Multivariate Curve Resolution Applied to Infrared Reflectance Measurements of Soil Contaminated with Organic Analyte

Neal B. Gallagher

nealg@eigenvector.com

Eigenvector Research, Inc., Manson, WA USA

Thomas A. Blake, Paul L. Gassman

Pacific Northwest National Laboratories

Richland, WA USA

EIGENVECTOR
RESEARCH INCORPORATED

Objective

- Can IR reflectance spectroscopy be used to detect low vapor-pressure organic analytes on soil?
  - environmental, monitoring for compliance
- Strategies for measurement and data analysis depend on how the spectra of organic analytes manifest on soil
- Use MCR to extract estimates of liquid spectra measured on soil contaminated with organic analyte

EIGENVECTOR
RESEARCH INCORPORATED

Multivariate Curve Resolution

- MCR is a method for extracting estimates of pure spectra and concentrations from measured spectra
  - often used for exploratory analysis when spectra and concentrations are unknown

  - Literature filled with examples from evolving data
    - LC-MS, GC-NIR, GC-GC …
  - Many newer examples include multivariate images
    - Image Mid-IR, NIR, UV-Vis, Raman, Remote Sensing ...

EIGENVECTOR
RESEARCH INCORPORATED

Why MCR?

- The MCR model is based on physics/chemistry and results are often easily interpretable
  - can be used for quantification with appropriate constraints
- Can often extract spectra from complex measurements
  - some spectra are for analytes that may never exist in a pure state making direct measurement impossible
- Good selectivity required for unique estimate
  - rotational, multiplicative ambiguity
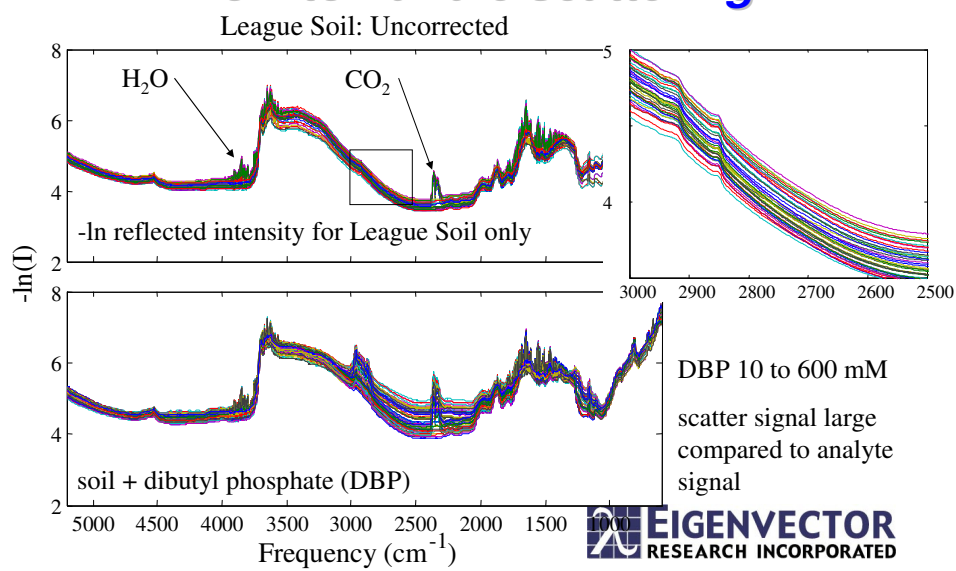 - many types of constraints used to obtain a unique estimate

EIGENVECTOR
RESEARCH INCORPORATED

2

MCR of Reflectance Measurements

- MCR used to examine analyte absorbed onto two different soils (Quincy and League)
    - are the extracted spectra different from liquid spectra?
    - can the results be used to help define strategies for detection and classification?
- Difficult problem because of scattering artifacts when measuring soils! And...
- atmospheric constituents (H_2O and CO_2) are also present!

EIGENVECTOR RESEARCH INCORPORATED

How to handle scattering?

