Estimation of Mars surface physical properties from hyperspectral images using the SIR method

Caroline Bernard-Michel, Sylvain Douté, Laurent Gardes and Stéphane Girard

Source: ESA

Laboratoire de Planétologie Grenoble

INRIA RHÔNE-ALPES

Outline

I. Context
   - Hyperspectral image data
   - Inverse problem

II. Dimension reduction
   - PCA
   - SIR

III. Regularization and estimation
   - Zhong et al., 2005
   - Tikhonov

IV. Validation on simulations

V. Application to the south polar cap of Mars

VI. Conclusion and future work

Outline

I. Context
- Hyperspectral image data
- Inverse problem

II. Dimension reduction
- PCA
- SIR

III. Regularization and estimation
- Zhong et al., 2005
- Tikhonov

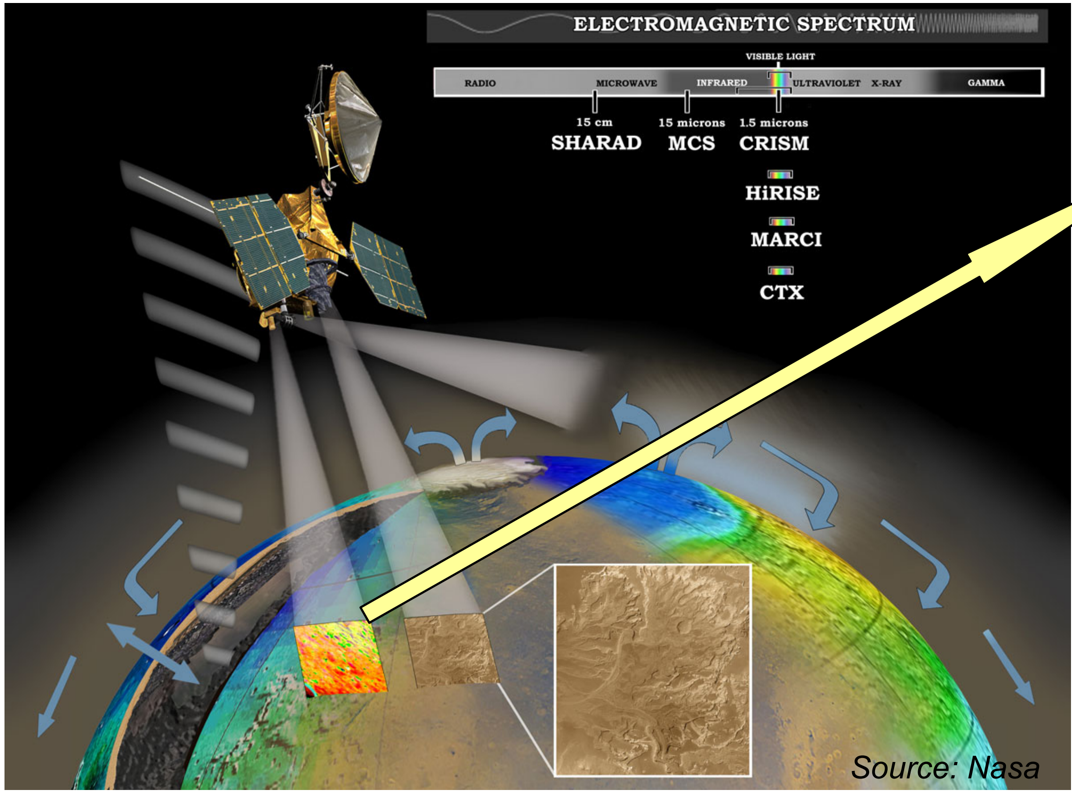
IV. Validation on simulations

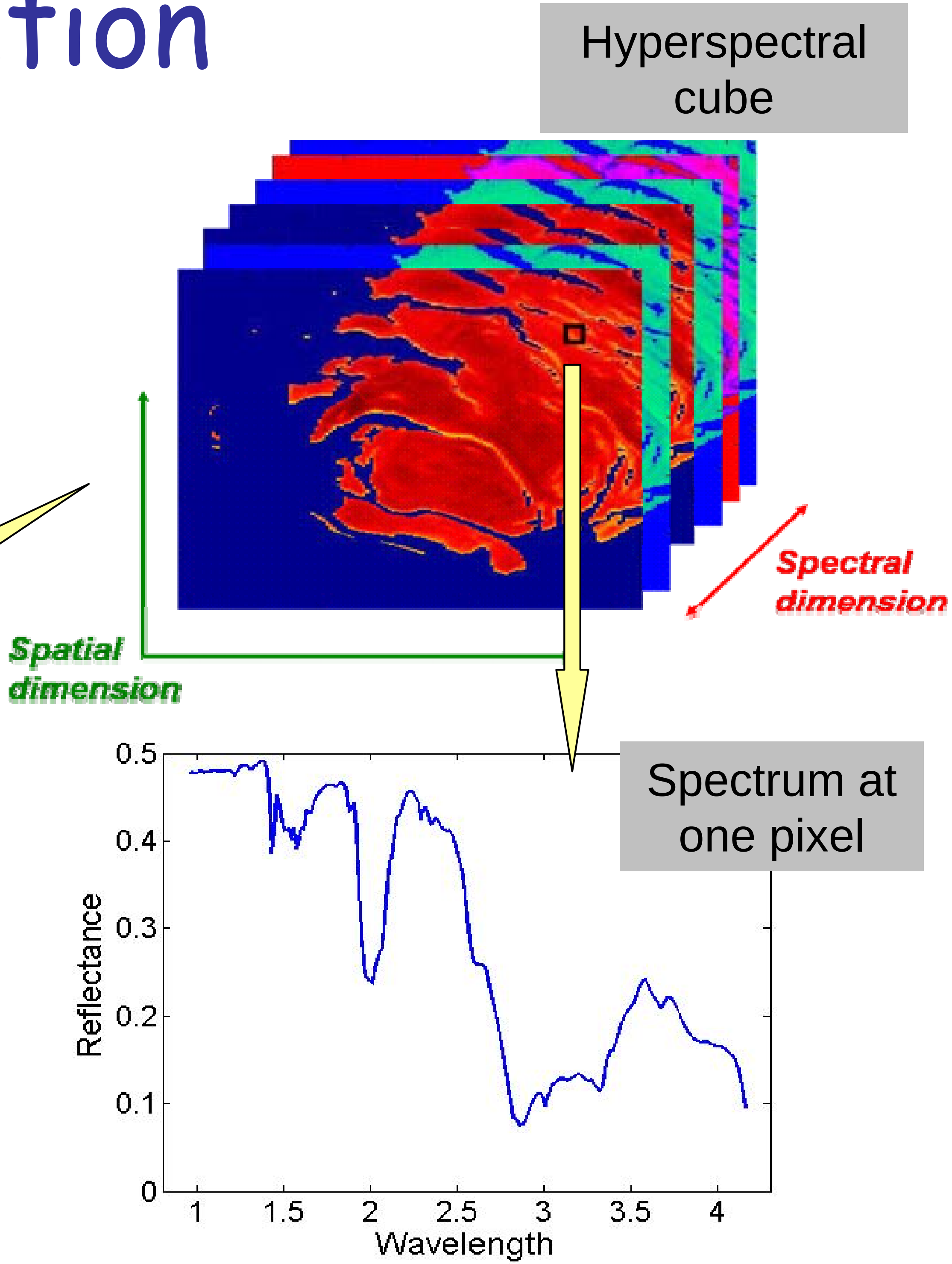
V. Application to the south polar cap of Mars

