Unsupervised Iterative Algorithms for Hyperpectral Image Processing

Yahya M. Masalmah, Ph.D.
Univ. Del Turabo

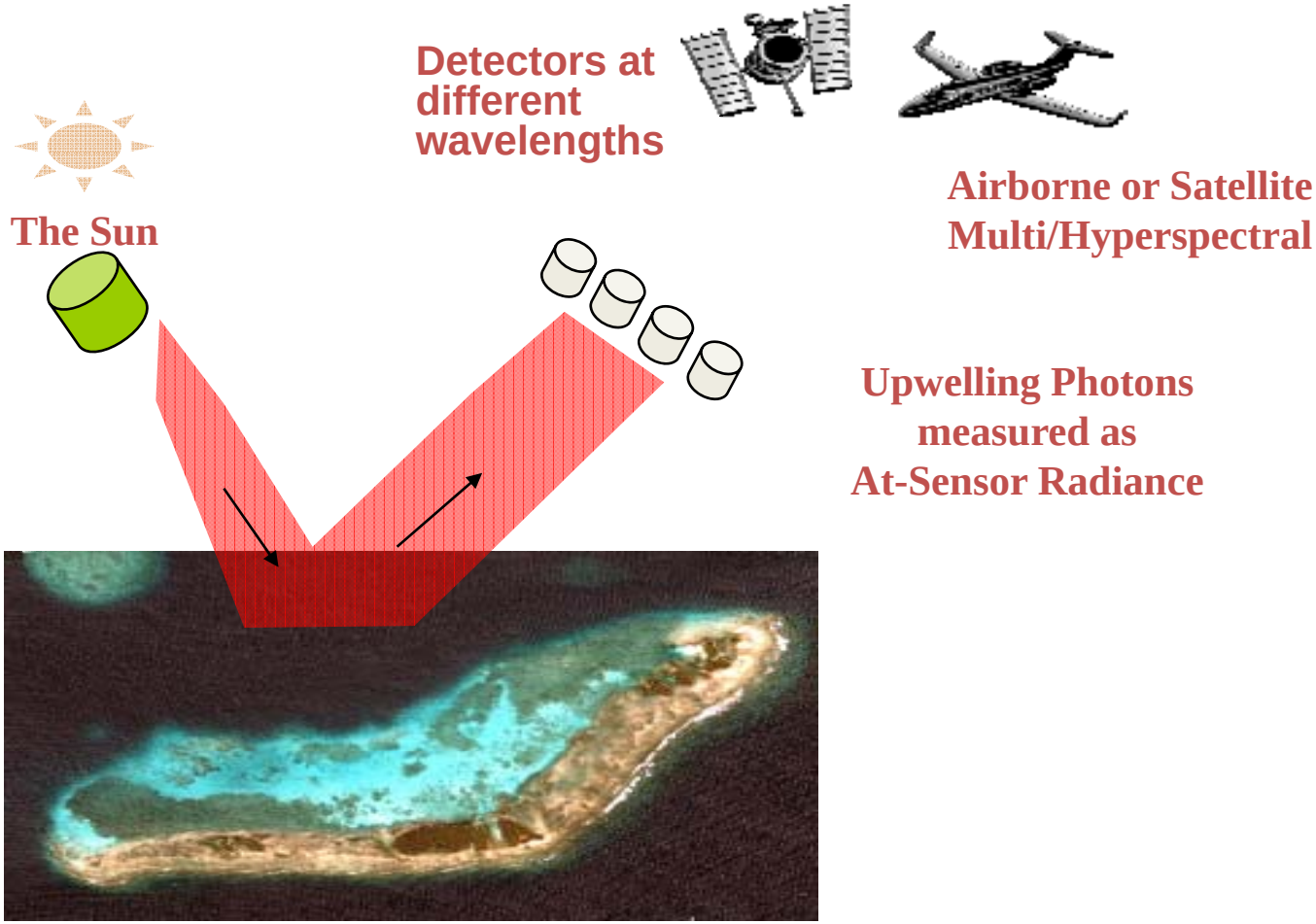
Advancements in Science and Engineering Symposium

March 26, 2009

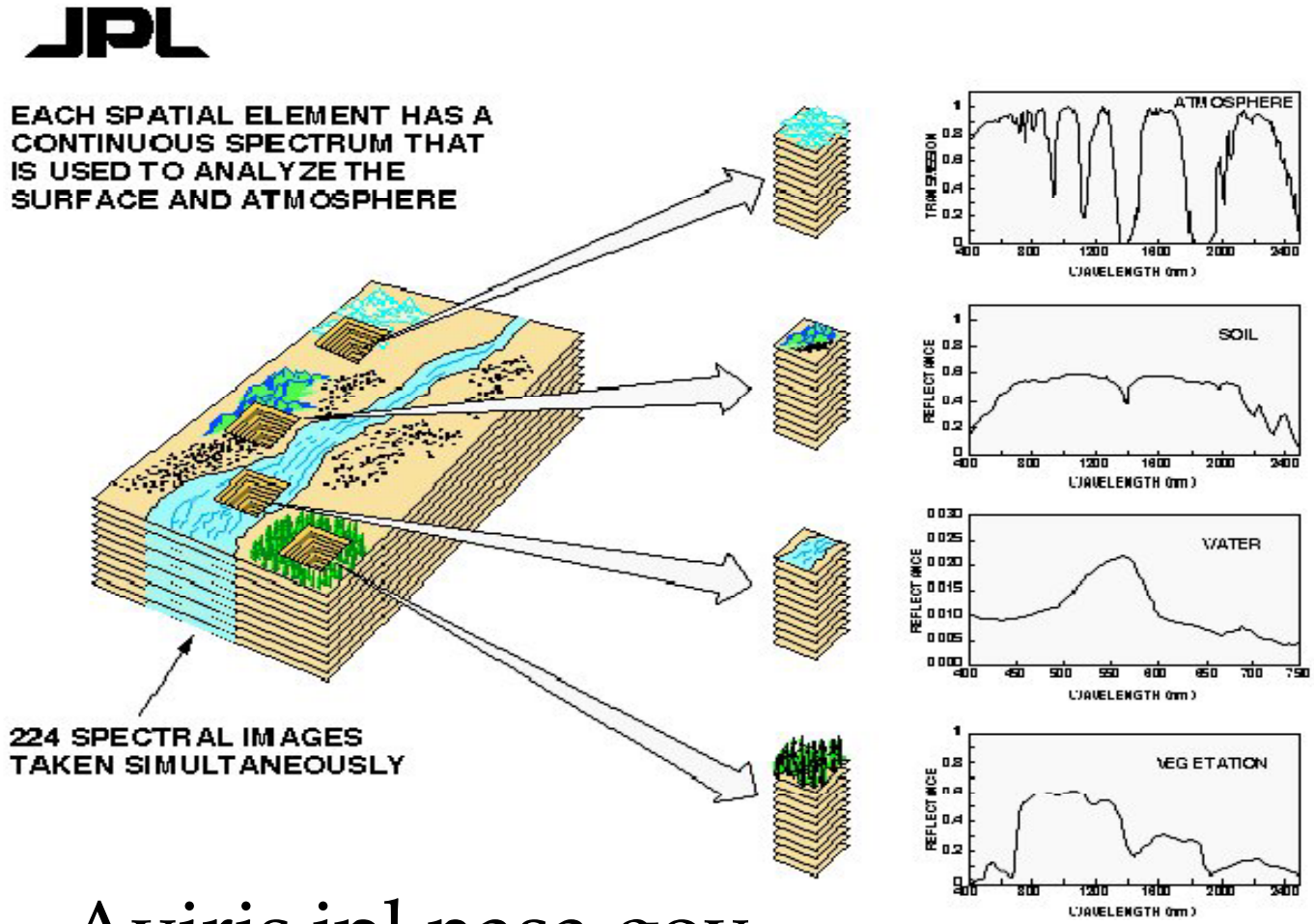
Overview

- Background
- Previous Work
- Problem Statement
- PMF
- Approach
- Results
- Conclusion

