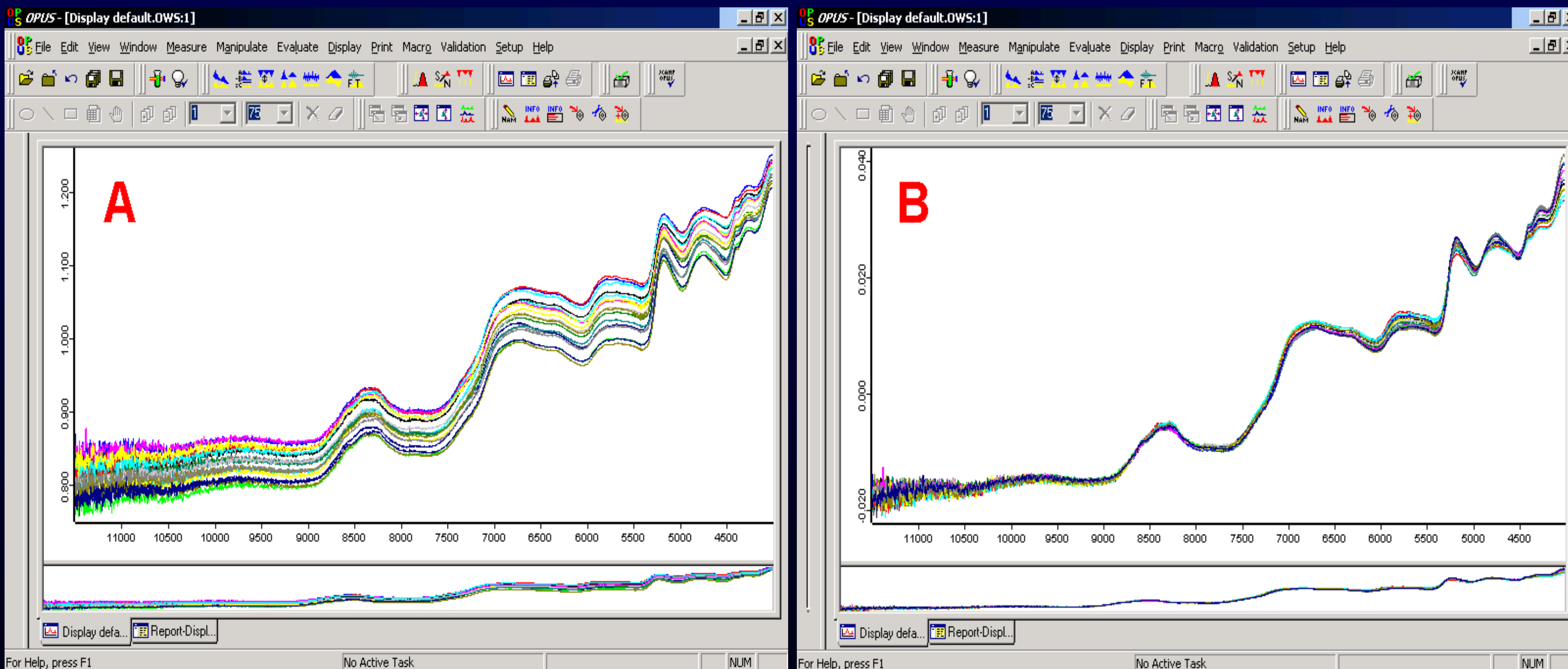
# DA-NIRS Spectra of Soybeans with Different Coat Colors Obtained with DA-7000



# The new and Improved Set of Standard Samples: Protein-Oil Inverse Correlation for the Year 2002 Calibration Standard Samples