VI. Conclusion and future work

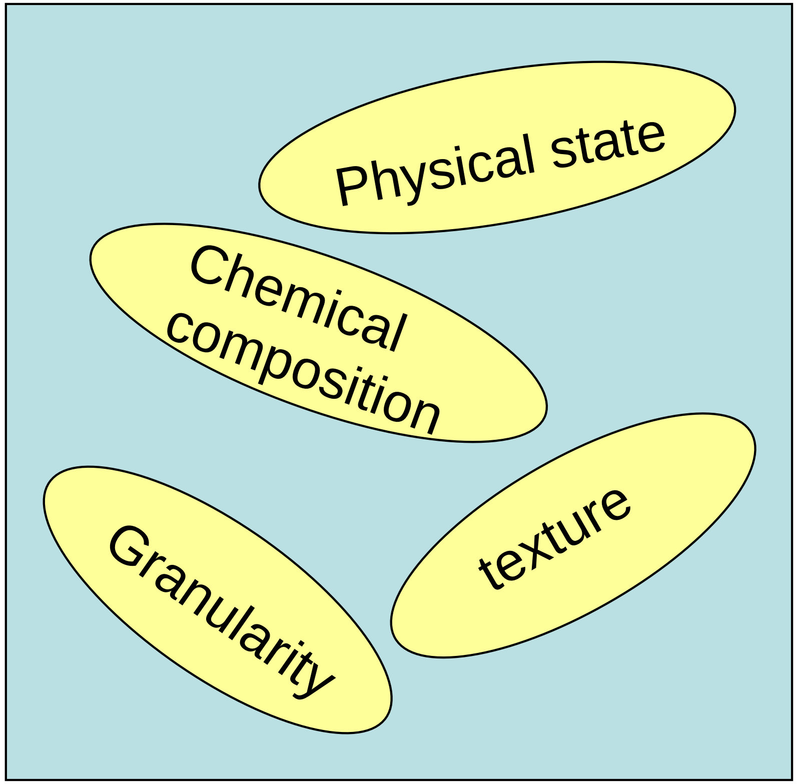
Introduction

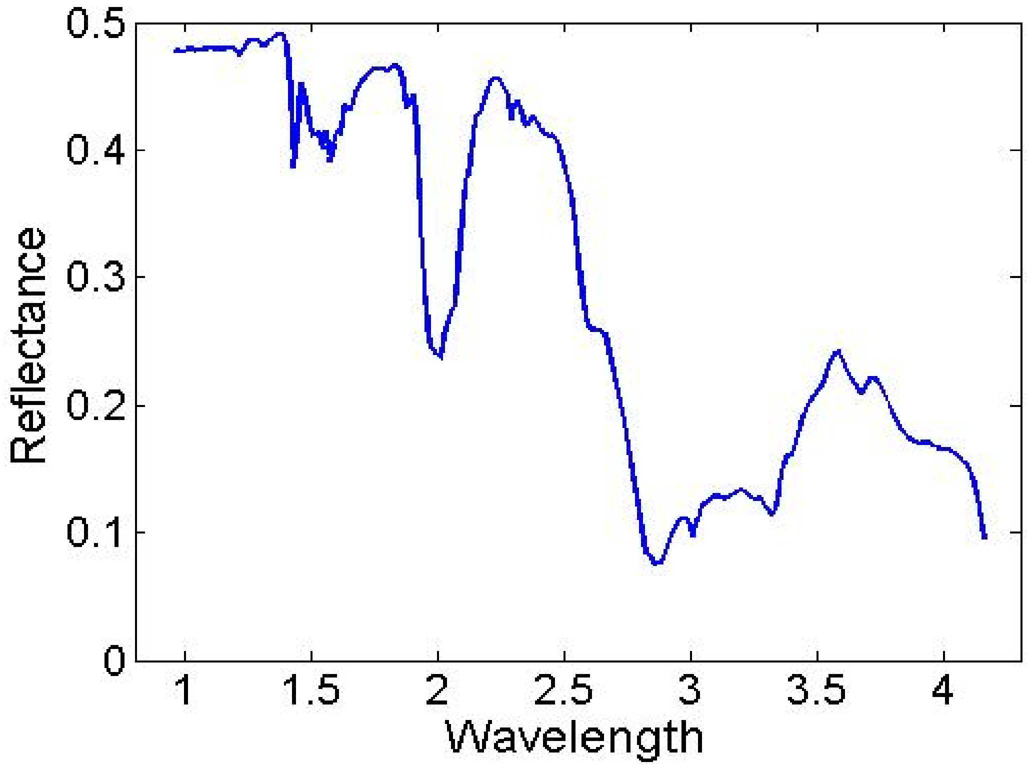
Spectrometer





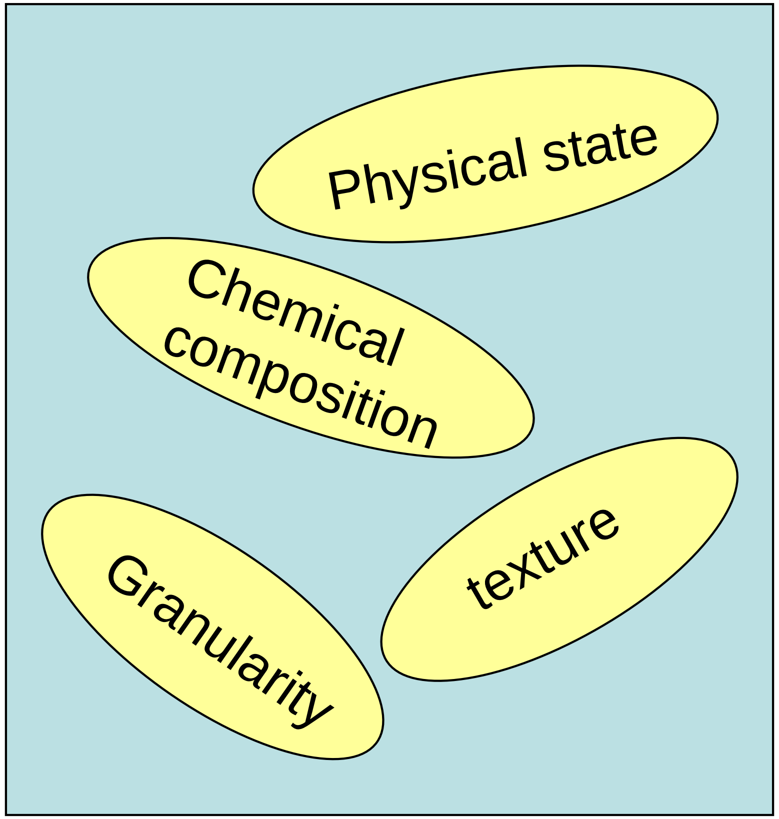
Radiative transfer model

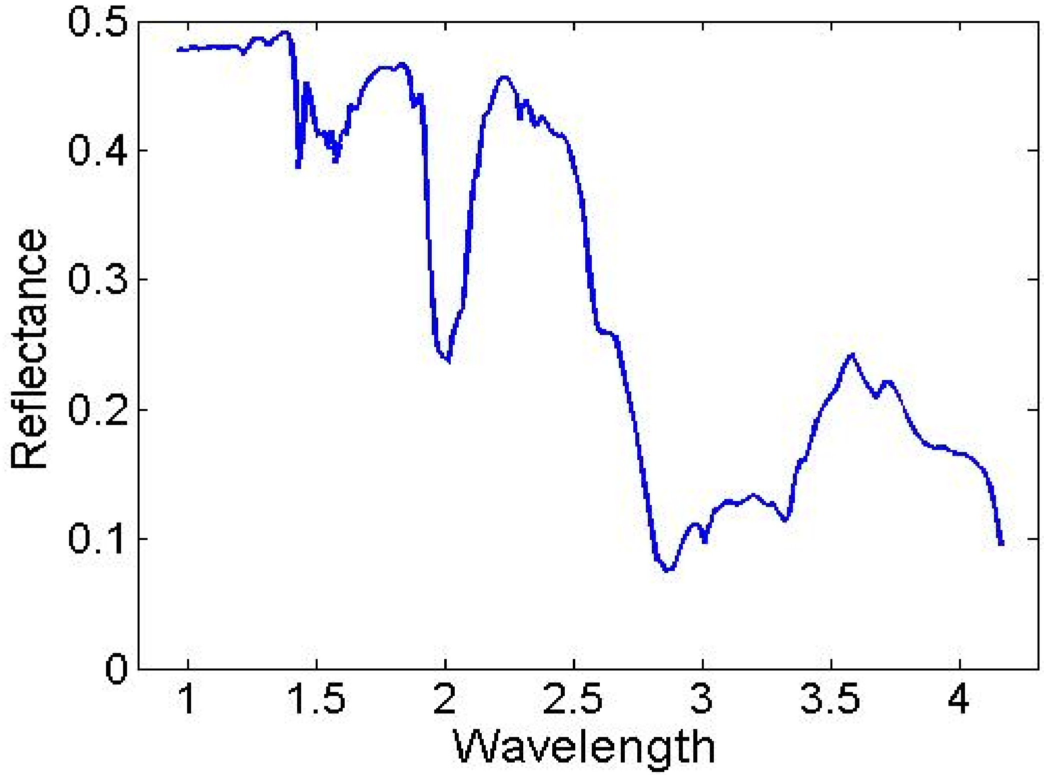




RADIATIVE TRANSFER MODEL: evaluates direct link between parameters and spectra. Allows the construction of a training data

Inverse problem

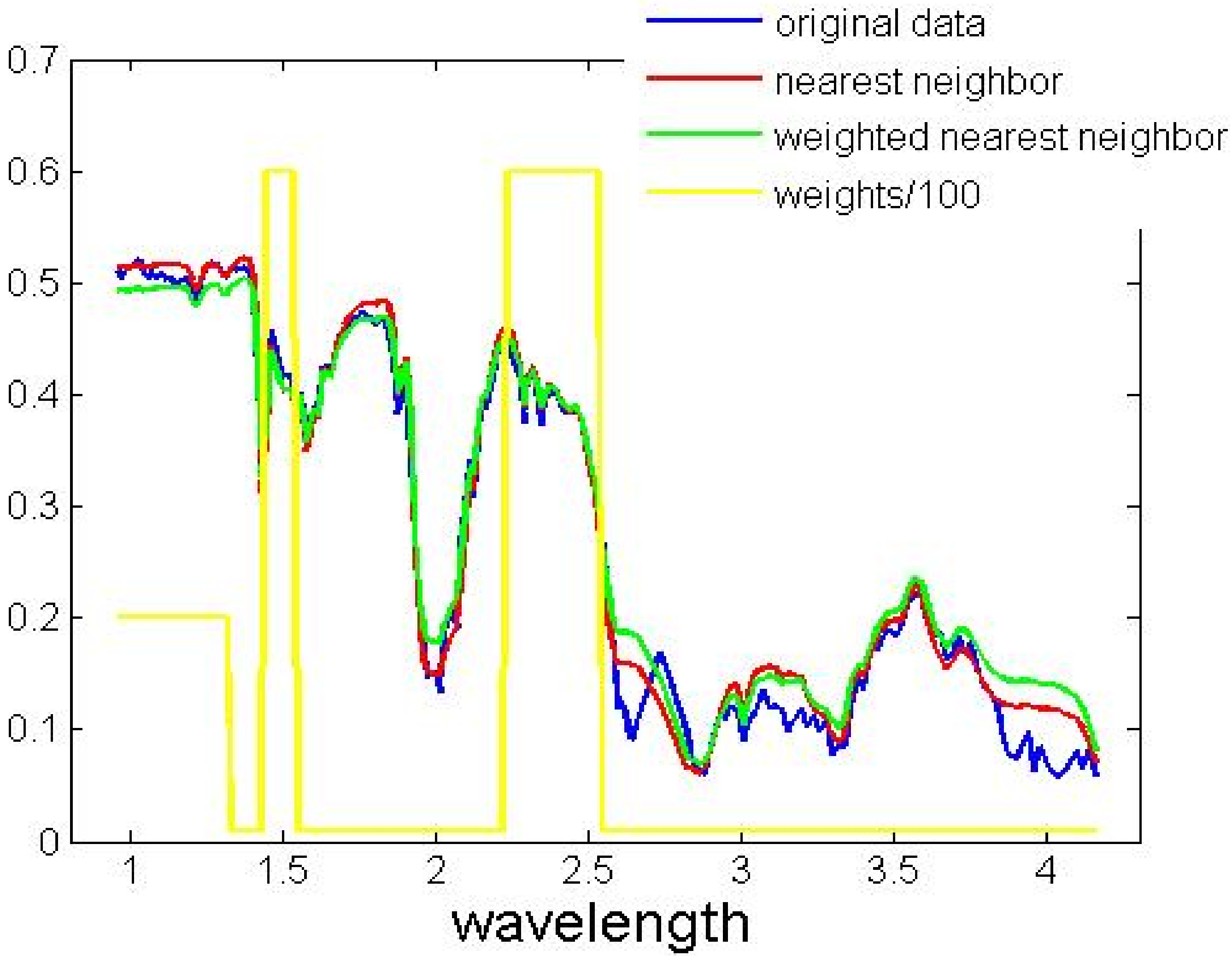




INVERSE PROBLEM: evaluates the properties of atmospheric and surface materials from the spectra