Background –Spectral Imaging



Imaging Spectrometry



Aviris.jpl.nasa.gov

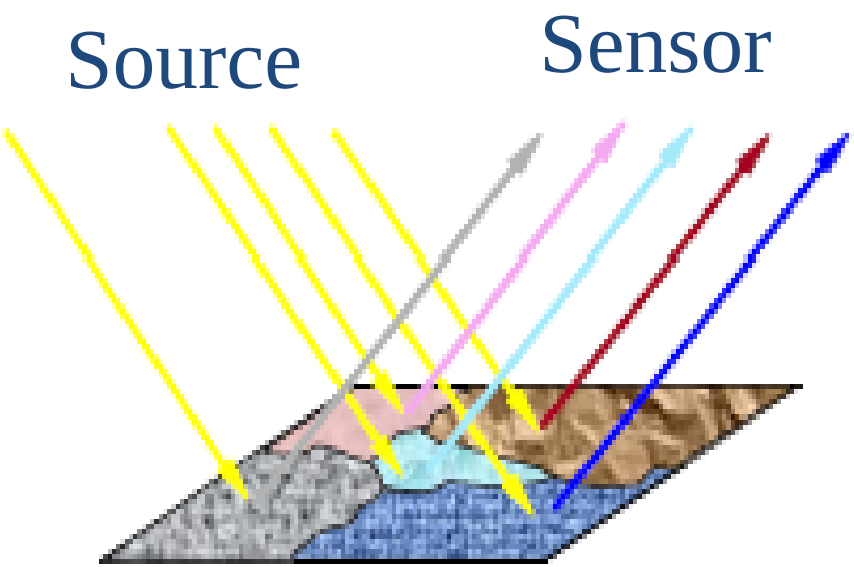
Challenge: Low Spatial Resolution in Hyperspectral Environmental RS

IKONOS Image
Multispectral Sensor
1 meter, 4 bands

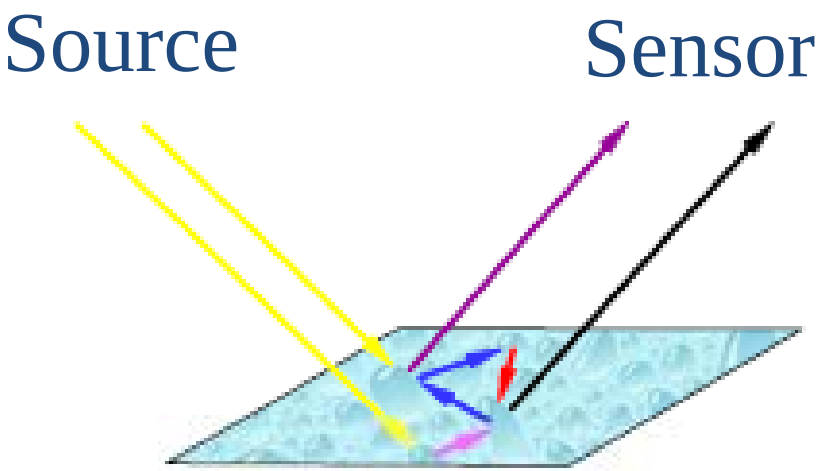
Hyperion Image
Hyperspectral Sensor
30 m, 220 bands, 10nm

Enrique Reef in Parguera, Lajas, PR

Mixing Models

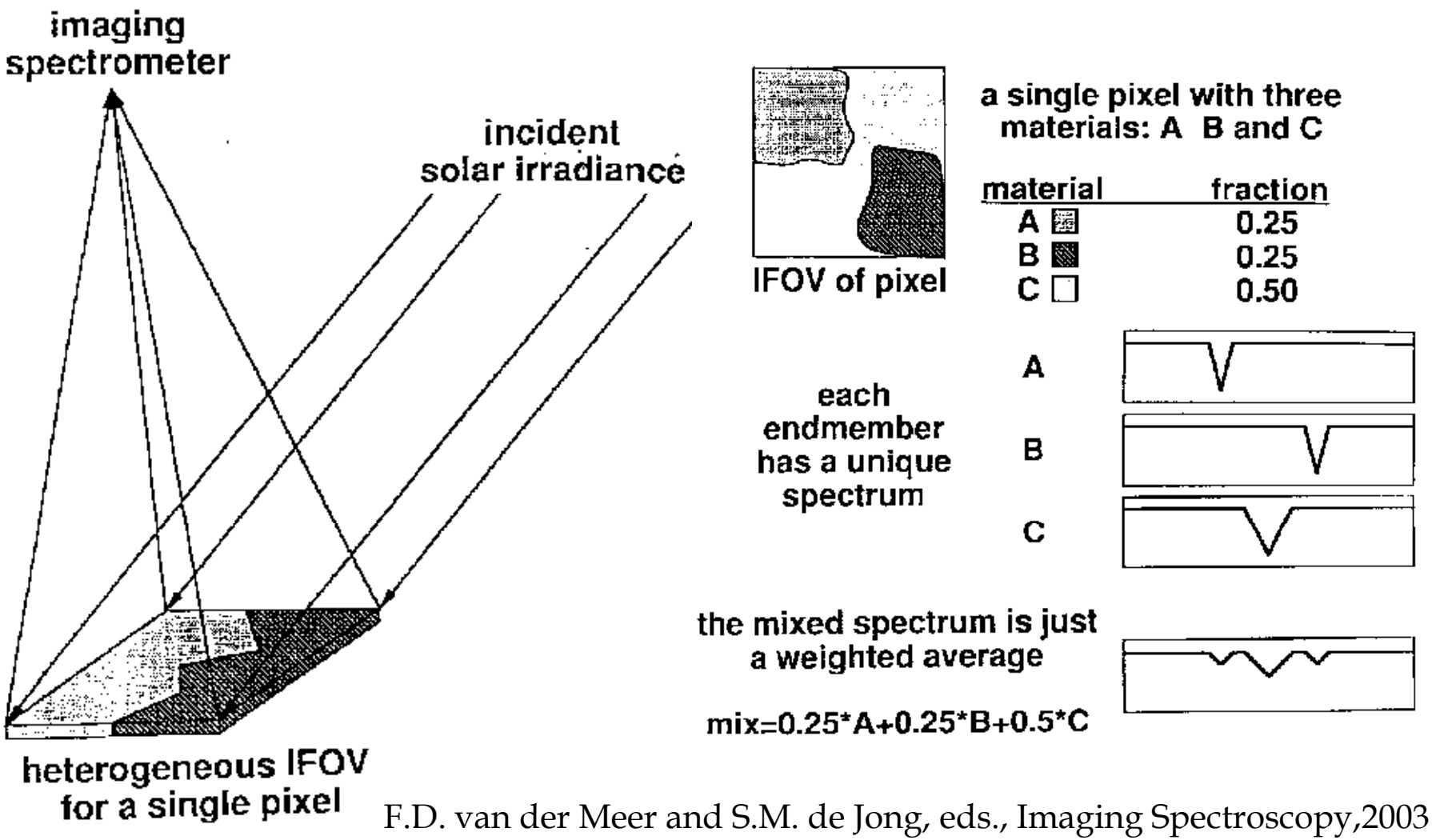


Linear Mixing



Nonlinear Mixing

Linear Mixing Model: Standard for Land Surface



F.D. van der Meer and S.M. de Jong, eds., Imaging Spectroscopy,2003

Background-Linear Mixing Model

$$x_j = \sum_{i=1}^{p} s_i a_{ij} + w_j = Sa_j + w_j \quad , j = 1,2,.....,n$$

where

- $\mathbf{x}_j \in \Re_+^m$ is the measured spectrum of the j-th pixel of interest,
- $\mathbf{s_i} \in \Re_+^m$ is the spectral signature of the i-th endmember,
- $\mathbf{w}_j \sim \mathbf{N(0,\sigma^2 I)}$ is the measurement noise,
- m is the number of spectral channels, and
- P is the number of endmembers.
- a_{ij} is the fractional abundance of the i-th endmember in the j-th pixel, $0 \le a_{ij} \le 1$ and $\sum_{i}^{p} a_{ij} = 1$

Background-Unmixing in HSI

$$\mathbf{X} = \mathbf{S}\mathbf{A} + \mathbf{W}$$

$$\mathbf{X} = [\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_N],\ S = [\mathbf{s}_1\ \mathbf{s}_2\ \ldots\ \mathbf{s}_P],\ \mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2\ \ldots \mathbf{a}_N]$$

