Inverting hyperspectral images with Gaussian Regularized Sliced Inverse Regression

INRIA

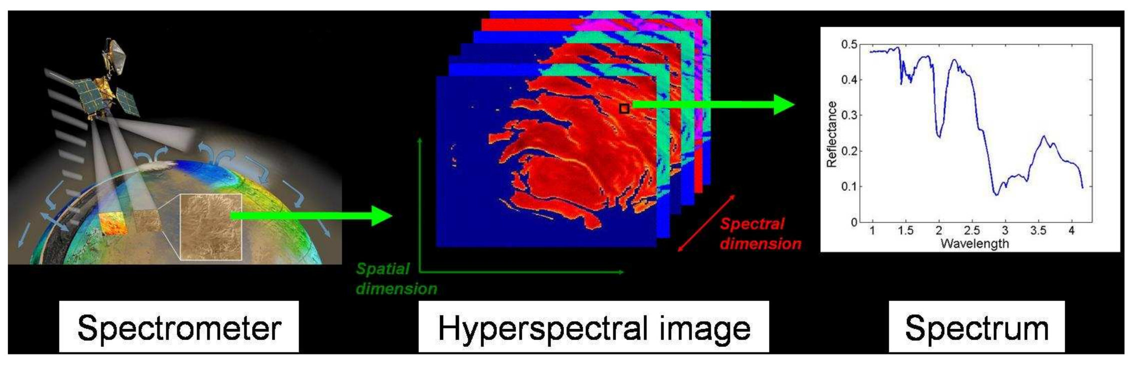
Caroline Bernard-Michel, Sylvain Douté, Laurent Gardes and Stéphane Girard

MISTIS - INRIA Rhône-Alpes
http://mistis.inrialpes.fr/

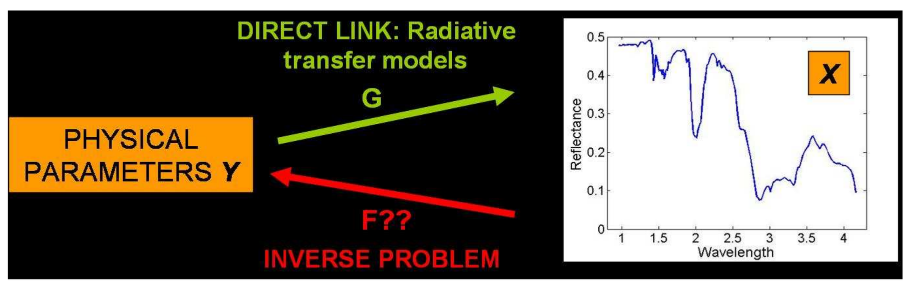
Laboratoire de Planétologie Grenoble

The inverse problem

- Visible and near infrared imaging spectroscopy allows the detection, mapping and characterization of minerals and ices by analyzing the solar light reflected in different directions by the surface materials.



- Modeling the direct link between some physical parameters Y and observable spectra X is called the **forward problem** and allows, for given values of the model parameters, to simulate the spectra that should be observed.
- Conversely, deducing the physical model parameters from the observed spectra is called an **inverse problem**.
- Application to OMEGA/MEX hyperspectral images observed on Mars [2].



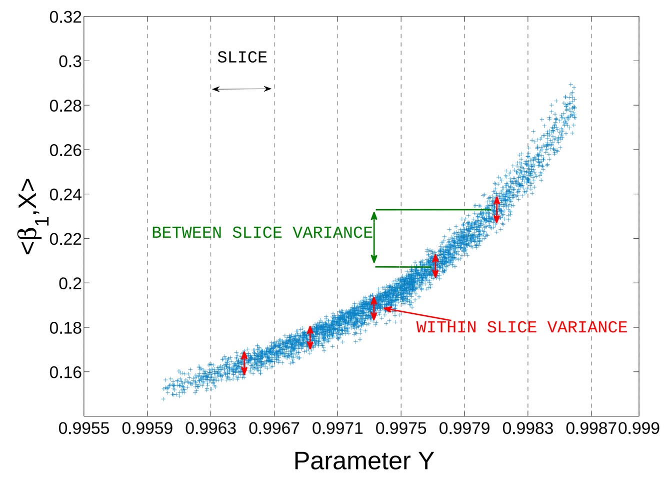
Our approach

Functional approach and dimension reduction

- Estimate the functional relationship F between the spectra $X \in \mathbb{R}^p$ and one parameter $Y \in \mathbb{R}$ ($p = 184$ wavelengths).
- Because of the curse of dimensionality, **dimension reduction** techniques are required.
- They rely on the assumption that the predictor X can be replaced by its projection on a subspace of smaller dimension K without loss of information. Denoting by $\beta_1, \dots, \beta_K$ a basis of this subspace, the functional relationship $Y = F(X)$ can be rewritten as $Y = f(\beta_1^t X, \dots, \beta_K^t X)$ where f is now a K-variate function.