Usual methods

- Nearest neighbor
- Weighted Nearest neighbor



Aim

- To establish functional relationships between:
  - Spectra $x \in \mathbb{R}^p$ (p=184) from Mars Express mission
  - Physical parameter $y \in \mathbb{R}$: proportion of water, proportion of dust, grain size…
  - Construct *f* in order to estimate parameters:

$$
\begin{aligned}
f : \mathbb{R}^p &\longrightarrow \mathbb{R} \\
x &\longrightarrow y
\end{aligned}
$$

Difficulties

- Curse of dimensionality (184 wavelengths): dimension of *x* has to be reduced
- Find projection axis $a \in \mathbb{R}^p$ (here, only the first axis will be retained)
- Instead of estimating f such as $y = f(x)$, we will suppose there exists $g : \mathbb{R} \to \mathbb{R}$ exists such that:

$$y = g(< a, x >, \varepsilon)$$

Outline

I. Context
  - Hyperspectral image data
  - Inverse problem

II. Dimension reduction
  - PCA
  - SIR

III. Regularization and estimation
  - Zhong et al., 2005
  - Tikhonov

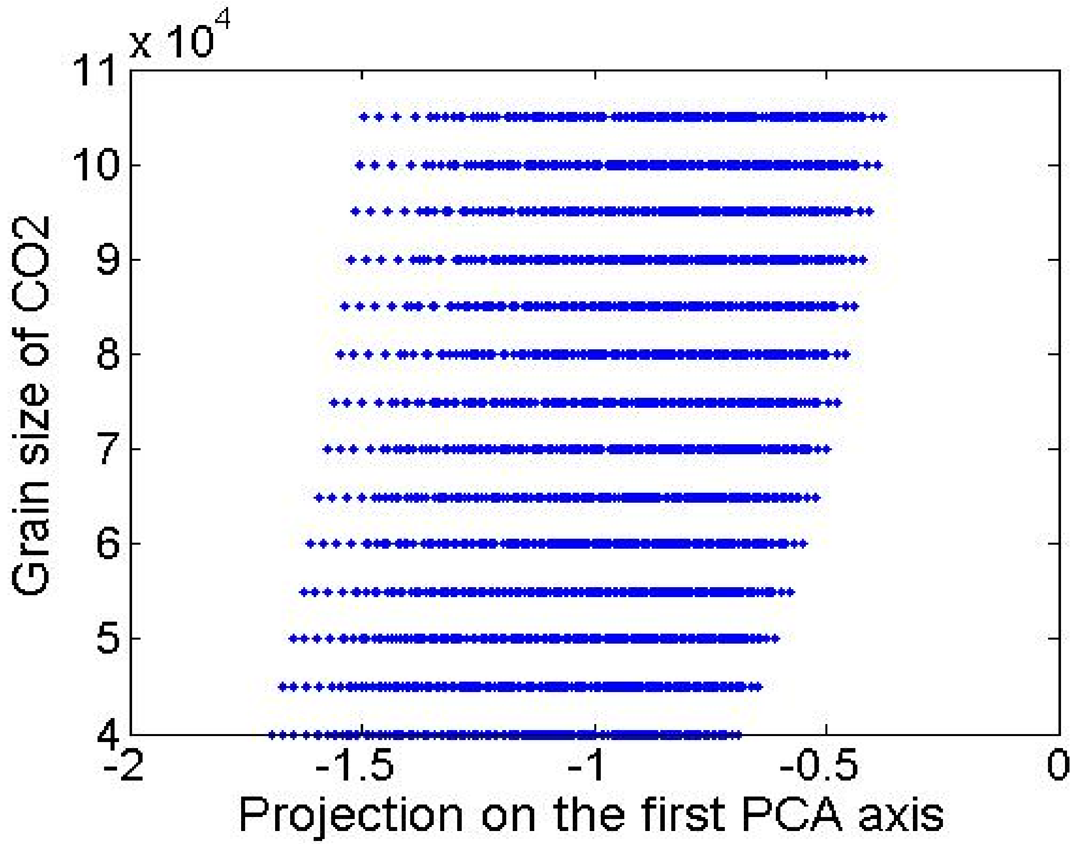
IV. Validation on simulations

V. Application to the south polar cap of Mars

VI. Conclusion and future work

Principal component analysis

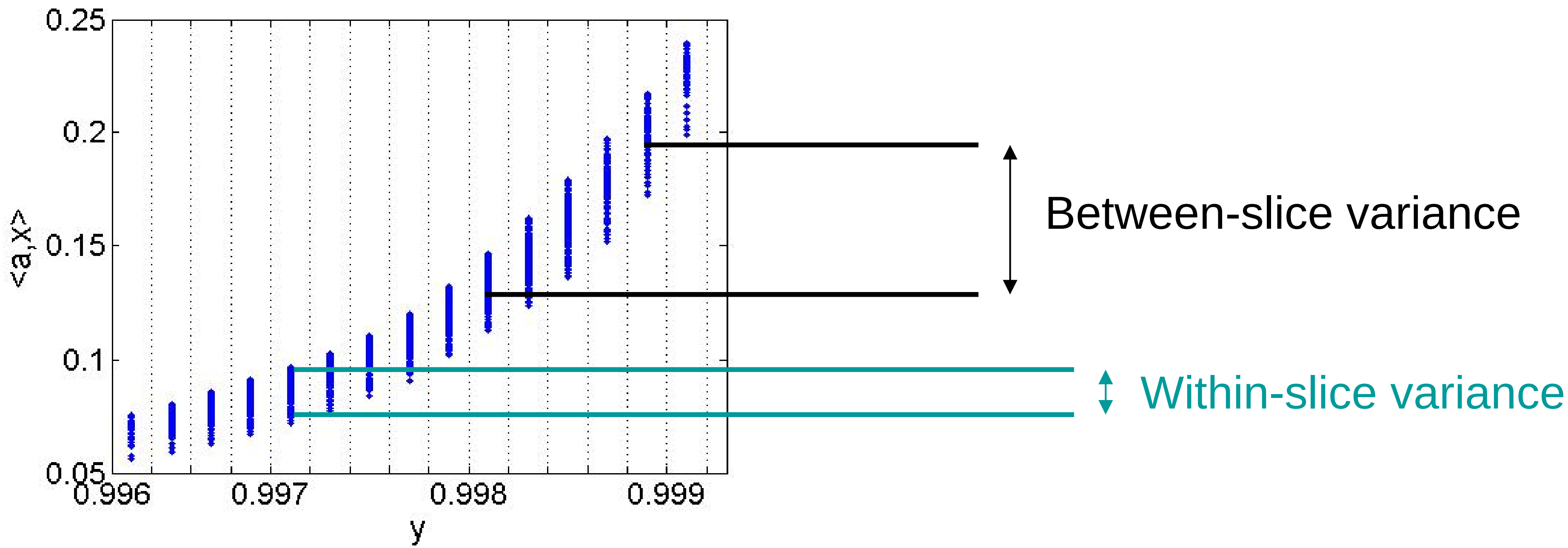
- **Maximizes the variance** of the projections of the observations *x*
- Does not take into account *y*



Sliced inverse regression

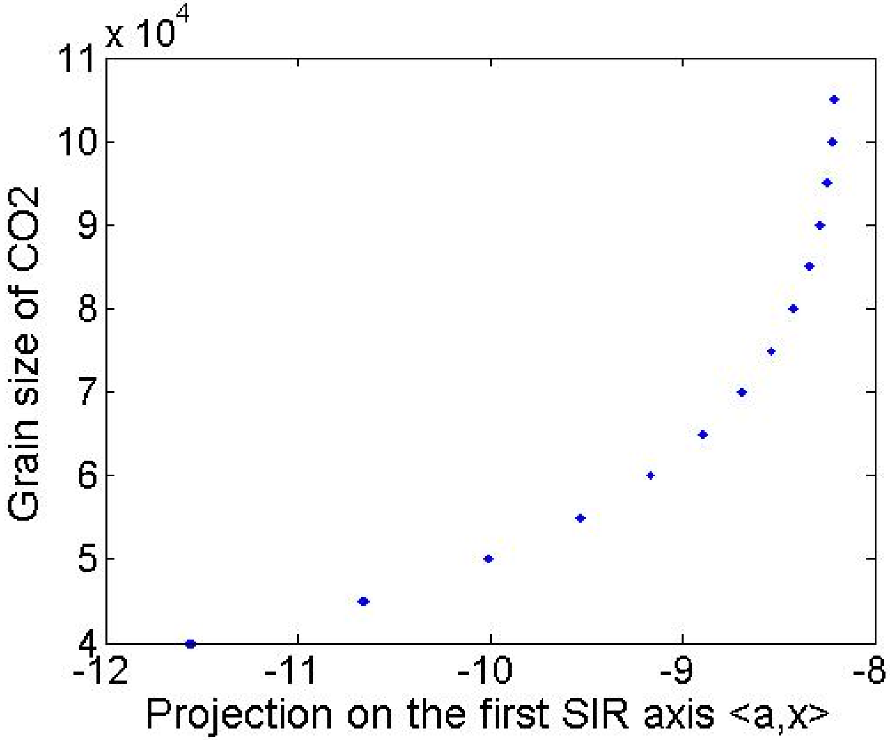
- Proposed by Li (1991)
- Maximizes the between-slice variance of projections
- PCA of E(*X*/*Y*) with $Z = \Sigma^{-1/2} X$
- Eigenvectors of $\Sigma^{-1}\Gamma$ with $\Gamma = \mathrm{var}(E(X/Y))$