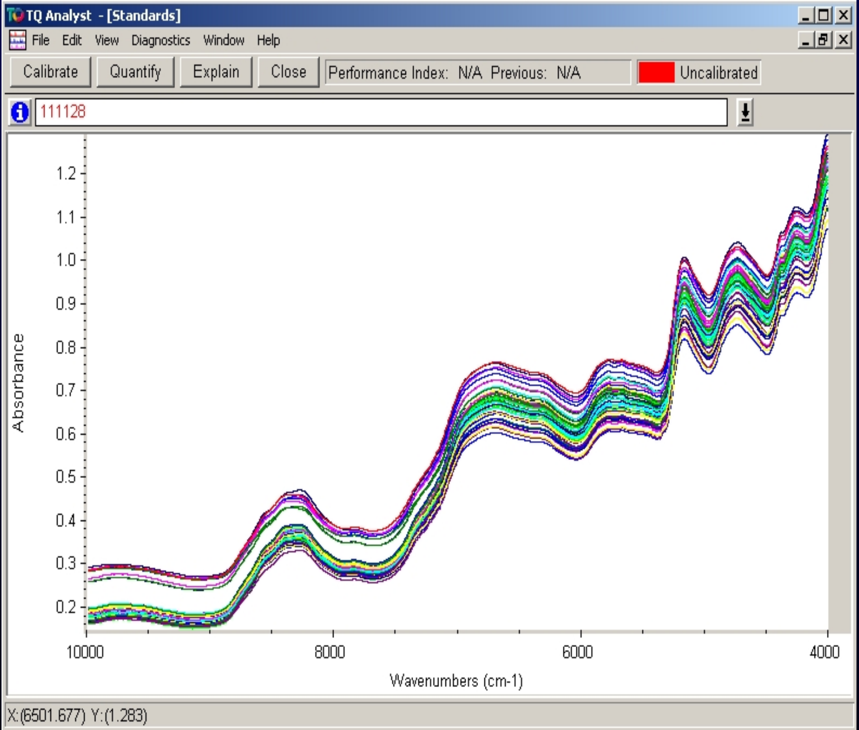


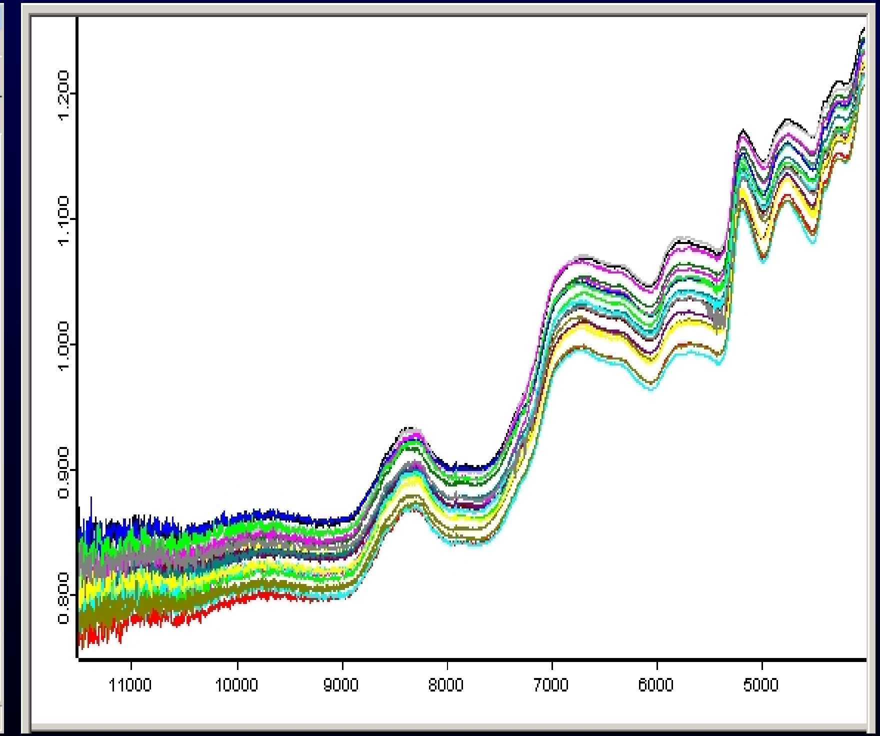
# Overlay Plot of FT-NIR Spectra of Single Soybean Seeds Obtained with the Bruker Vector-33 Instrument. A: Raw, B: MSC



# Overlay Plot of NIR Spectra of Soybeans

## Left: Vector-22, Right: Vector-33



# ATR