League Soil: Uncorrected

H_2O CO_2

-ln reflected intensity for League Soil only

soil + dibutyl phosphate (DBP)

DBP 10 to 600 mM

scatter signal large compared to analyte signal

-ln(I)

Frequency (cm^{-1})

EIGENVECTOR RESEARCH INCORPORATED

3

Typical MCR Model

- Based on the classical least squares (CLS) model, attempt to estimate **C** and **S** given **X**:

$$\mathbf{X} = \mathbf{CS}^T + \mathbf{E}$$

   X *M*x*N* measured responses,

   C *M*x*K* pure analyte contributions,

   S *N*x*K* pure analyte spectra, and

   E *M*x*N* residuals.

EIGENVECTOR
RESEARCH INCORPORATED

MCR Algorithms

- Most popular algorithms uses an alternating constrained least squares procedure (easy to code)
  - Non-negativity on **C** and **S** is most common
- Other algorithms
  - Geometric approach (e.g. N-FINDR, SIMPLSMA, DISTSLCT)
    - often fast, good first guess but doesn't solve problem in least squares sense, difficult to include many types of constraints
  - Gauss-Newton, Levenberg-Marquardt (PMF, ICE, dGN)
    - typically not as fast as ALS, use penalty functions or Lagrangian multiplier methods, ~easy to modify to include new constraints
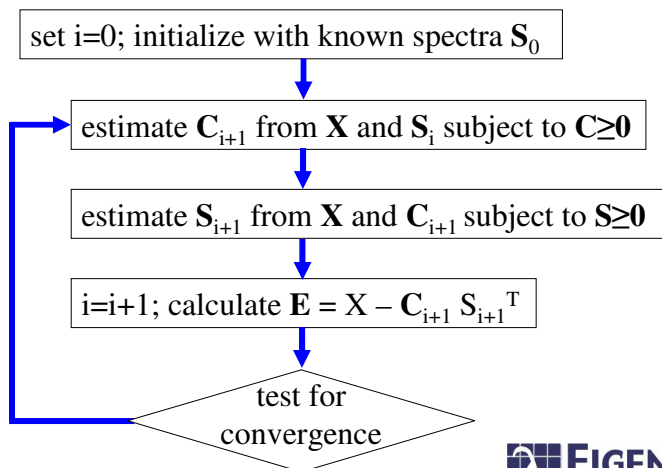
EIGENVECTOR
RESEARCH INCORPORATED

MCR Decomposition

$$\mathbf{X} = \mathbf{C}\,\mathbf{S}^{\mathrm{T}} + \mathbf{E}$$

$\mathbf{C} \geq 0$; $\mathbf{S} \geq 0$

EIGENVECTOR
RESEARCH INCORPORATED

Alternating Least Squares

set i=0; initialize with known spectra $\mathbf{S}_0$

estimate $\mathbf{C}_{i+1}$ from $\mathbf{X}$ and $\mathbf{S}_i$ subject to $\mathbf{C} \geq 0$

estimate $\mathbf{S}_{i+1}$ from $\mathbf{X}$ and $\mathbf{C}_{i+1}$ subject to $\mathbf{S} \geq 0$

i=i+1; calculate $\mathbf{E} = \mathbf{X} - \mathbf{C}_{i+1}\,\mathbf{S}_{i+1}^{\mathrm{T}}$

test for convergence

EIGENVECTOR
RESEARCH INCORPORATED

Extended Mixture MCR Model

$$\mathbf{X} = \begin{bmatrix} \mathbf{C}_{Chem} & \mathbf{C}_{Atm} & \mathbf{C}_{Soil} \end{bmatrix} \begin{bmatrix} \mathbf{S}_{Chem} & \mathbf{S}_{Atm} & \mathbf{S}_{Soil} \end{bmatrix}^T + \mathbf{E}$$