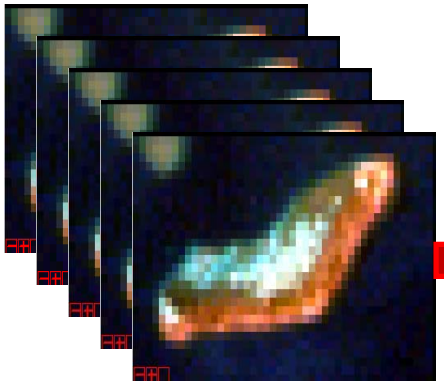
where

N is the number of pixels in the image

P is the number of endmembers

Challenge -- to estimate both A and S, given X

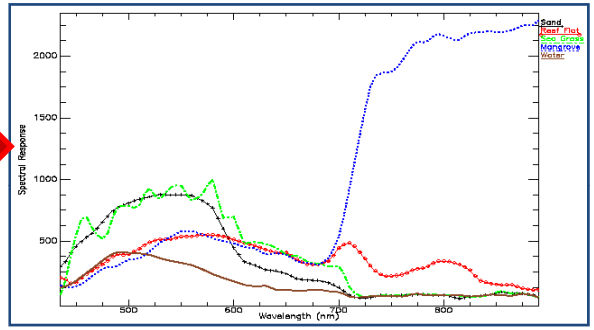
Unmixing Process



Hyperspectral Image

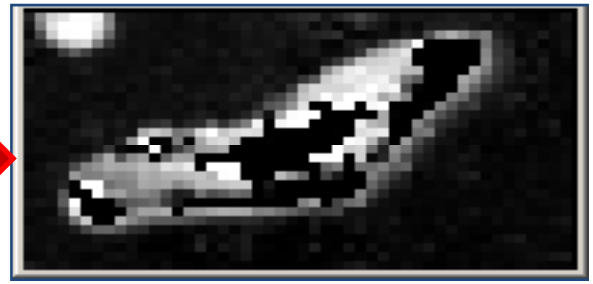
Unmixing

$\hat{p}, \hat{\mathbf{S}}$



Endmember Signatures

$\hat{\mathbf{A}}$



Abundance Maps

Standard Unmixing using a Multistage Approach

- Linear dimension
- Virtual dimensionality
- Positive Rank

- PPI
- N-FINDR

- LS
- NNLS
- NNSTOLS

Previous Work: Unmixing Approaches

- **Two Stage**
  - Endmember determination
    - Supervised or unsupervised
  - Abundance Estimation
- **Single Stage**
  - Estimating endmembers and abundances together by solving an optimization problem.
    - Unsupervised

Unsupervised Unmixing

$$\left(\hat{\mathbf{A}}, \hat{\mathbf{S}}, \hat{p}\right) = \arg \min_{\substack{a_{ij} \geq 0,\ \mathbf{A}^{\mathrm{T}} \mathbf{1_p} = \mathbf{1_N} \\ \mathbf{S} \geq \mathbf{0}}} \| \mathbf{X} - \mathbf{SA} \|_F^2$$

Endmembers, **S**, and their abundances, **A**, are extracted simultaneously from the data by solving this optimization problem.

Positive Matrix Factorization: Definition

- Given a nonnegative matrix $\mathbf{X} \in \Re^{M \times N}$
  - find a positive integer P < min{M, N}, and
  - nonnegative matrices $\mathbf{S} \in \Re^{M \times P}$ and $\mathbf{A} \in \Re^{P \times N}$ such that

