Support Vectors Machines Regression for Estimation of Mars Surface Physical Properties

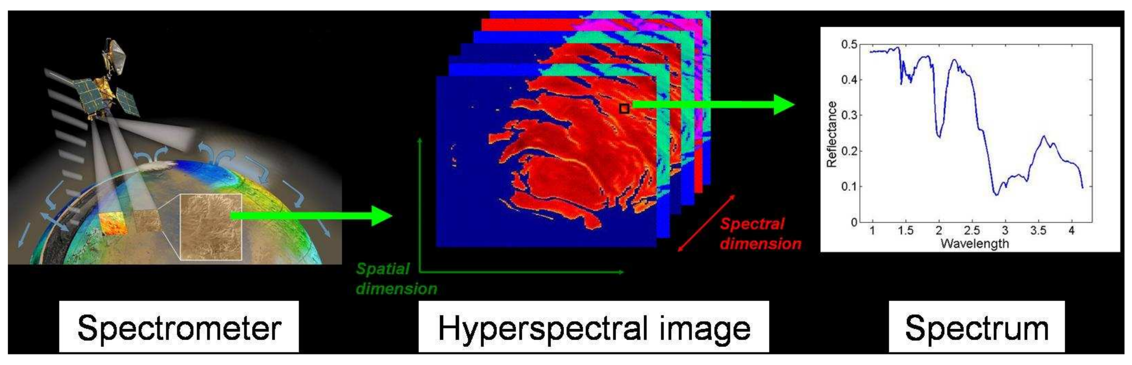
Caroline Bernard-Michel, Sylvain Douté, Mathieu Fauvel, Laurent Gardes and Stéphane Girard

MISTIS - INRIA Rhône-Alpes & Laboratoire Jean Kuntzmann - Laboratoire de Planénotologie de Grenoble

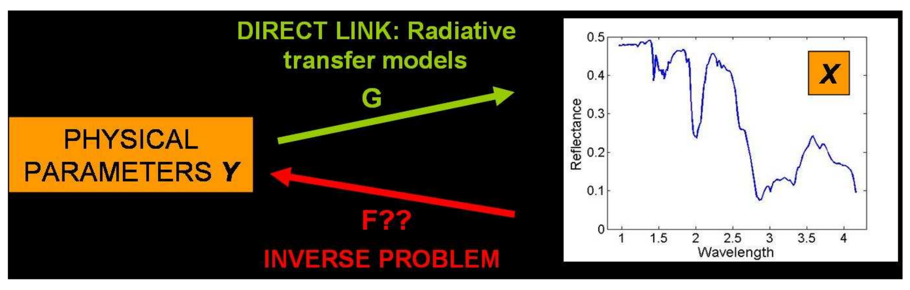
http://mistis.inrialpes.fr/

I. 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 [1].



II. Our approach

Regression Problem

- Estimate the functional relationship f between the spectra $\mathbf{x} \in \mathbb{R}^d$ and one parameter $y \in \mathbb{R}$ ($d = 184$ wavelengths).
- Because of the **curse of dimensionality**, parameters estimation are difficult.
- Model free approaches based on **statistical learning theory** are a good alternative to parametric ones.

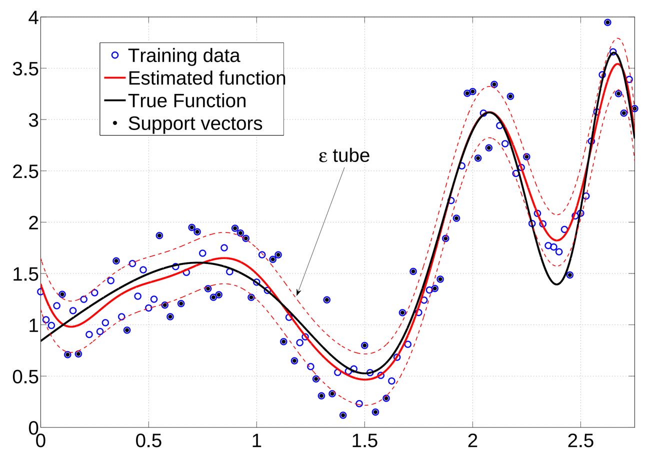
Support Vectors Machines Regression

- Structural risk minimization [2]:

$$\min_f \left[\frac{1}{n}\sum_{i=1}^{n} l\big(f(\mathbf{x}_i), y_i\big) + \lambda\|f\|^2\right] \text{ with } l\big(f(\mathbf{x}), y\big) = \begin{cases} 0 \text{ if } |f(\mathbf{x}) - y| \leq \epsilon \\ |f(\mathbf{x}) - y| - \epsilon \text{ otherwise.} \end{cases}$$

- Learn f of the form: $f(\mathbf{x}) = \sum_{i=1}^{n} \alpha_i k(\mathbf{x}, \mathbf{x}_i) + b$.
- $((\alpha_i)_{i=1,\dots,n}, b)$ found by convex optimization.
- k is a **kernel function**: f might be non-linear.