$\mathbf{C}_{Chem}$ $M \times K_C$ pure analyte contributions

$\mathbf{C}_{Atm}$ $M \times K_S$ atmosphere analyte contributions

$\mathbf{C}_{Soil}$ $M \times K_R$ soil scattering contributions

$\mathbf{S}_{Chem}$ $N \times K_C$ pure analyte spectra

$\mathbf{S}_{Atm}$ $N \times K_S$ atmosphere analyte spectra, H_2O & CO_2

$\mathbf{S}_{Soil}$ $N \times K_R$ soil scattering "spectra"

EIGENVECTOR
RESEARCH INCORPORATED

Also Works for 2nd Derivative Spec

2nd derivative adds selectivity, but $\mathbf{S}''$ are not ≥ 0

$$\mathbf{X}'' = \begin{bmatrix} \mathbf{C}_{Chem} & \mathbf{C}_{Atm} & \mathbf{C}_{dSoil} \end{bmatrix} \begin{bmatrix} \mathbf{S}''_{Chem} & \mathbf{S}''_{Atm} & \mathbf{S}''_{Soil} \end{bmatrix}^T + \mathbf{E}$$
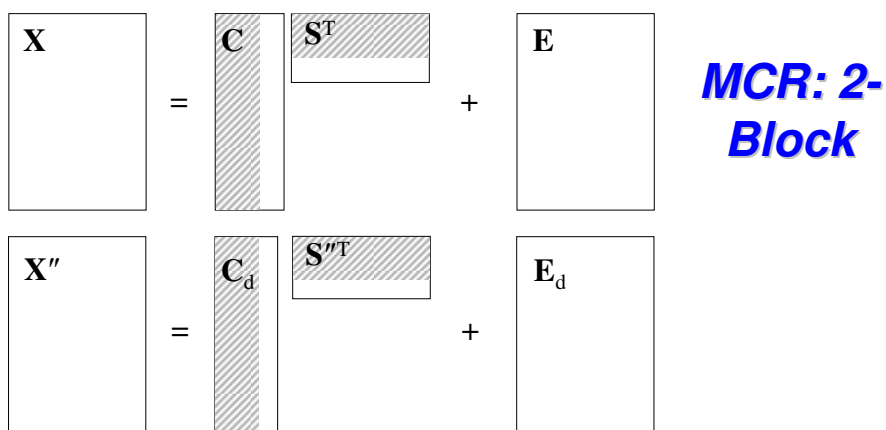
$\mathbf{C}_{Chem}$ $M \times K_C$ pure analyte contributions

$\mathbf{C}_{Atm}$ $M \times K_S$ atmosphere analyte contributions

$\mathbf{C}_{dSoil}$ $M \times K_{dR}$ soil scattering contributions

$\mathbf{S}''_{Chem}$ $N \times K_C$ pure analyte 2nd derivative spectra

$\mathbf{S}''_{Atm}$ $N \times K_S$ atmosphere analyte 2nd derivative spectra, H_2O & CO_2

$\mathbf{S}''_{Soil}$ $N \times K_{dR}$ soil scattering 2nd derivative "spectra"

EIGENVECTOR
RESEARCH INCORPORATED

How to get S_R?

- $\mathbf{S}_R$ is a sub-space that spans scatter
  - measure multiple reflectance spectra of soil samples that do not contain analyte $\rightarrow \mathbf{X}_R$
  - perform typical MCR on scatter data
    $$\mathbf{X}_R = \mathbf{C}_{R0}\mathbf{S}_R{}^T + \mathbf{E}$$
  - use $\mathbf{S}_R$ to characterize scatter
- $\mathbf{S''}_R$ is a sub-space that spans scatter
  - use mean $\mathbf{X''}_R$ to characterize 2nd derivative scatter

EIGENVECTOR
RESEARCH INCORPORATED

| $\mathbf{X}$ | $=$ | $\mathbf{C}$ | $\mathbf{S}^T$ | $+$ | $\mathbf{E}$ | *MCR: 2-Block* |

| $\mathbf{X''}$ | $=$ | $\mathbf{C}_d$ | $\mathbf{S''}^T$ | $+$ | $\mathbf{E}_d$ |