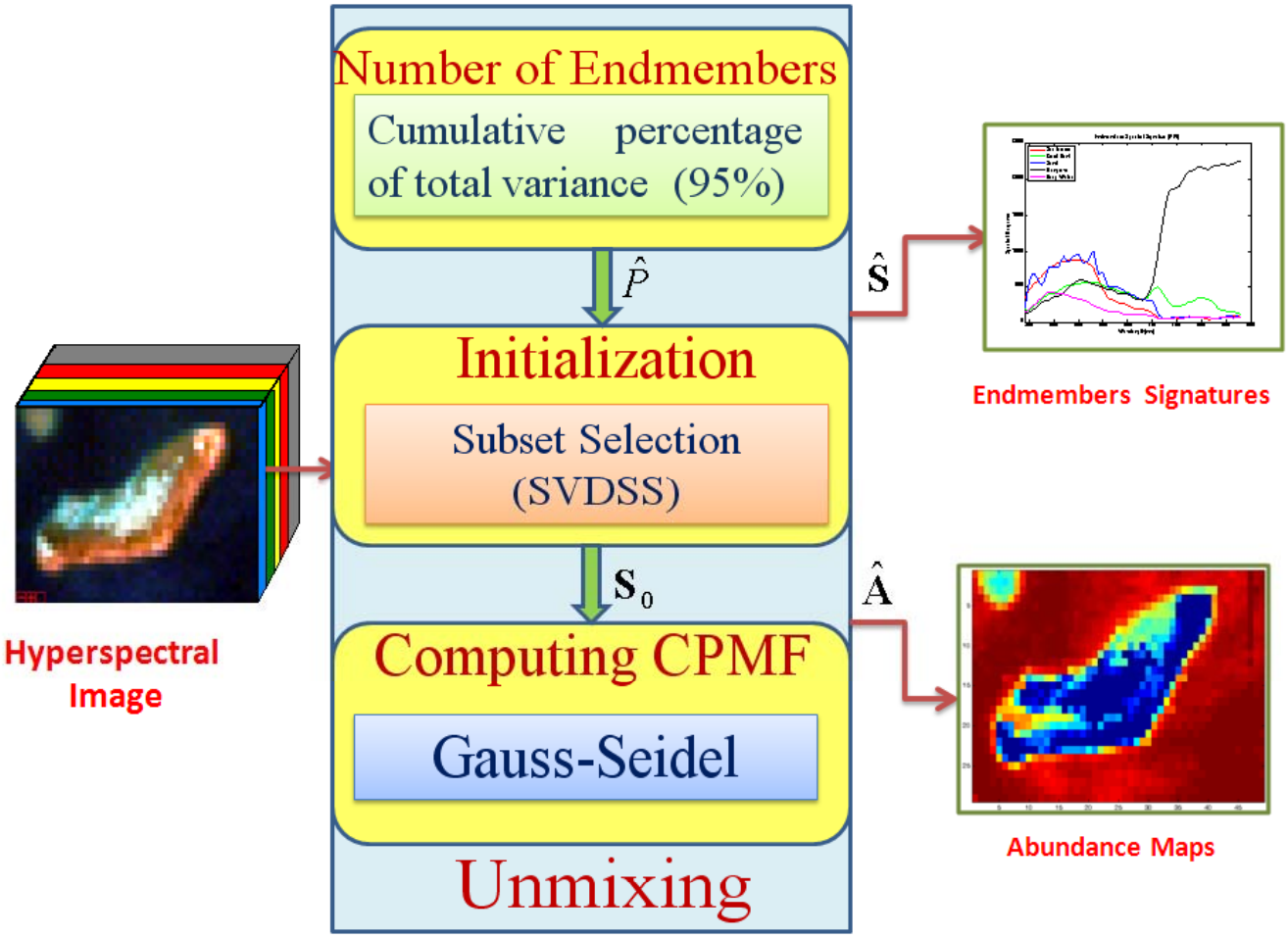
$$\mathbf{X} = \mathbf{SA}$$

- Approximated PMF

$$\mathbf{X} \approx \mathbf{SA}$$

Unsupervised unmixing is a constrained PMF problem

Unsupervised Unmixing Algorithm



Number of Endmembers

- Cumulative Percentage of Total Variation

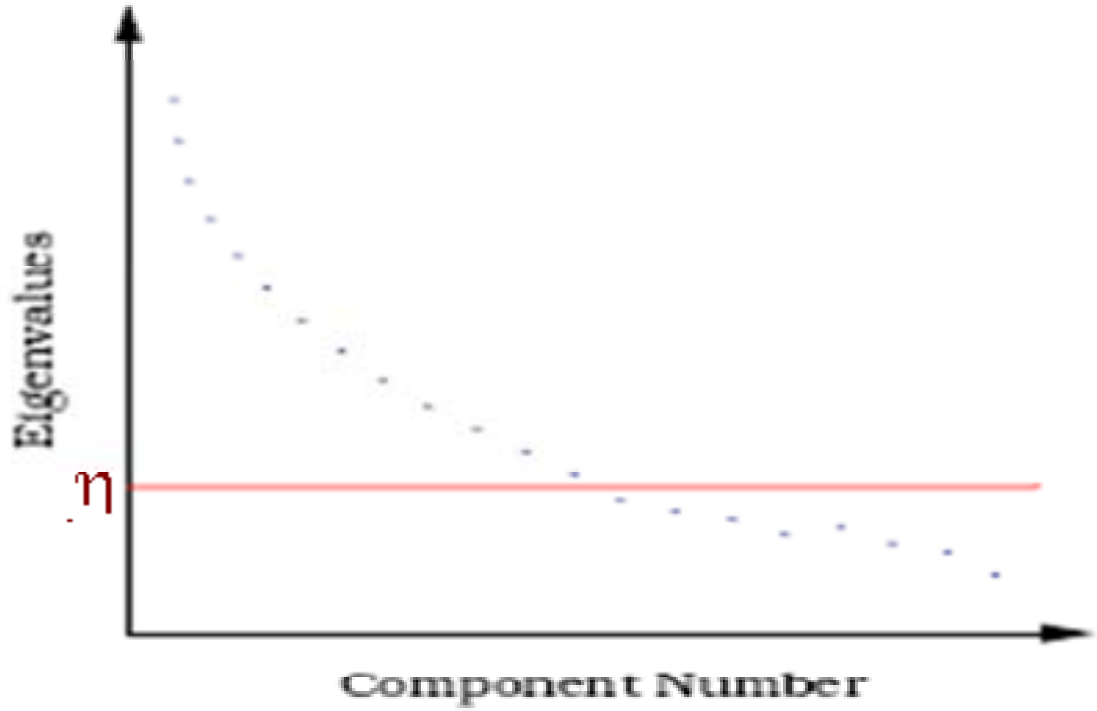
$$\% \; Var = 100 \times \left(\frac{\sum_{i=1}^{\hat{P}_{CV}} \lambda_i}{\sum_{i=1}^{M} \lambda_i} \right)$$

- Continuous Significant Dimension

$$\hat{P}_{CSD} = \sum_{j=1}^{M} \min(\lambda_j, 1)$$

Number of Endmembers

- Size of Variance (Kaiser-Guttman)



Correlation matrix eigenvalues are used

Initialization

- Singular Value Subset Selection (SVDSS)

  - Selects the most independent pixels in the image.

    (Masalmah07)

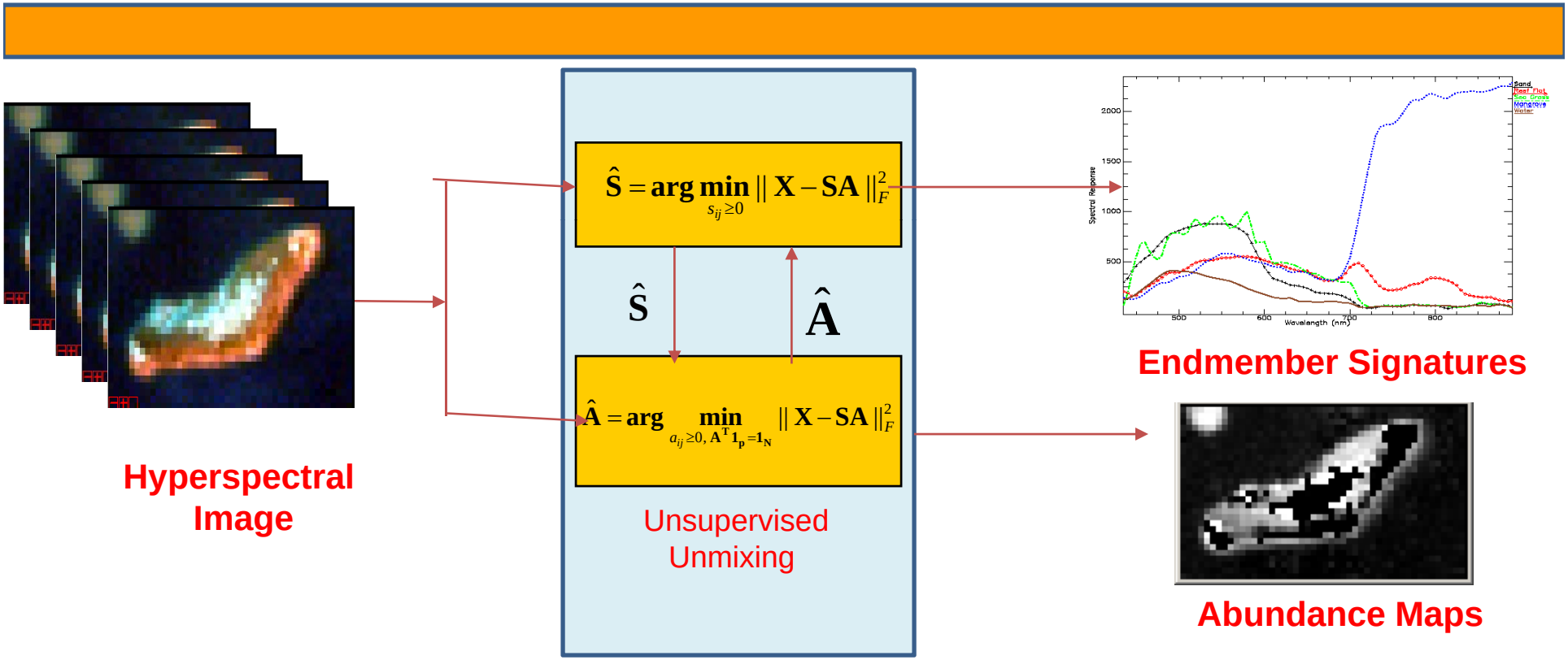
Gauss-Seidel Method

$$\hat{\mathbf{A}} = \arg \min_{a_{ij} \geq 0,\, \mathbf{A}^{\mathbf{T}} \mathbf{1}_{\mathbf{p}} = \mathbf{1}_{\mathbf{N}}} \| \mathbf{X} - \mathbf{SA} \|_F^2 = \mathbf{T}_1(\mathbf{S})$$

$$\hat{\mathbf{S}} = \arg \min_{s_{ij} \geq 0} \| \mathbf{X} - \mathbf{SA} \|_F^2 = \mathbf{T}_2(\mathbf{A})$$

Each sub problem is a linear problem

Unsupervised Unmixing-Two-stage



Stopping Criteria

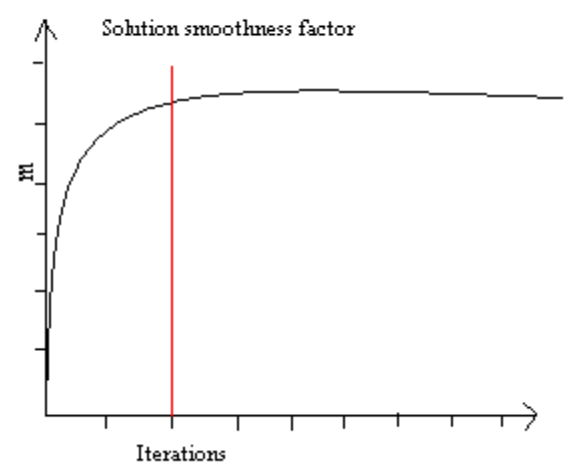
- Maximum Iteration
- Solution successive iteration difference