Sliced Inverse regression

- Introduced by Li [4]
- Find the directions $b = (\beta_1, \dots, \beta_K)$ such that $b^t X$ best explains Y.
- Find the directions b minimizing the variations of $b^t X$ given Y.
- In practice, the range of Y is partitioned into h slices and one needs to calculate the eigenvectors of $\Sigma^{-1}\Gamma$ where Σ is the spectra covariance matrix and Γ the slice mean spectra covariance matrix.



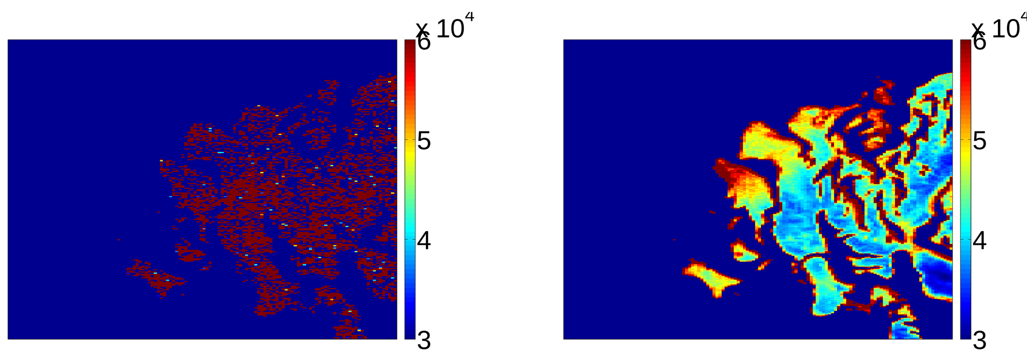
Gaussian Regularized Sliced Inverse Regression (GRSIR)

Limits of SIR

- In inverse problems, Σ is generally **ill-conditioned** or **singular**.
- In presence of noise, estimates are hugely biased.
- A **regularization** is required.

Idea of GRSIR

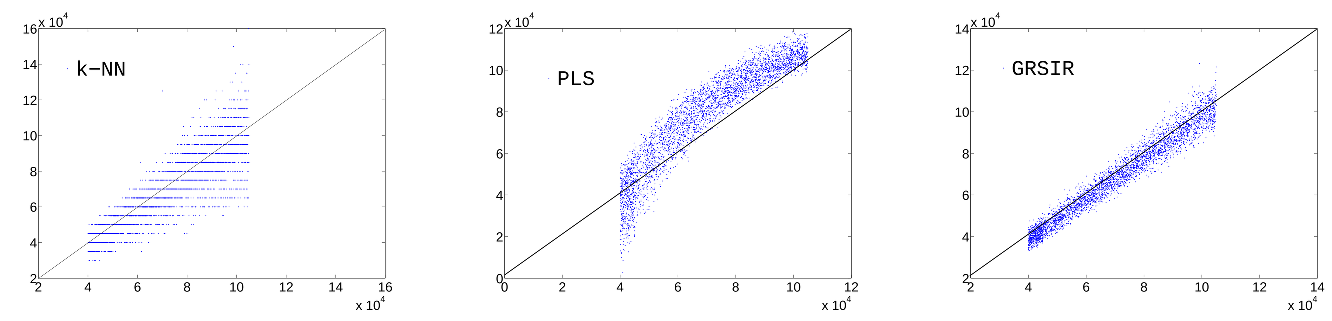
- Incorporate some prior information on the projections in order to dampen the effect of the noise [1].
- Instead of computing the eigenvectors of $\Sigma^{-1}\Gamma$, we propose to compute the eigenvectors of $(\Sigma^2 + \delta I_p)^{-1}\Sigma\Gamma$ in a manner similar to Tikhonov regularization. δ is called the regularization parameter. It makes a compromise between improving estimations and maintaining the functional relationship.



Grain size of CO_2 ice estimated by SIR (left) and GRSIR (right) on a hyperspectral image observed on the CO_2 ice of the south polar cap of Mars during orbit 61

Validation on simulations

- One GRSIR axis is sufficient.
- the k-nearest neighbors methodology (k-NN) is very unstable.
- GRSIR gives the best results in terms of Normalized Root Mean Square Errors (NRMSE) for most parameters.
- PLS [3] does not seem suited because the relationship is non linear.
- There is still a small bias with GRSIR due to the choice of the learning database.