$\mathbf{C}, \mathbf{C}_d \geq 0$ $\mathbf{S} \geq 0$
$\mathbf{C}_{dChem} = \mathbf{C}_{Chem}$ $\mathbf{S}_S, \mathbf{S''}_S$ constrained to H_2O and CO_2
$\mathbf{C}_{dAtm} = \mathbf{C}_{Atm}$ $\mathbf{S}_R, \mathbf{S''}_R$ constrained to estimated subspace

EIGENVECTOR
RESEARCH INCORPORATED

ALS w/ 2 Blocks

set i=0; initialize $S_0 = [\ S_{C,0}\ S_{S,0}\ S_R\ ]$, $C_{dR,0} = 1$

est C_{i+1} from X and S_i w/ $C \geq 0$

$C''_{C,i} = C_{C,i+1}$; $C''_{S,i} = C_{S,i+1}$

est S''_{i+1} from X'' and C''_{i+1} w/ S''_R, S''_S constrained

est C''_{i+1} from X'' and S''_{i+1} w/ $C'' \geq 0$

$C_{C,i+1} = C''_{C,i+1}$; $C_{S,i+1} = C''_{S,i+1}$

calculate $E = X - C_{i+1} S_{i+1}^T$

test for convergence

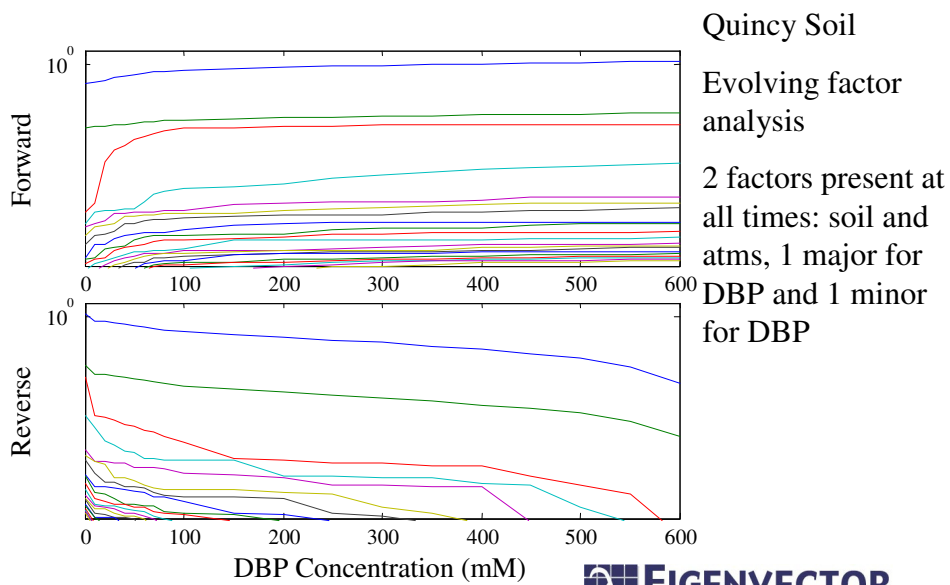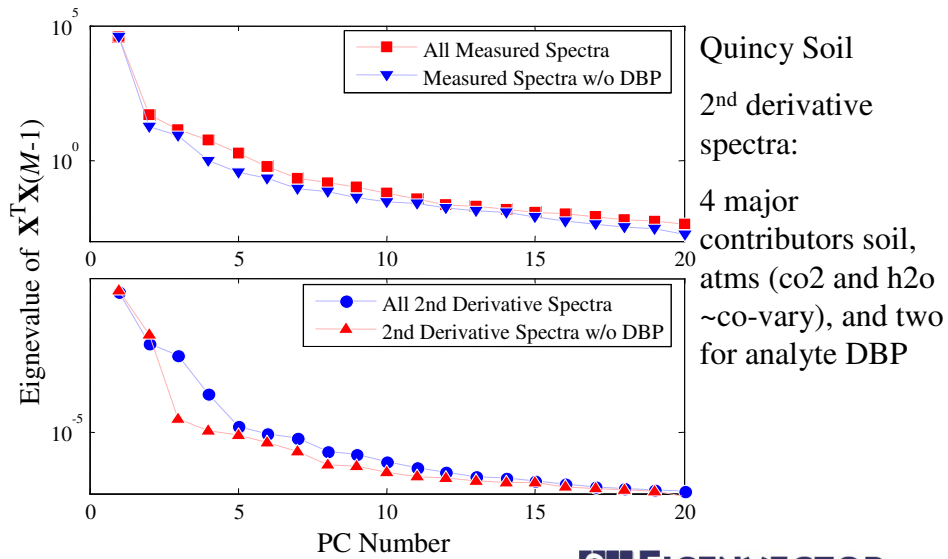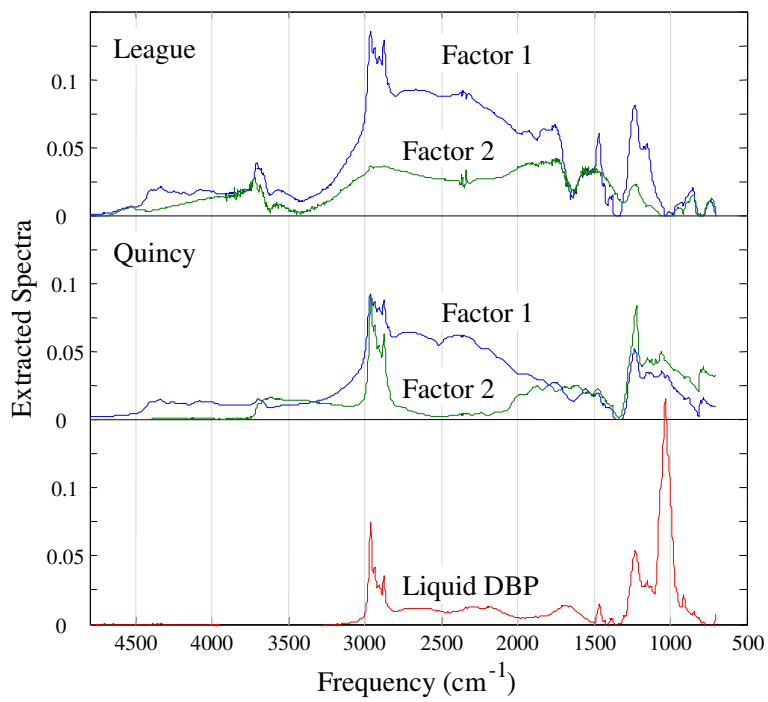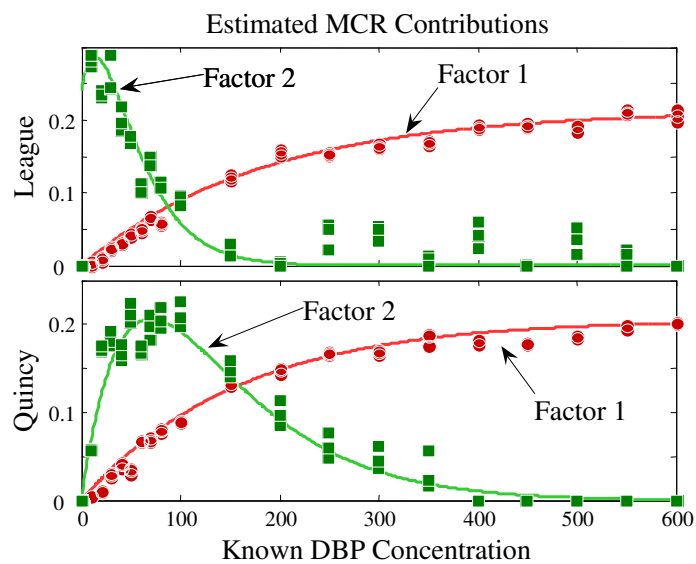est S_{i+1} from X and C_{i+1} w/ $S \geq 0$ w/ S_R, S_S constrained
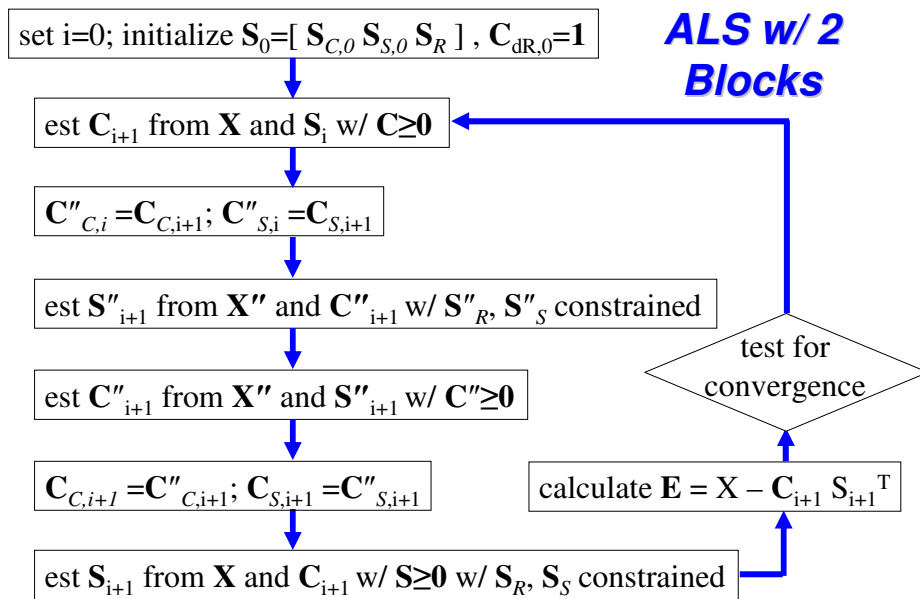
Soil/Analyte Samples