$$\| \mathbf{A}^{k+1} - \mathbf{A}^{k} \|_F \leq \varepsilon_A$$

$$\| \mathbf{S}^{k+1} - \mathbf{S}^{k} \|_F \leq \varepsilon_S$$

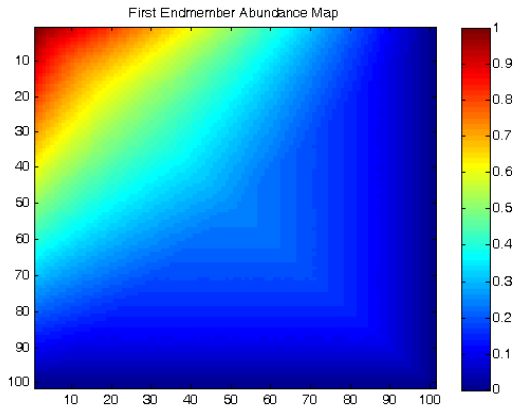
- Solution Smoothing Factor

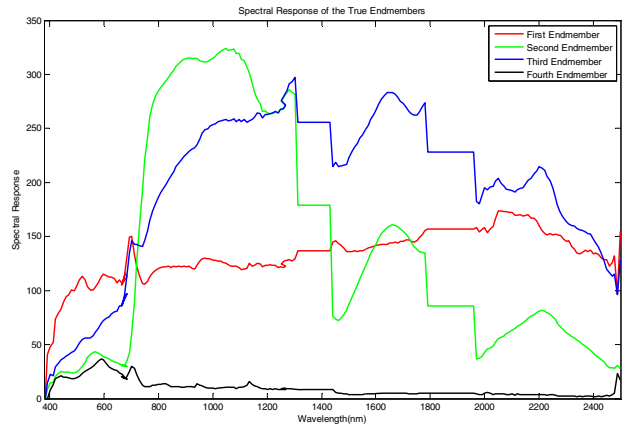
$$m_s = \log(17.09\sigma_s)$$

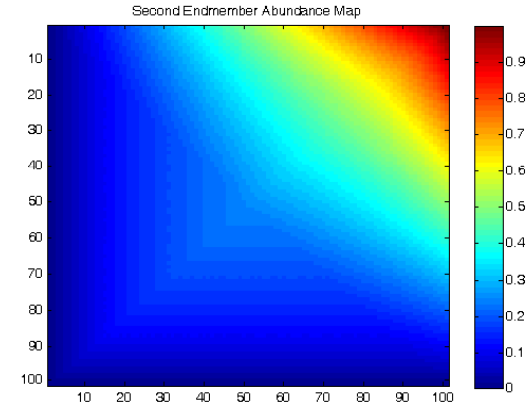


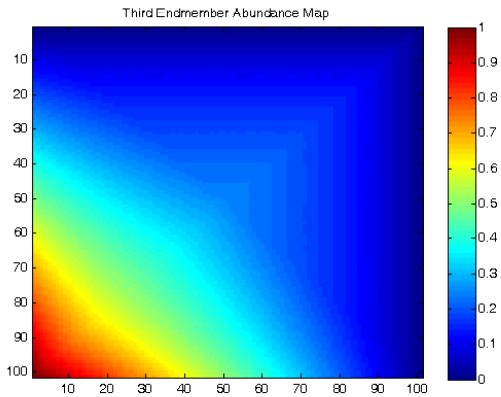
Experimental Results

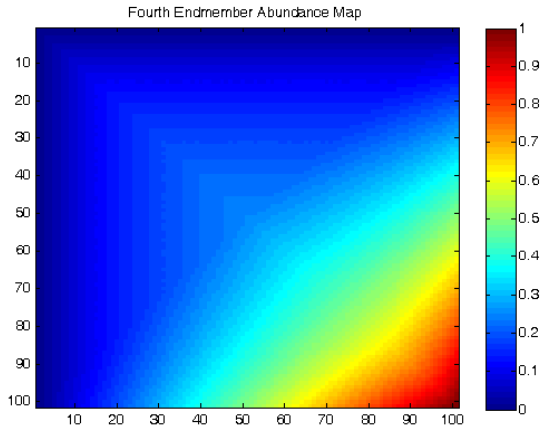
Simulated Data -Cube











Quality of Estimates Metrics (For Simulated Data Only)

- Percentage Error of Euclidean Distance

$$error(s_i, \hat{s}_i) = \frac{\| \hat{s}_i - s_i \|}{\| s_i \|}$$

- Spectral Similarity (spectral Angle)

$$\theta(s_i, \hat{s}_i) = \cos^{-1}\left(\frac{\hat{s}_i^T s_i}{\| \hat{s}_i \| . \| s_i \|} \right)$$

$\hat{s}_i$ is the i - th estimated endmember spectrum

s_i is the i - th original endmember spectrum

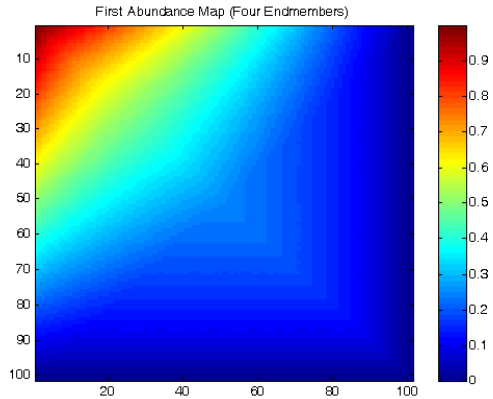
Measure of Fit Metrics

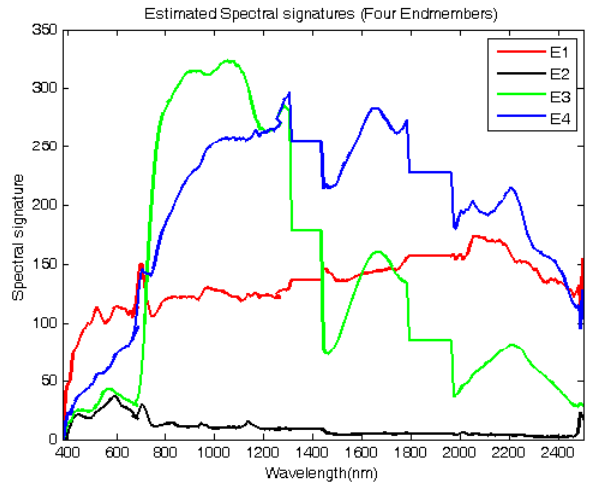
- Sample Correlation (R-squared)
- $$R_j^2 = 1 - \frac{\sum_{i=1}^{M} r_{ij}^2}{\sum_{i=1}^{M} x_{ij}^2}$$ where $$r_j = x_j - \hat{x}_j$$
- Root-Mean-Square(RMS)

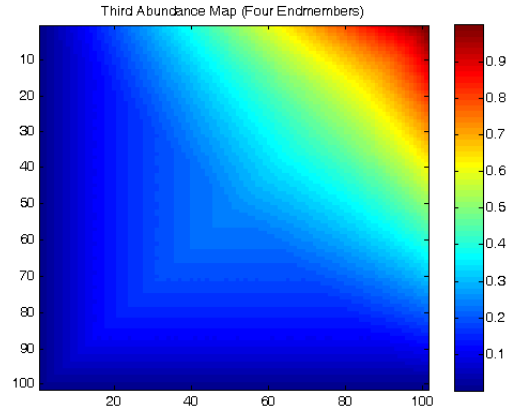
$$RMS_j = \sqrt{\frac{1}{M}\sum_{i=1}^{M} r_{ij}^2}$$

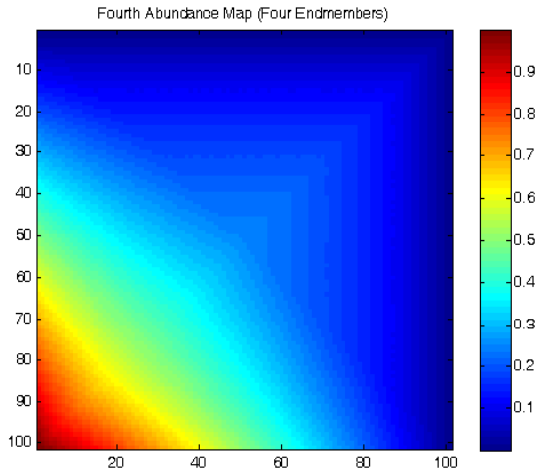
Results- Cube Simulated Data

(Four Endmembers)