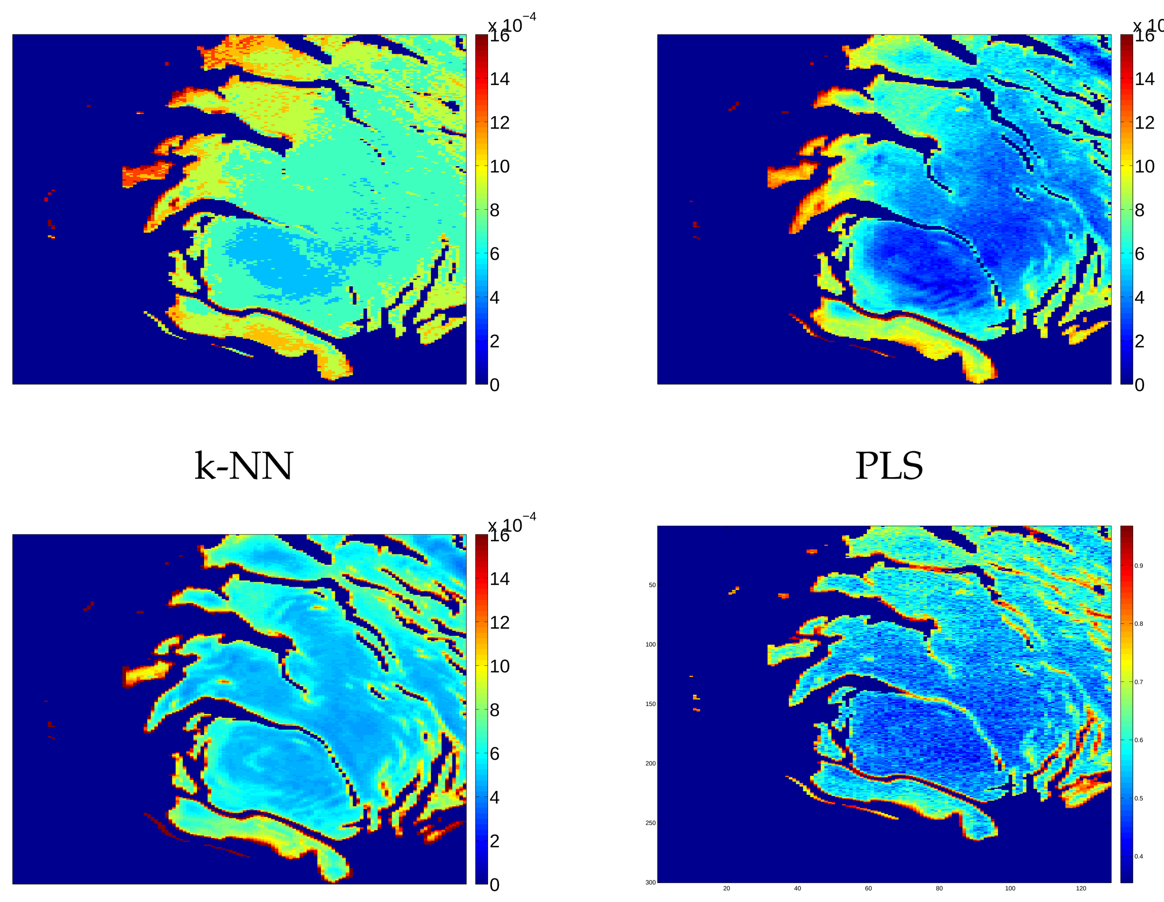
Estimation of the grain size of CO_2 ice (Y-axis) by k-NN, PLS and GRSIR versus real values (X-axis)

Parameters	k-NN	PLS	GRSIR	
	NRMSE	NRMSE	NRMSE	SIRC
Proportion of water	0.86	0.50	0.40	0.90
Proportion of CO_2	0.88	0.58	0.30	0.98
Proportion of dust	0.44	0.48	0.17	0.99
Grain size of water	0.43	0.65	0.54	0.84
Grain size of CO_2	0.53	0.41	0.22	0.95

Normalized Root Mean Square Errors for k-NN, PLS and GRSIR methodologies on simulations. The NRMSE criterion quantifies the importance of estimation errors (must be close to zero). The SIR criterion (SIRC) quantifies the quality of the relationship f (must be close to one).

Inversions of real hyperspectral images

- Validation is difficult because no ground truth data is available.
- GRSIR first axis does put weights on key spectral points according to researchers in planetary physics.
- GRSIR estimations vary continuously and seem to be spatially coherent.
- GRSIR map is more detailed.
- GRSIR is in accordance with the Wavanglet physical approach whereas in some regions, k-NN and PLS give conflicting estimations. Wavanglet is a supervised classification method that allows the detection and quantification of major compounds on hyperspectral images [5].
- Images from different orbits but analyzing the same portion of Mars give similar GRSIR estimates.
- When spectra cannot be inversed by GRSIR, it generally means they correspond to another physical model.



Proportion of dust estimated by k-NN, PLS, GRSIR on a hyperspectral image observed on Mars during orbit 41. Wavanglet: Cosinus of the spectral angle between each spectrum and a given reference spectrum.

Bibliography

- [1] C. Bernard-Michel, L. Gardes, and S. Girard (2007), Gaussian regularized sliced inverse regression, Technical report, INRIA.
- [2] J-P. Bibring et al. (2004), Perennial water ice identified in the south polar cap of mars, *Nature*, 428:627–630.
- [3] T. Hastie, R. Tibshirani, and J. Friedman (2001), *The elements of Statistical Learning*, Springer.
- [4] K.C. Li (1991), Sliced inverse regression for dimension reduction, *Journal of the American Statistical Association*, 86:316-327.
- [5] F. Schmidt, S. Douté, and B. Schmitt (2007). WAVANGLET: An efficient supervised classifier for hyperspectral images, *Geoscience and Remote Sensing, IEEE Transactions*, 45(5):1374–1385.