  $\Sigma = \mathrm{var}(X)$

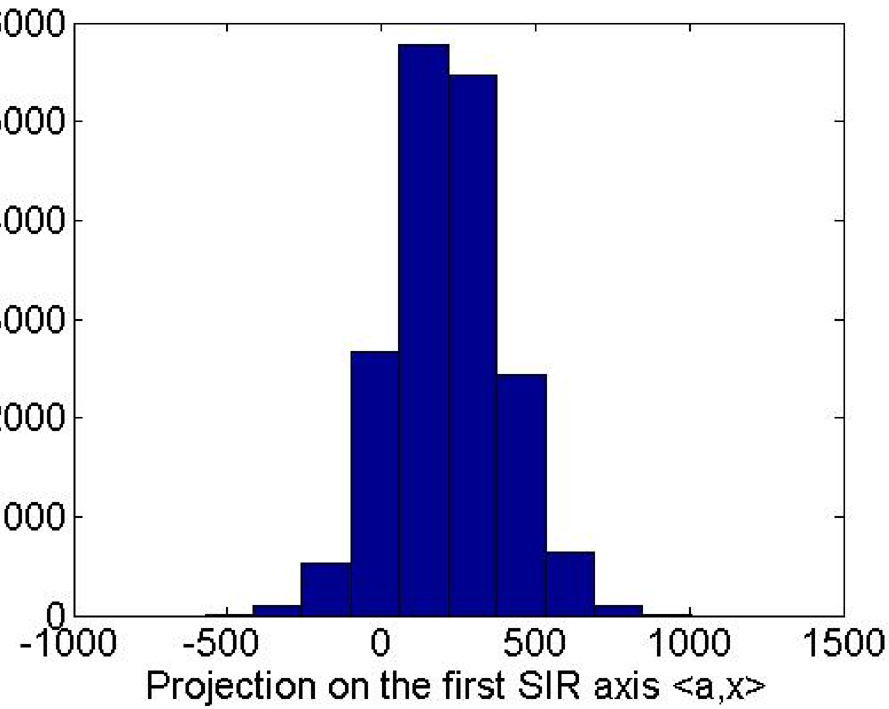


Application of SIR

TRAINING DATA



OBSERVED DATA



Problem

- Covariance matrix is ill-conditioned
  - Bad estimations of the directions
  - Sensitivity to noise
- Can be solved using regularization

DATA (x , y)

⇩

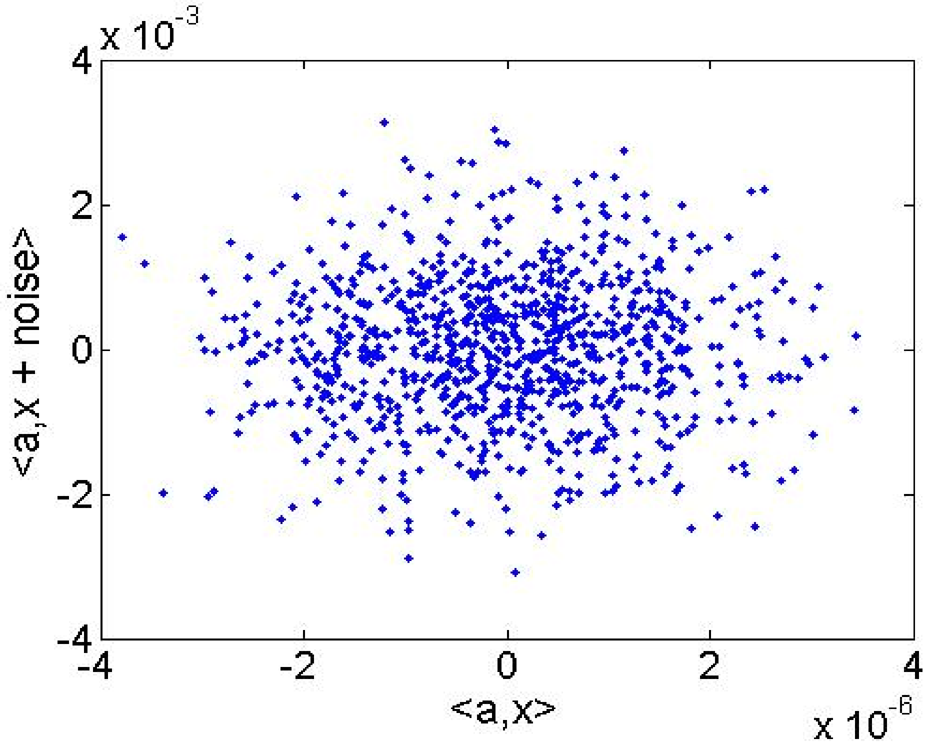
SLICED INVERSE REGRESSION

⇩

AXIS a

↙ ↘

< a, x > | < a , x + noise >



Outline

I. Context
   - Hyperspectral image data
   - Inverse problem
II. Dimension reduction
   - PCA
   - SIR
III. Regularization and estimation
   - Zhong et al., 2005
   - Tikhonov
IV. Validation on simulations
V. Application to the south polar cap of Mars
VI. Conclusion and future work

Regularization (1)

- Usual SIR:
  - eigenvectors of $\Sigma^{-1}\Gamma$
- Regularized SIR:
  - Zhong et al., 2005: eigenvectors of

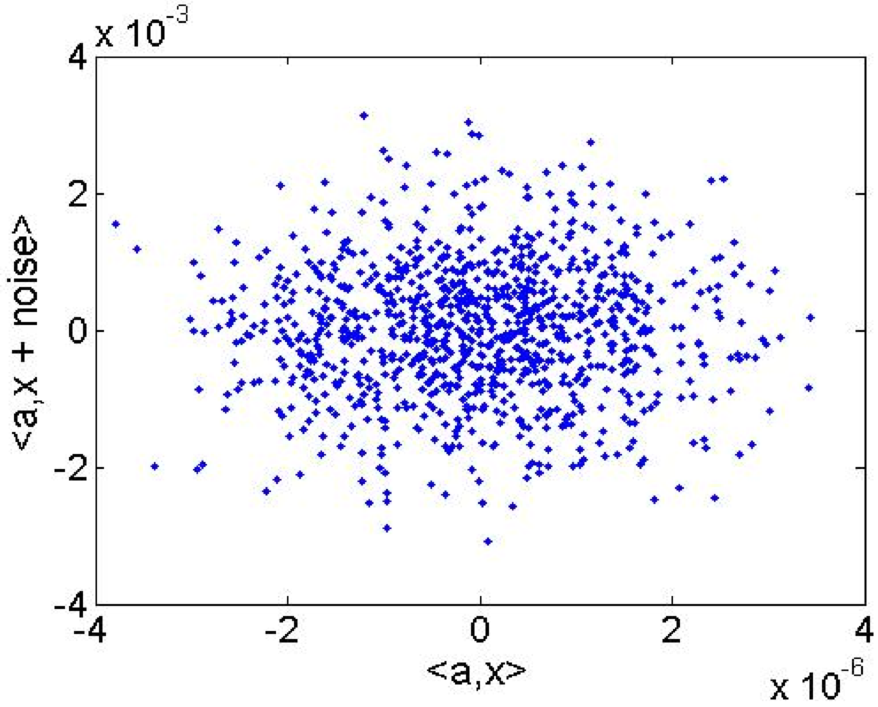
$$\left(\Sigma + \lambda Id\right)^{-1}\Gamma$$

  - Tikhonov regularization: eigenvectors of

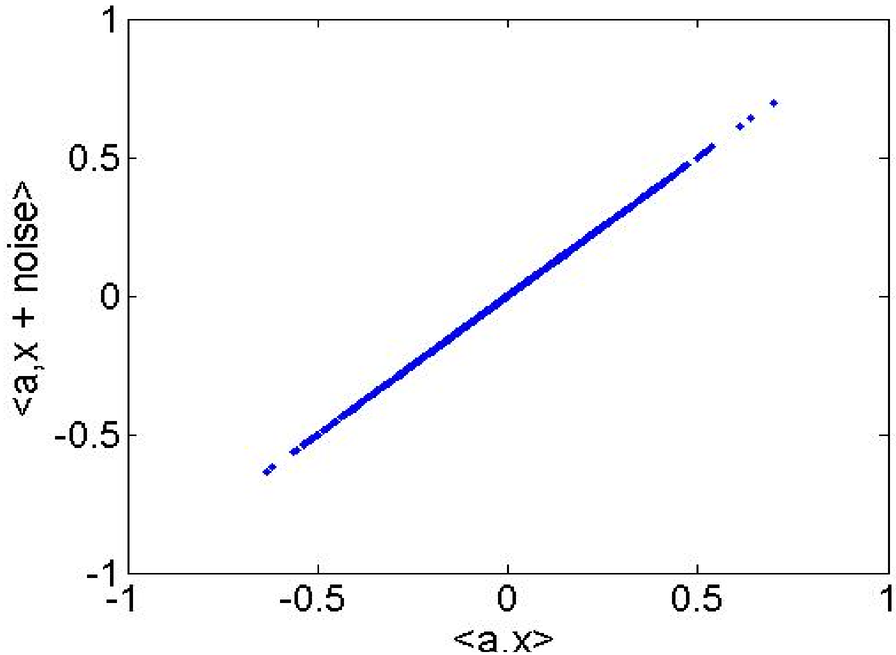
$$\left(\Sigma^2 + \lambda Id\right)^{-1}\Sigma\,\Gamma$$