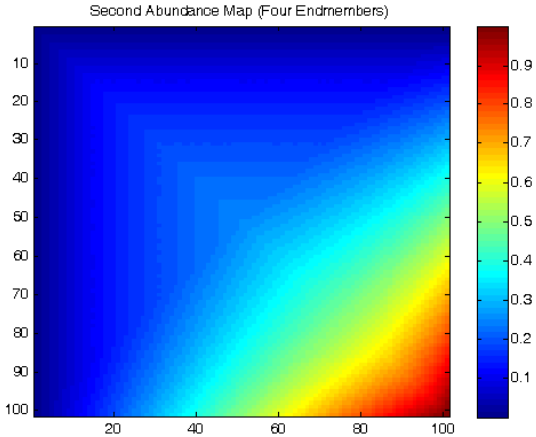




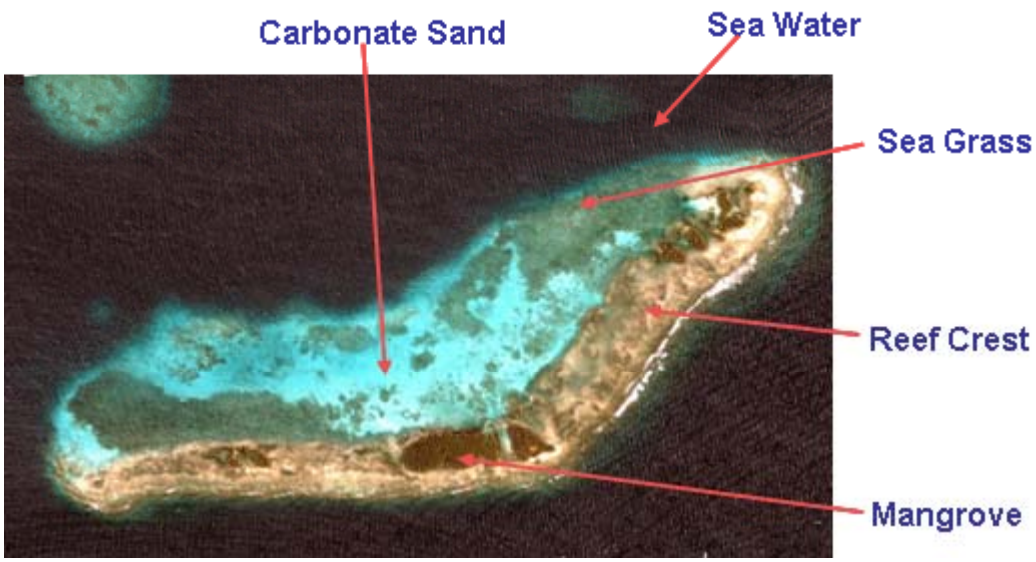




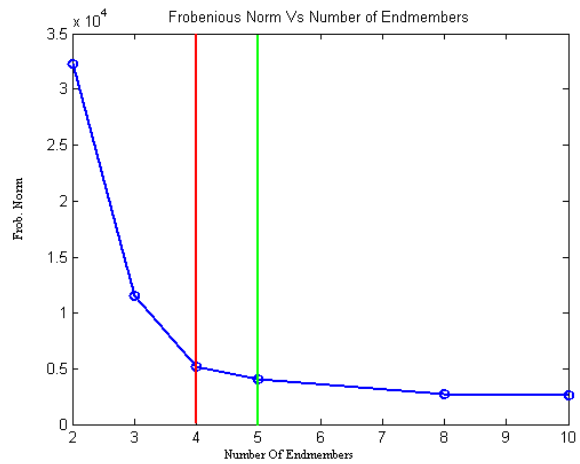
Approach	SVD	KG	CSD
Dimension	4	3	3



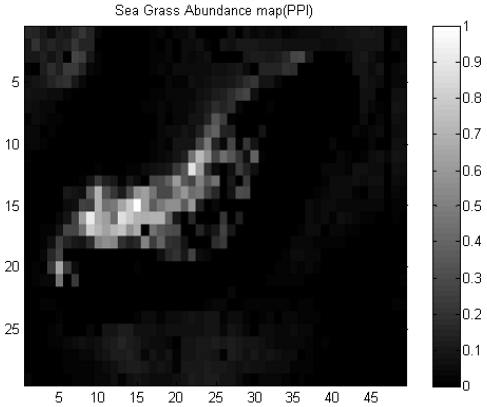
Results-Real data (Enrique Reef)



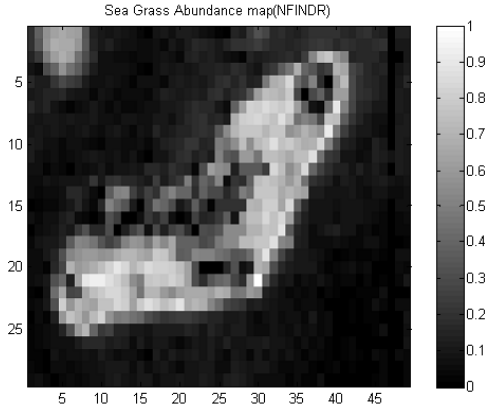
Approach	SVD	KG	CSD
Dimension	5	4	5



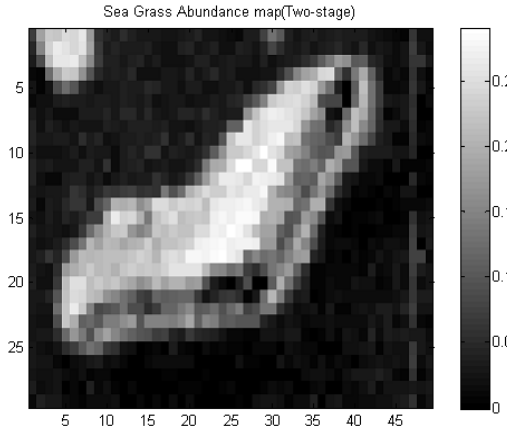
Results- Real Data (Enrique Reef)



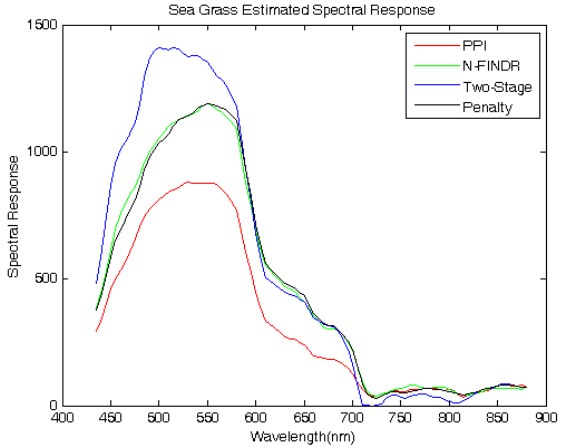
PPI



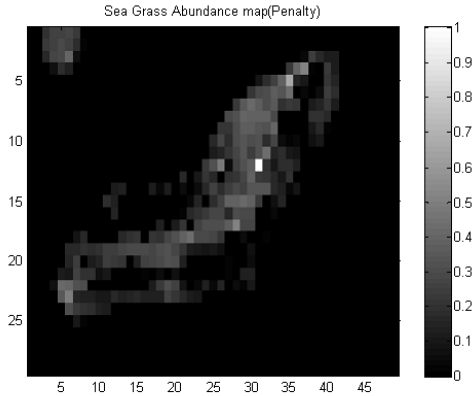
N-FINDR



Gauss-Seidel

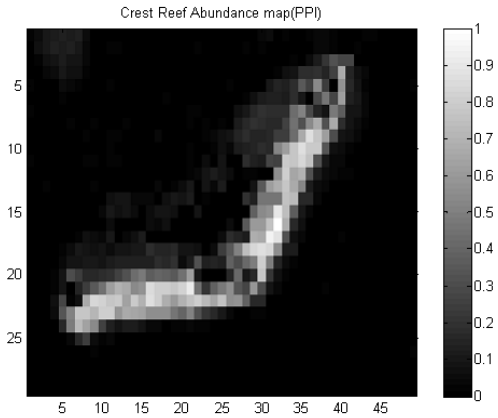


Sea Grass

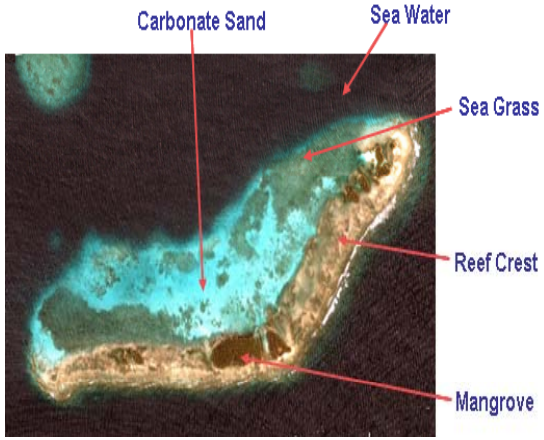


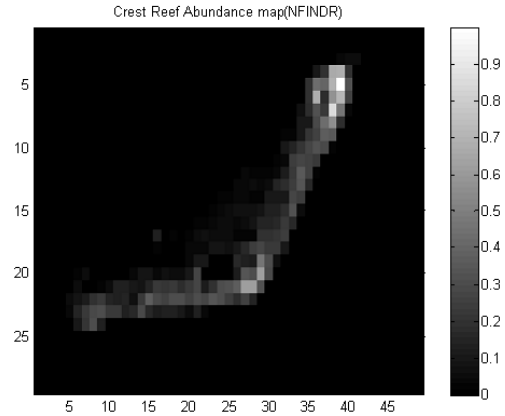
Penalty

Results- Real Data (Enrique Reef)