- League Soil (44% clay, 42% silt, 14% sand)
- Quincy Soil (7% clay, 17% silt, 76% sand)
- Analyte: Dibutyl phosphate in 2-Methyl Butane
  - 0, 10-600 mM dripped onto soil sample
  - 2 MB highly volatile, evaporates quickly
  - measure spectra w/ and w/o dry-N_2 purge
  - sample (DBP concentration) randomized

EIGENVECTOR
RESEARCH INCORPORATED

League Soil

2nd derivative spectra:

4 major contributors soil, atms (co2 and h2o ~co-vary), and two for analyte DBP

League Soil

Evolving factor analysis

2 factors present at all times: soil and atms, 1 major for DBP and 1 minor for DBP

Quincy Soil

2nd derivative spectra:

4 major contributors soil, atms (co2 and h2o ~co-vary), and two for analyte DBP

Quincy Soil

Evolving factor analysis

2 factors present at all times: soil and atms, 1 major for DBP and 1 minor for DBP

10

Estimated MCR Contributions

Conclusions 1/2

- Factor 1
  - increases and "saturates" (consistent w/ EMSC results)
  - C-H in 2800-3000 cm^{-1} similar to liquid, but not identical
  - increase in broad features in 1500-2800 cm^{-1}
- Factor 2
  - present only at low concentrations
  - for Quincy (~sand), C-H increasing faster than broad features
- major peak at 1032 cm^{-1} in liquid spectra missing

EIGENVECTOR
RESEARCH INCORPORATED

Conclusions 2/2

- Some differences between liquid spectra and spectra on soil
  - some minor (e.g. 2800-3000) some major (1032 cm^{-1})
  - spectra are slightly different <100 mM compared to those observed >100 mM
- Uncertain as to physical/chemical cause for differences, additional study required
- Helps to develop sensing strategies
  - e.g. pre-processing approaches

EIGENVECTOR
RESEARCH INCORPORATED