Regularization (2)

Usual SIR



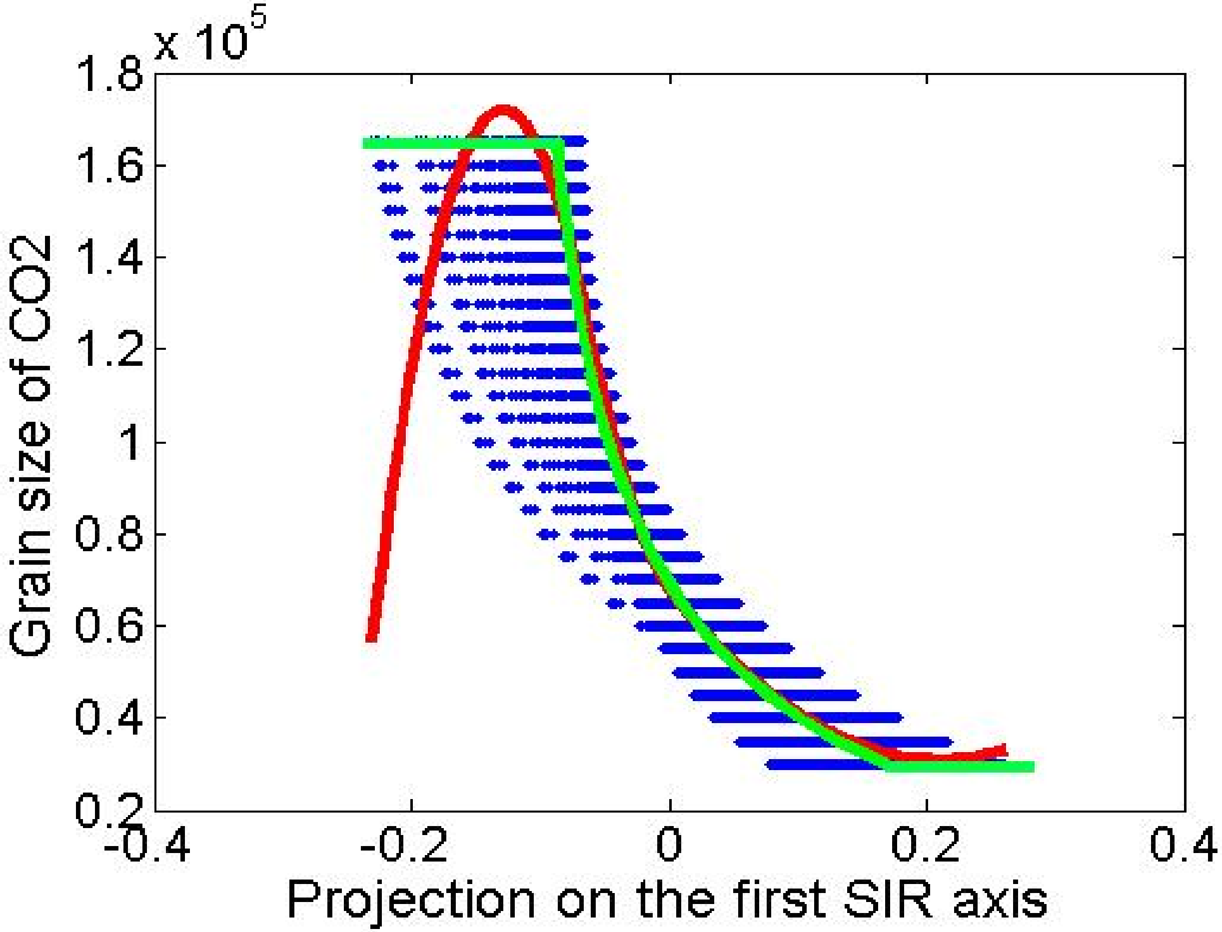
Regularized SIR (Tikhonov)



- *Depends on the regularization parameter λ*
- *The condition number of the matrix decreases when λ increases*
- *The estimation bias increases when λ increases*

Estimation

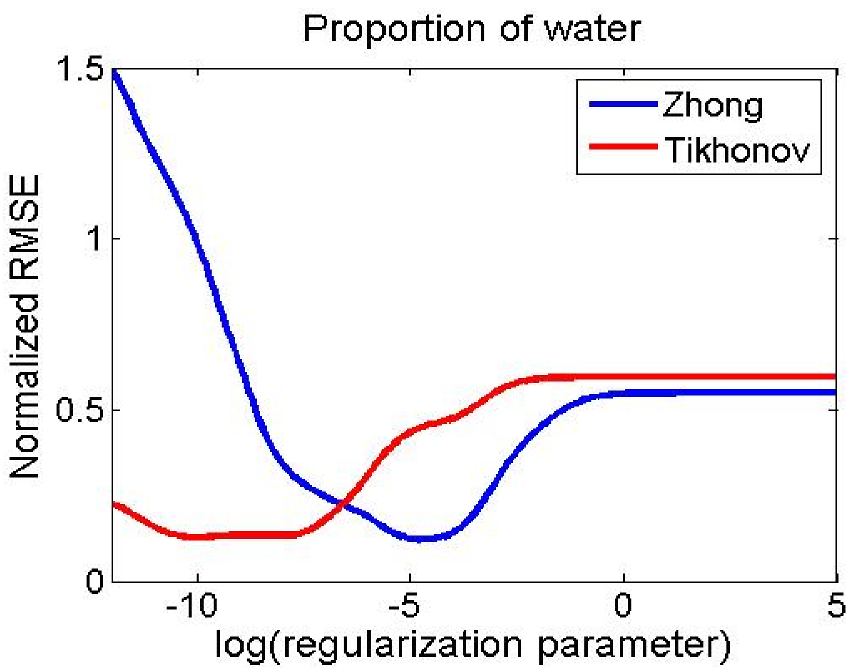
- Nearest neighbors (quite long!)
- Spline functions (choice of new parameters, boundaries)
- **Linear interpolation**

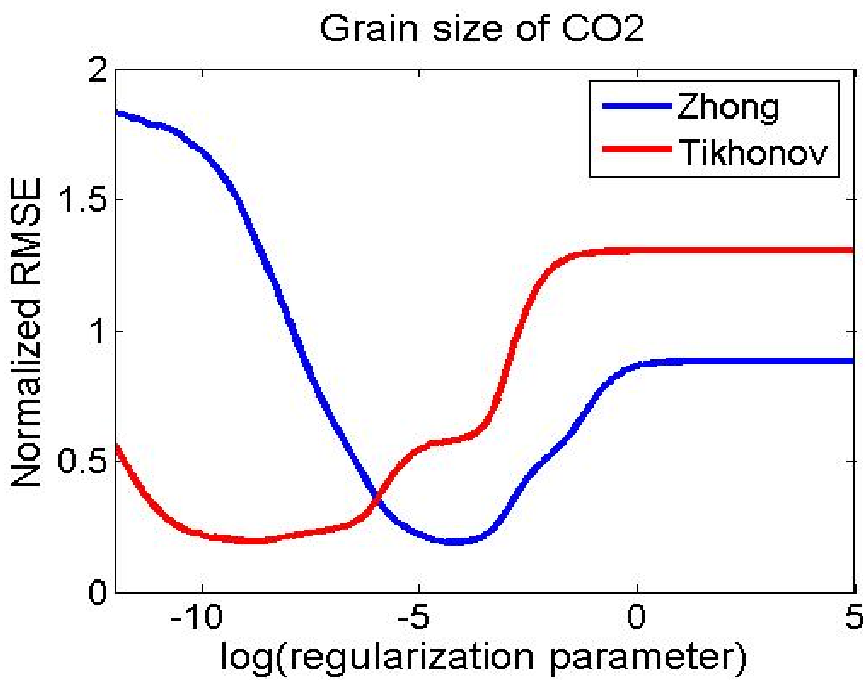


Choice of the regularization parameter

- By minimization of “Normalized RMSE” criterion

$$\frac{\|\hat{y}-y\|}{\|y-\bar{y}\|}=\frac{\text{Residuals sum of square}}{\text{Total sum of square}}$$