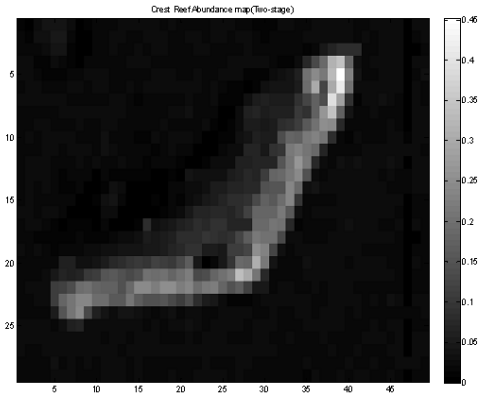


PPI

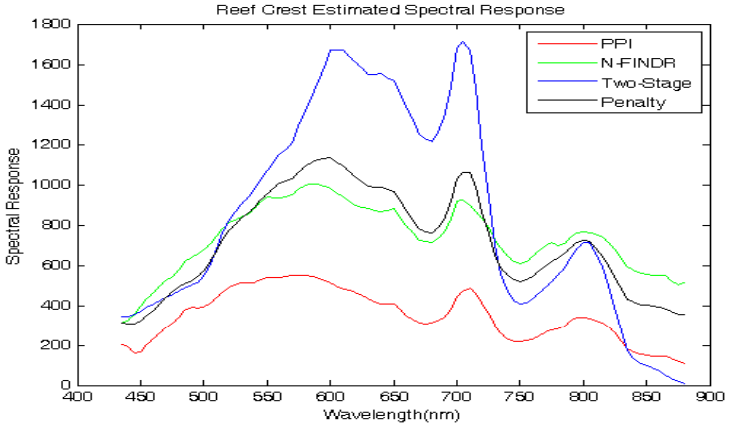




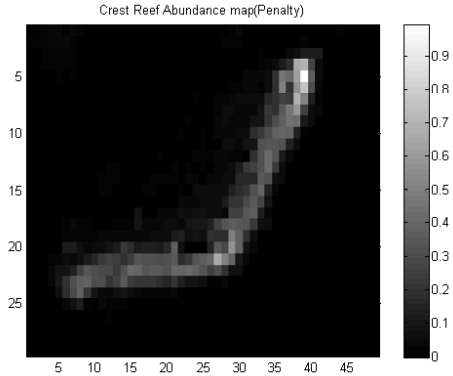
N-FINDR



Gauss-Seidel

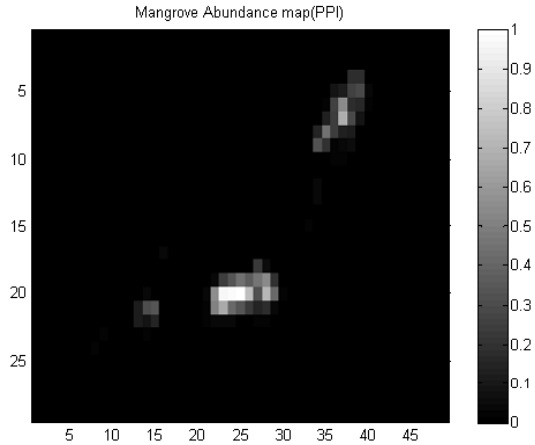


Reef Crest

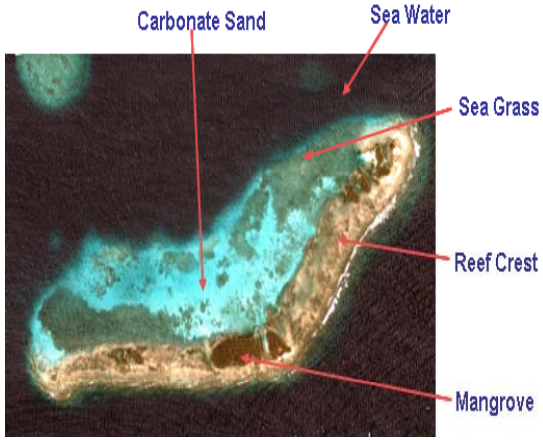


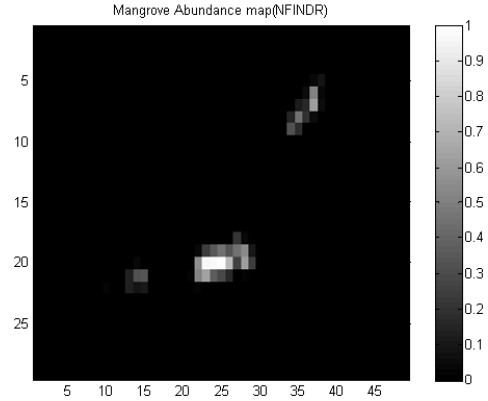
Penalty

Results- Real Data (Enrique Reef)

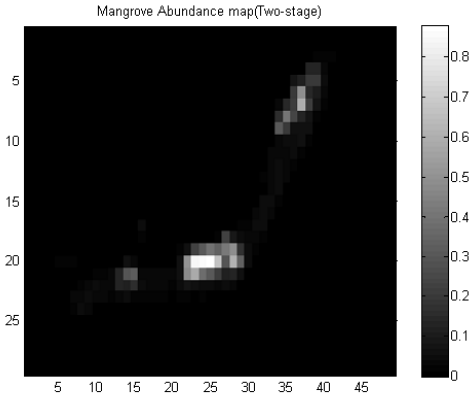


PPI

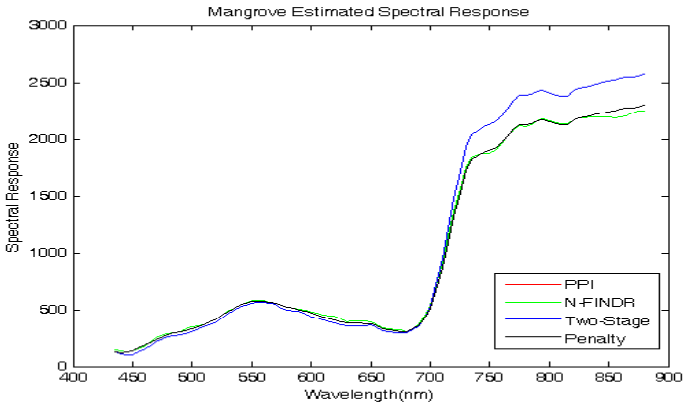




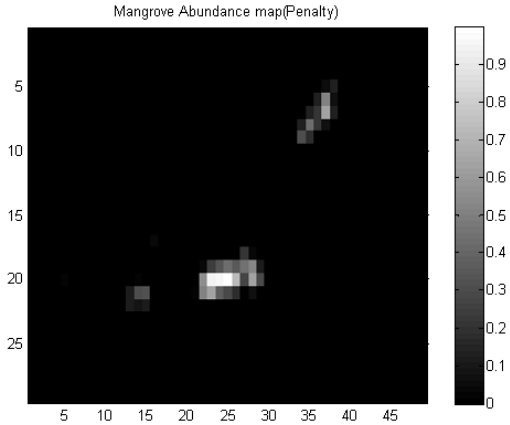
N-FINDR



Gauss-Seidel

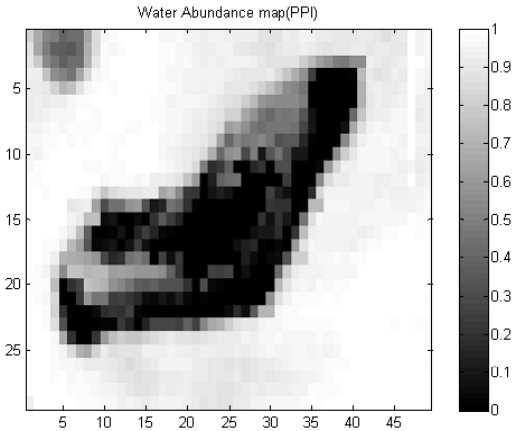


Mangrove

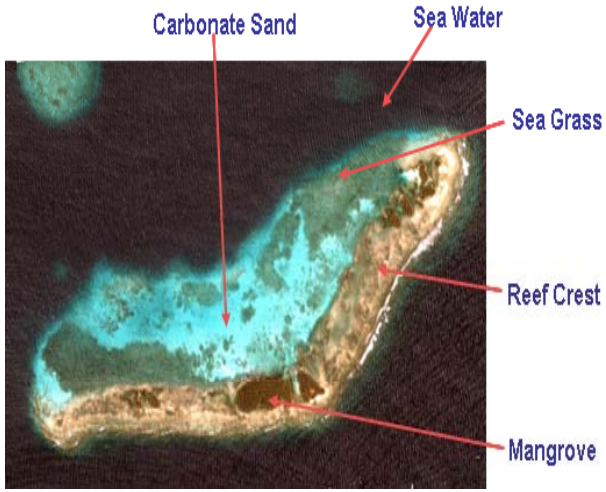


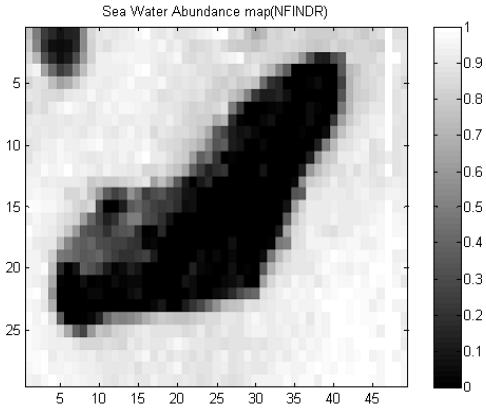
Penalty

Results- Real Data (Enrique Reef)