Outline

I. Context
- Hyperspectral image data
- Inverse problem

II. Dimension reduction
- PCA
- SIR

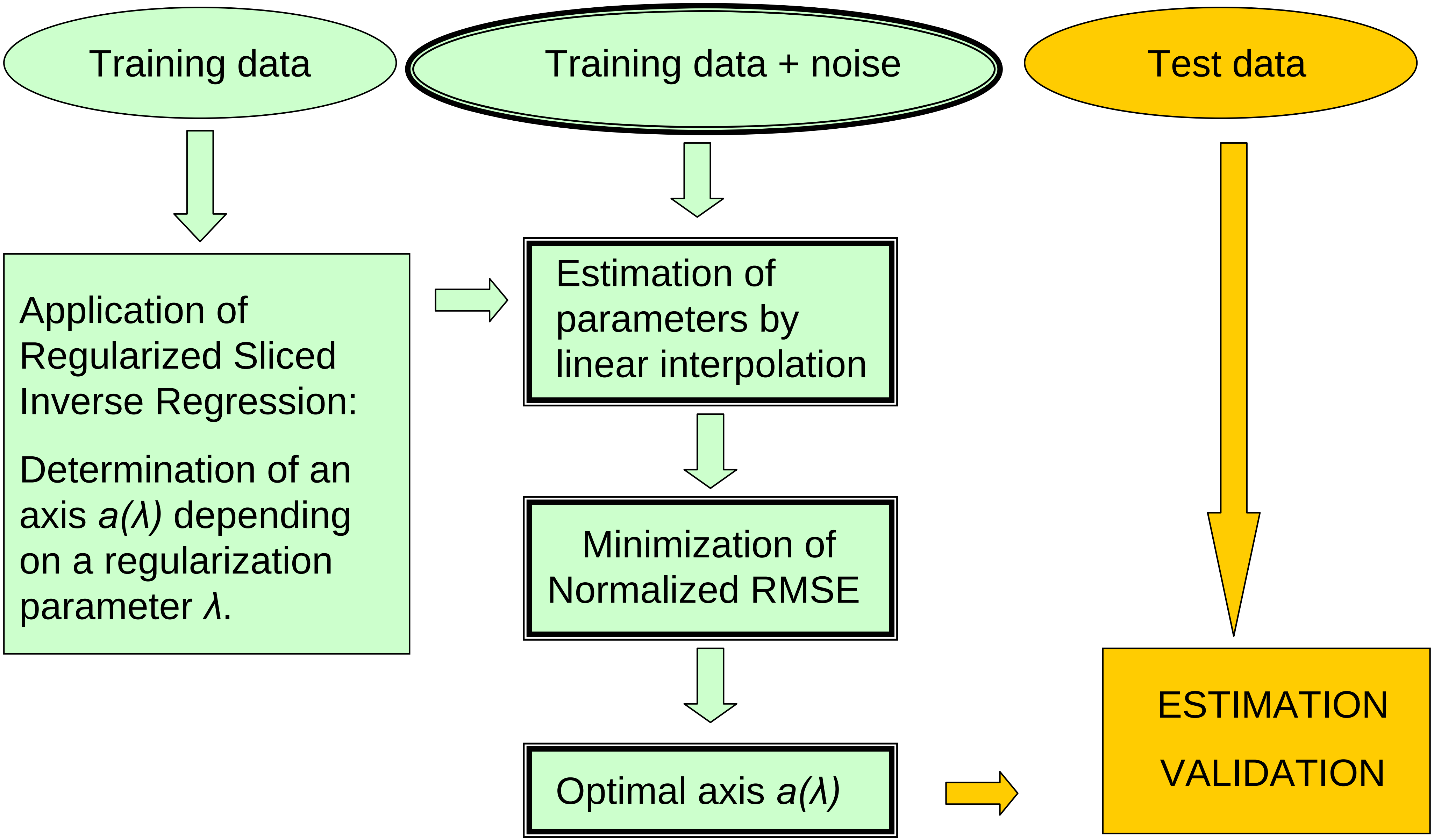
III. Regularization and estimation
- Zhong et al., 2005
- Tikhonov

IV. Validation on simulations

V. Application to the south polar cap of Mars

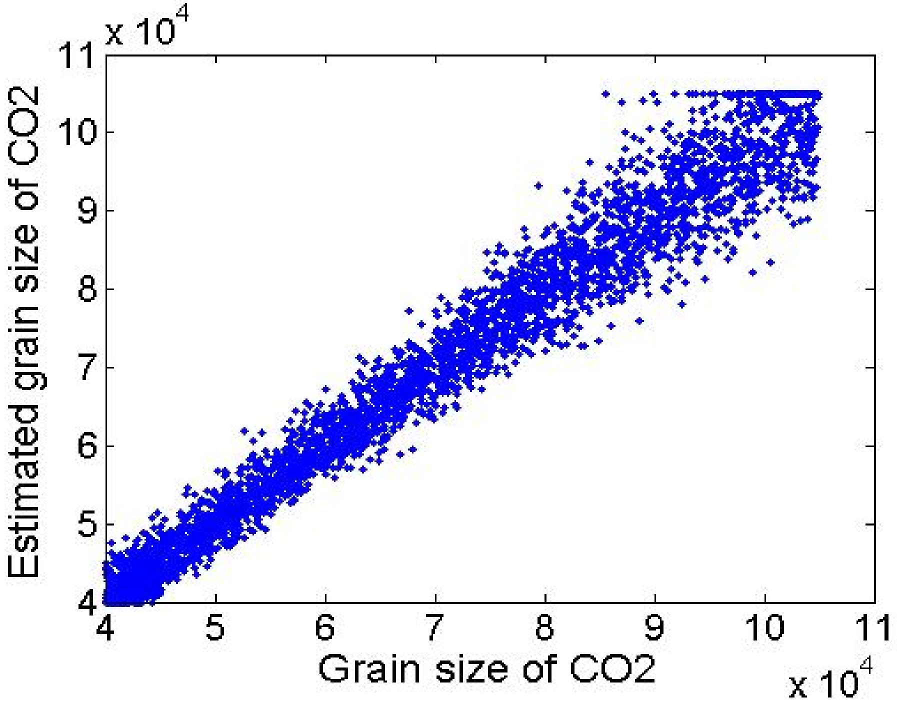
VI. Conclusion and future work

Validation (1)

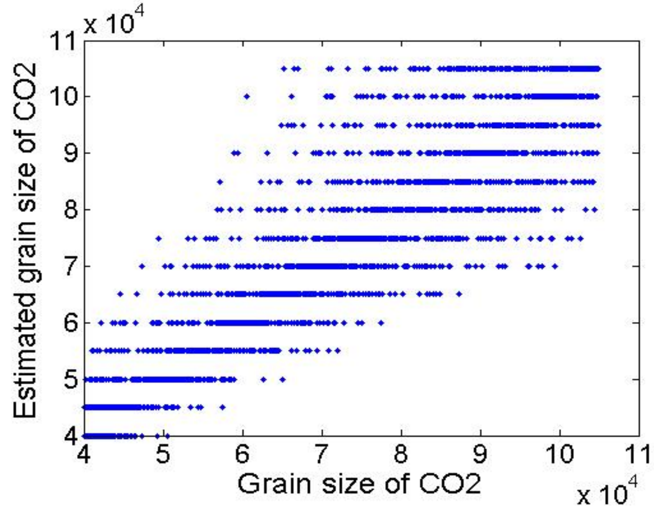


Validation (2)

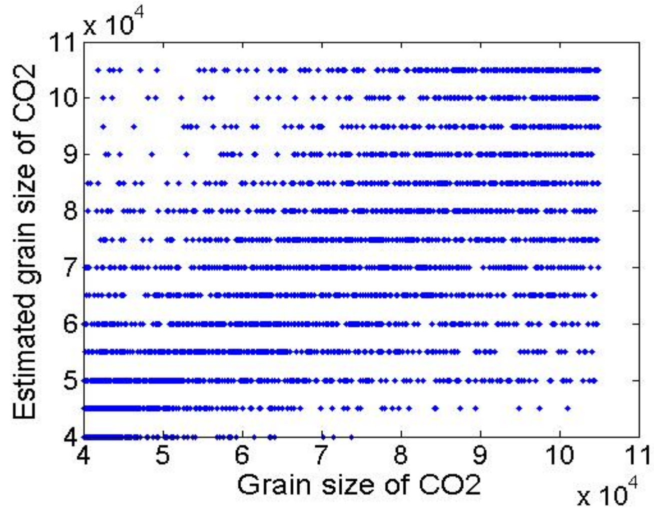
Regularized Sliced inverse regression (Tikhonov)



Nearest neighbors



Weighted nearest neighbors



Validation (3)

Normalized RMSE criterion:

$$\frac{\|\hat{y}-y\|}{\|y-\bar{y}\|} = \frac{\text{Residuals sum of square}}{\text{Total sum of square}}$$

Better when close to 0

SIR criterion:

$$\frac{{}^t a\,\Gamma\,a}{{}^t a\,\Sigma\,a} = \frac{\text{between-slice variance}}{\text{total variance}}$$

Better when close to 1

		Tikhonov		Zhong		Nearest neighbor	Weighted nearest neighbor
Parameter	**Variation interval**	$\frac{\|\hat{y}-y\|}{\|y-\bar{y}\|}$	$\frac{{}^t a\,\Gamma\,a}{{}^t a\,\Sigma\,a}$	$\frac{\|\hat{y}-y\|}{\|y-\bar{y}\|}$	$\frac{{}^t a\,\Gamma\,a}{{}^t a\,\Sigma\,a}$	$\frac{\|\hat{y}-y\|}{\|y-\bar{y}\|}$	$\frac{\|\hat{y}-y\|}{\|y-\bar{y}\|}$
Proportion of dust	[0.0006 0.002]	0.33	0.92	0.33	0.92	0.56	0.53
Proportion of CO2	[0.9960 0.9988]	0.27	0.96	0.23	0.94	0.56	0.54
Proportion of water	[0.0006 0.002]	0.13	1.00	0.12	1.00	0.27	0.28
Grain size of water	[100 400]	0.37	0.92	0.38	0.87	0.40	0.68
Grain size of CO2	[40000 105000]	0.19	0.99	0.18	0.98	0.38	0.82

- SIR gives better results than nearest neighbor classification
- Tikhonov and Zhong regularizations are equivalent
- With Tikhonov regularization, minimal normalized RMSE is reached on a larger interval than with Zhong's.

Outline

I. Context
- Hyperspectral image data
- Inverse problem

II. Dimension reduction
- PCA
- SIR

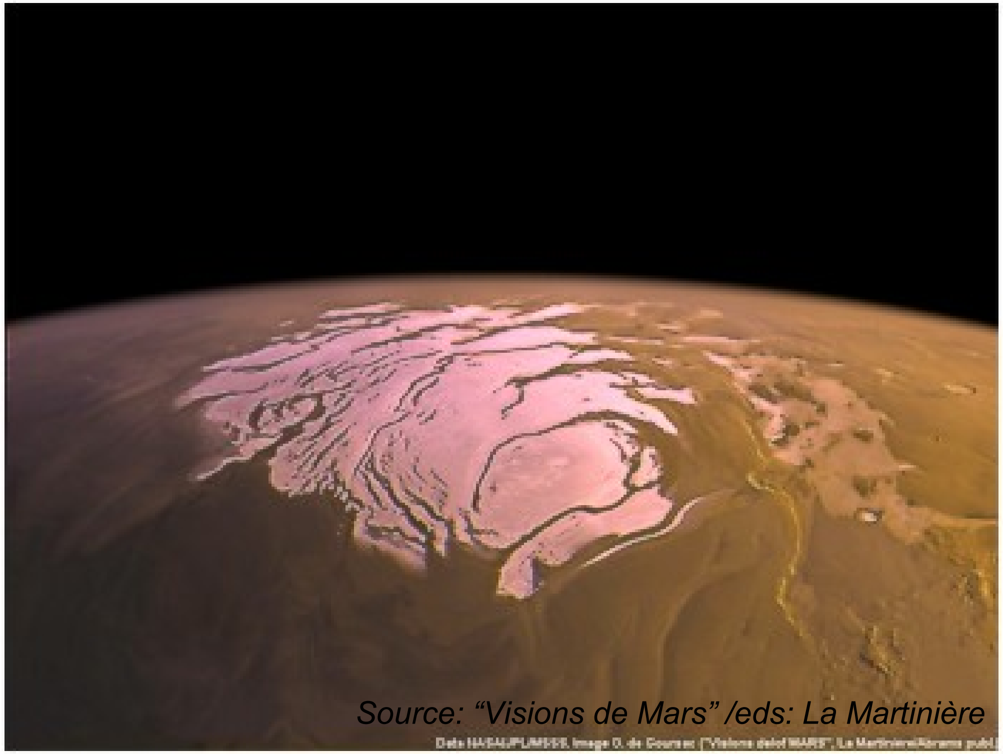
III. Regularization and estimation
- Zhong et al., 2005
- Tikhonov

IV. Validation on simulations

V. Application to the south polar cap of Mars

VI. Conclusion and future work

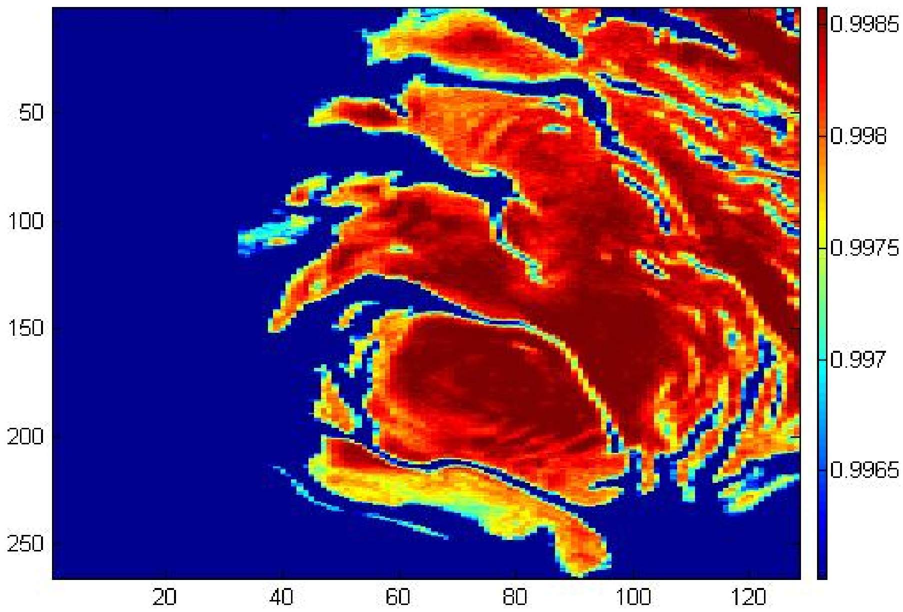
Application to south polar cap of Mars


Source: “Visions de Mars” /eds: La Martinière

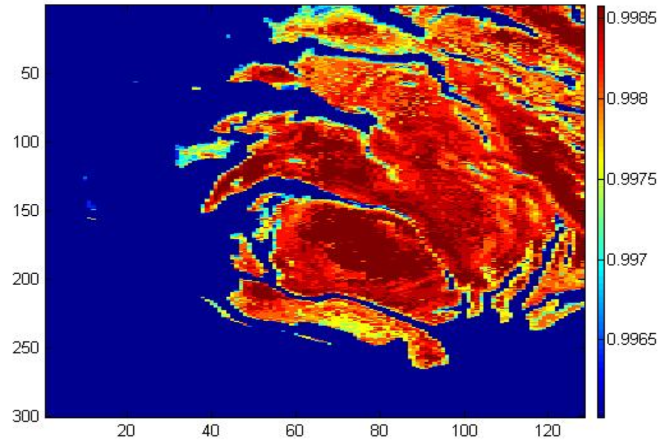
- Model determined by physicists (water + CO_2 + dust)
- 17753 spectra
- 184 wavelengths
- Training data simulated by radiative transfer model
- 5 parameters to study : proportions of water, dust and CO_2, grain sizes of CO_2 and water.

Proportion of CO2

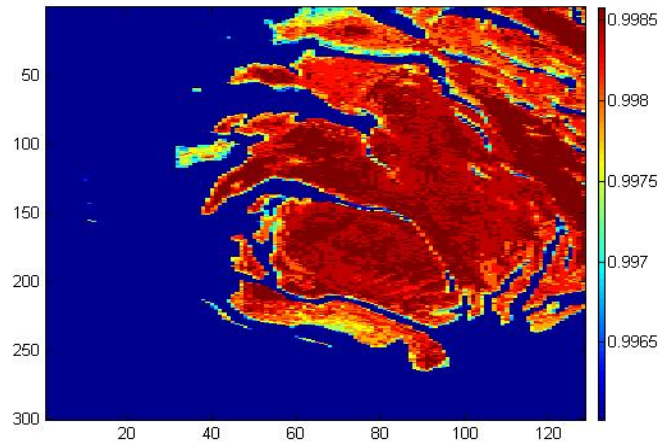
Regularized Sliced Inverse Regression (Tikhonov)



Nearest neighbors

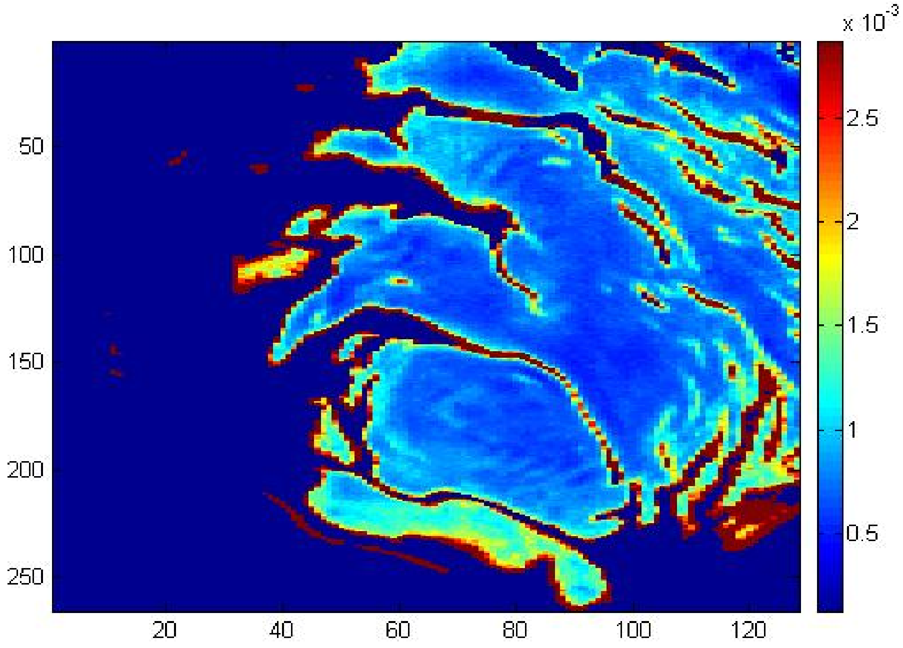


Weighted nearest neighbors

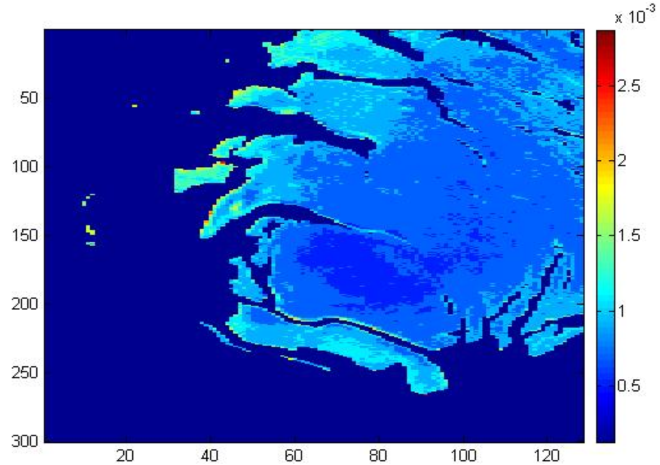


Proportion of water

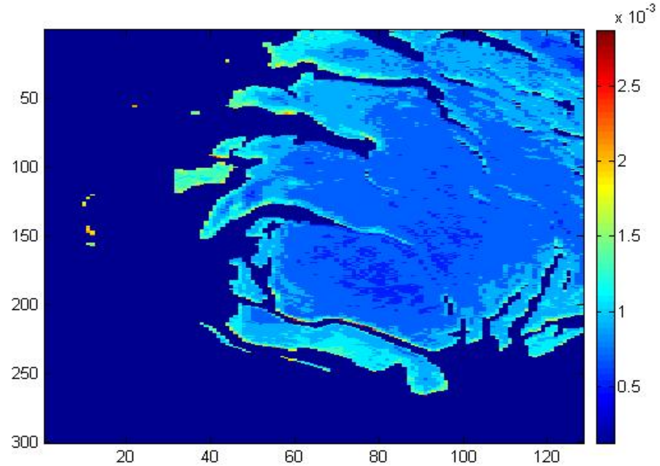
Regularized Sliced Inverse Regression (Tikhonov)



Nearest neighbors



Weighted nearest neighbors



Outline

I. Context
- Hyperspectral image data
- Inverse problem

II. Dimension reduction
- PCA
- SIR

III. Regularization and estimation
- Zhong et al., 2005
- Tikhonov

IV. Validation on simulations

V. Application to the south polar cap of Mars

VI. Conclusion and future work

Conclusion and further work

- Good results on simulations.
- Realistic results on real data.
- Validation is difficult because of the lack of ground measurements.
- SIR for data streams.