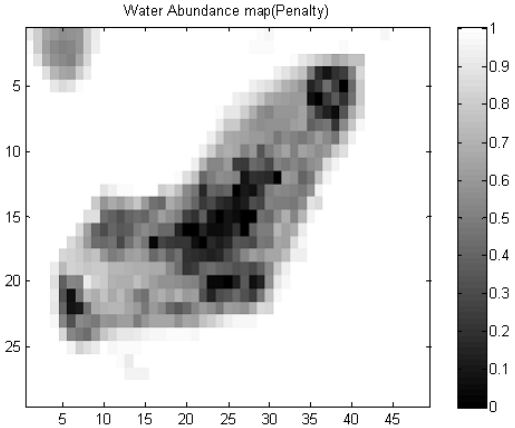


PPI

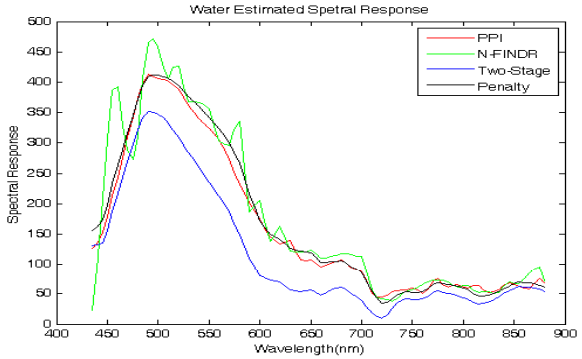




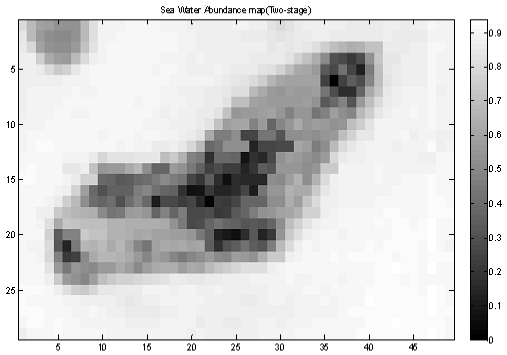
N-FINDR



Gauss-Seidel

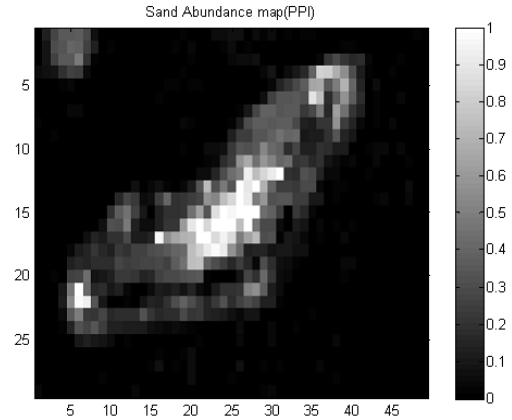


Water

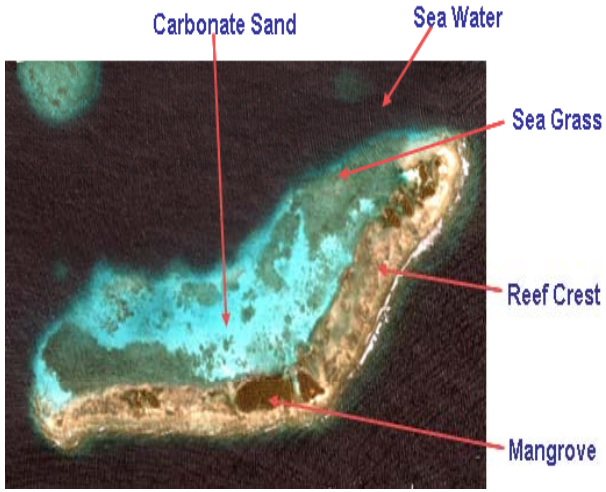


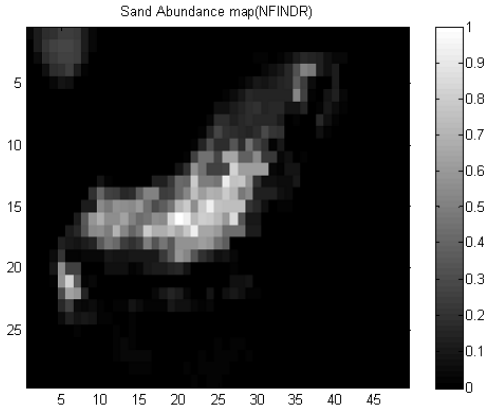
Penalty

Results- Real Data (Enrique Reef)

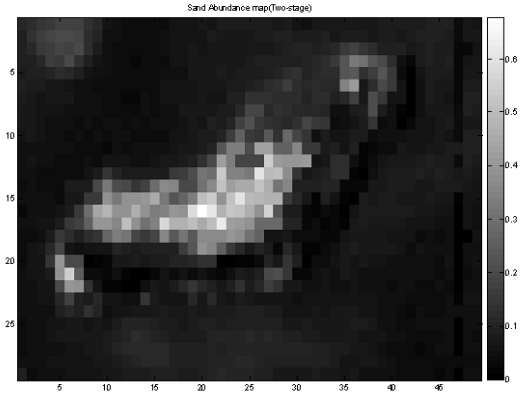


PPI

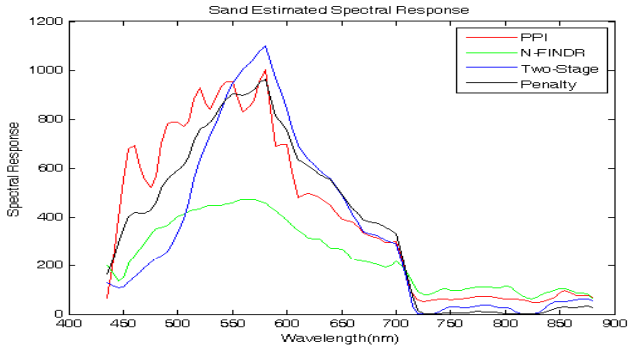




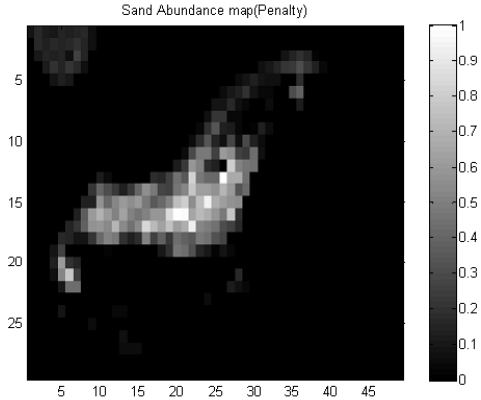
N-FINDR



Gauss-Seidel



Sand



Penalty

Results- Real Data

(Enrique Reef)

Approach	Min R^2 value	Max R^2 value	Average R^2 value	Standard deviation
PPI	0.8820	1.000	0.9938	0.0086
N-FINDR	0.8590	1.000	0.9780	0.0170
Two-stage	0.9773	0.9999	0.9983	0.0016
Penalty	0.9409	0.9999	0.9954	0.0054

Approach	Min rms value	Max rms value	Average rms	Standard deviation
PPI	0	259.05	21.06	21.25
N-FINDR	0	78.03	32.49	9.17
Two-stage	4.69	49.73	10.10	4.35
Penalty	5.85	100.42	15.92	7.40

Conclusions

- The developed algorithm solved the unsupervised unmixing problem.
- Results obtained from simulated data agree completely with the known results.
- Results obtained from real data were better than other known approaches.
- The developed algorithms were applied to different data sets.

Questions ?