SIR for data streams

- We consider data arriving sequentially by blocks in a stream.
- Each data block $j = 1, \ldots, J$ is a n-sample from the regression model $y = g(< a, x >, \varepsilon)$.
- **Goal** : Update the estimation of the direction a at each arrival of a new block of observations.

Joint work with team CQFD, INRIA Bordeaux Sud-Ouest

Method

- Compute the **individual directions** $\hat{a}_j$ on each block $j = 1, \ldots, J$ using regularized SIR.
- Compute a **common direction** as

$$\hat{a} = \underset{||a||=1}{\operatorname{argmax}} \sum_{j=1}^{J} \cos^2(\hat{a}_j, a) \cos^2(\hat{a}_j, \hat{a}_J).$$

Idea : If $\hat{a}_j$ is close to $\hat{a}_J$ then $\hat{a}$ should be close to $\hat{a}_j$.
Explicit solution : $\hat{a}$ is the eigenvector associated to the largest eigenvalue of

$$M_J = \sum_{j=1}^{J} \hat{a}_j \hat{a}_j^t \cos^2(\hat{a}_j, \hat{a}_J).$$

Advantages of SIRdatastream

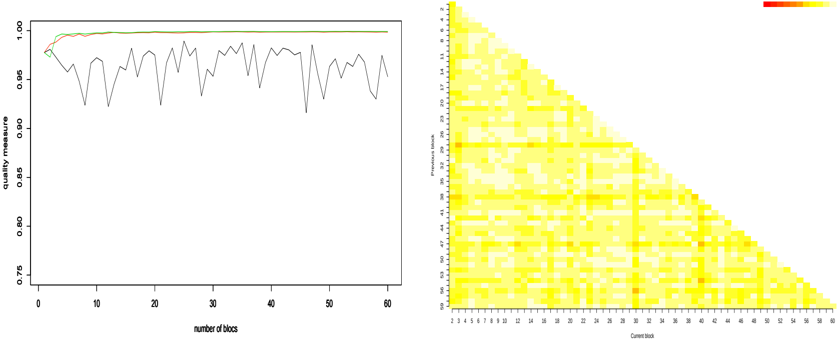
- Computational complexity $O(Jnp^2)$ v.s. $O(J^2np^2)$ for the brute-force method which would consist in applying regularized SIR on the union of the j first blocks for $j = 1, \dots, J$.
- Data storage $O(np)$ v.s. $O(Jnp)$ for the brute-force method.

(under the assumption $n >> \max(J, p)$).

- Interpretation of the weights $\cos^2(\hat{a}_j, \hat{a}_J)$.

Illustration on simulations

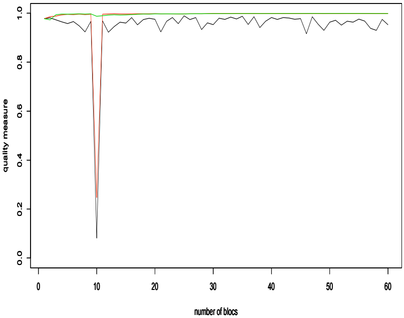
Scenario 1 : A common direction in all the 60 blocks.

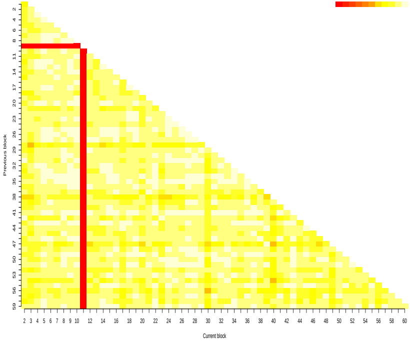


Left : $\cos^2(\hat{a}, a)$ for SIRdatastream, SIR brute-force and SIR estimators at each time t. *Right :* $\cos^2(\hat{a}_j, \hat{a}_J)$. The lighter (yellow) is the color, the larger is the weight. Red color stands for very small squared cosines.

Illustration on simulations

Scenario 2 : The 10th block is an outlier.

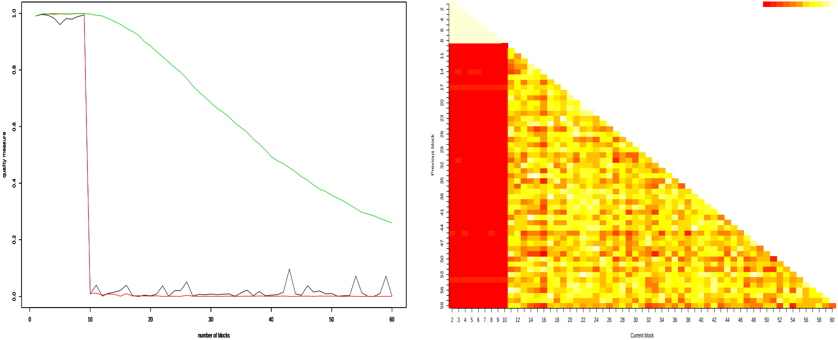




Left : $\cos^2(\hat{a}, a)$ for SIRdatastream, SIR brute-force and SIR estimators at each time t. *Right :* $\cos^2(\hat{a}_j, \hat{a}_J)$. The lighter (yellow) is the color, the larger is the weight. Red color stands for very small squared cosines.

Illustration on simulations

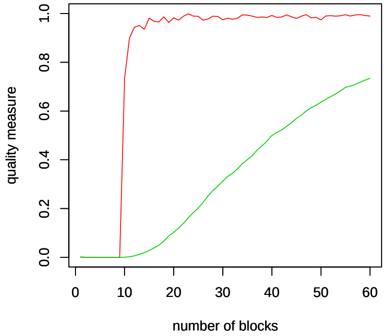
Scenario 3 : A drift occurs from the 10th block (a to a')

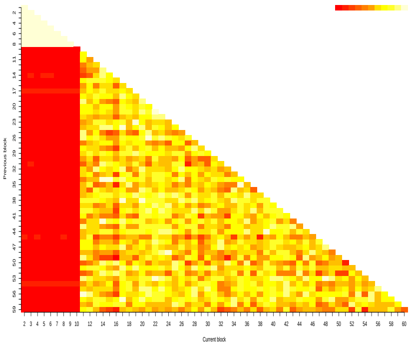


Left : $\cos^2(\hat{a}, a)$ for SIRdatastream, SIR brute-force and SIR estimators at each time t. *Right :* $\cos^2(\hat{a}_j, \hat{a}_J)$. The lighter (yellow) is the color, the larger is the weight. Red color stands for very small squared cosines.

Illustration on simulations

Scenario 3 (cont'd) : A drift occurs from the 10th block (a to a')

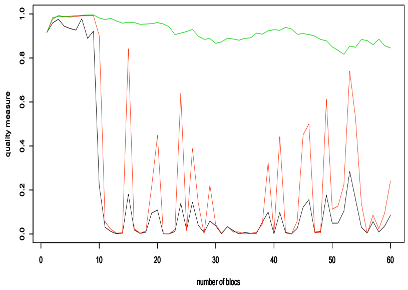


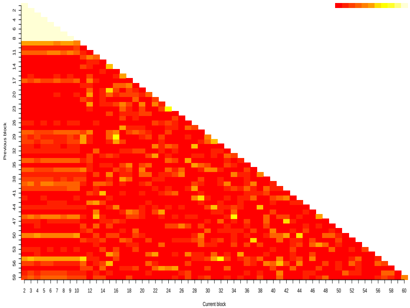


Left : $\cos^2(\hat{a}, a')$ for SIRdatastream and SIR brute-force. *Right :* $\cos^2(\hat{a}, a')$

Illustration on simulations

Scenario 4 : From the 10th block to the last one, there is no common direction.





Left : $\cos^2(\hat{a}, a)$ for SIRdatastream, SIR brute-force and SIR estimators at each time t. *Right :* $\cos^2(\hat{a}_j, \hat{a}_J)$. The lighter (yellow) is the color, the larger is the weight. Red color stands for very small squared cosines.

Related research themes in Mistis

Dimension reduction

- Dimension reduction for regression (this talk),
- Unsupervised dimension reduction (nonlinear PCA),
- Dimension reduction for classification and clustering.

Classification

- Robust classification and outlier detection,
- Classification with missing data,
- Classification of spatial data,
- Classification of heterogeneous data.

References on SIR

- [Li, 1991] Li, K.C. (1991). Sliced inverse regression for dimension reduction. *Journal of the American Statistical Association*, **86**, 316–327.
- [Cook, 2007]. Cook, R.D. (2007). Fisher lecture : Dimension reduction in regression. *Statistical Science*, **22**(1), 1–26.
- [Zhong et al, 2005] : Zhong, W., Zeng, P., Ma, P., Liu, J.S. and Zhu, Y. (2005). RSIR : Regularized Sliced Inverse Regression for motif discovery. *Bioinformatics*, **21**(22), 4169–4175.
- [Chiaromonte et al, 2002] : Chiaromonte, F. and Martinelli, J. (2002). Dimension reduction strategies for analyzing global gene expression data with a response. *Mathematical Biosciences*, **176**, 123–144.

References on this work

- Bernard-Michel, C., Douté, S., Fauvel, M., Gardes, L. et Girard, S. (2009). Retrieval of Mars surface physical properties from OMEGA hyperspectral images using Regularized Sliced Inverse Regression. *Journal of Geophysical Research - Planets*, **114**, E06005
- Bernard-Michel, C., Gardes, L. et Girard, S. (2009). Gaussian Regularized Sliced Inverse Regression, *Statistics and Computing*, **19**, 85–98.
- Bernard-Michel, C., Gardes, L. et Girard, S. (2008). A Note on Sliced Inverse Regression with Regularizations, *Biometrics*, **64**, 982–986.