Kernel		Parameters
Linear	$\langle \mathbf{x}, \mathbf{z}\rangle$	-
Polynomial	$\big(\langle \mathbf{x}, \mathbf{z}\rangle + q\big)^p$	$q \geq 0, p \in \mathbb{N}^+$
Gaussian	$\exp\big(-\gamma\|\mathbf{x} - \mathbf{z}\|^2\big)$	$\gamma \in \mathbb{R}^+$
Spectral	$\exp\big(-\gamma \operatorname{acos}\big((\langle \mathbf{x}, \mathbf{z}\rangle)/(\|\mathbf{x}\|\|\mathbf{z}\|)\big)^2\big)$	$\gamma \in \mathbb{R}^+$



Simulated data: $\mathbf{x}_i$ corresponding to non-zero α_i are called ***support vectors***

III. Validations on simulations

Data sets

- Simulated by radiative transfert; 3584 training samples & 3528 test samples
- 5 parameters : proportion of CO_2, H_2O & dust - grain size of CO_2 & H_2O

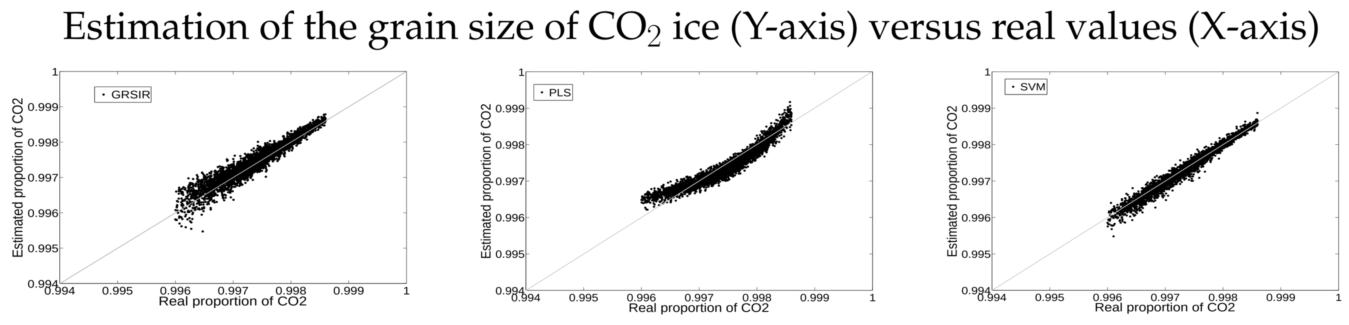
Results

- Competing methods: Gaussian Regularized Sliced Inverse Regression (GRSIR) [3] and Partial Least Squares (PLS) [2].
- Optimal parameters selected by cross-validation.

Parameter	GRSIR	PLS	SVM						
			lin.	Gauss.	Spect.	0-Pol	0.5-Pol	1-Pol	2-Pol
Prop. of H_2O	0.28	0.32	0.31	0.14	0.25	0.24	0.17	0.14	0.13
Prop. of CO_2	0.19	0.31	0.30	0.15	0.27	0.27	0.18	0.16	0.15
Prop. of dust	0.11	0.22	0.22	0.09	0.19	0.19	0.11	0.10	0.10
Grain size of H_2O	0.34	0.39	0.39	0.15	0.34	0.33	0.23	0.19	0.18
Grain size of CO_2	0.16	0.24	0.25	0.11	0.21	0.20	0.14	0.12	0.11
CPU time (s)	0.16	0.66	3.57	10.30	5.89	5.98	10.20	60.30	478

NRMSE and computing time for GRSIR, PLS and SVM with various kernels. "x-Pol" is $q = x$ in the polynomial kernel. The power of the polynomial kernel was fixed to 9 for each parameter, after cross-validation. The NRMSE quantifies the importance of estimation errors (must be close to zero). The bottom line of the table corresponds to the training time of parameter "Prop. of H_2O" after the selection of optimal hyperparameters.

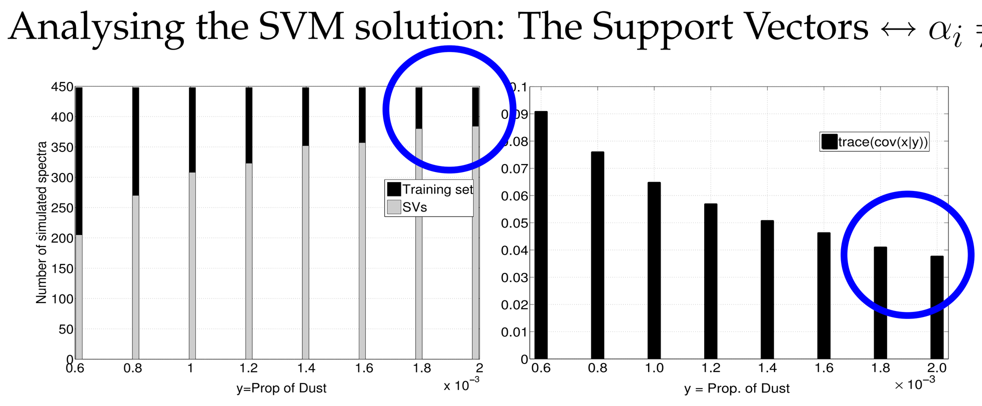
Estimation of the grain size of CO_2 ice (Y-axis) versus real values (X-axis)



GRSIR PLS SVM (Gaussian Kernel)

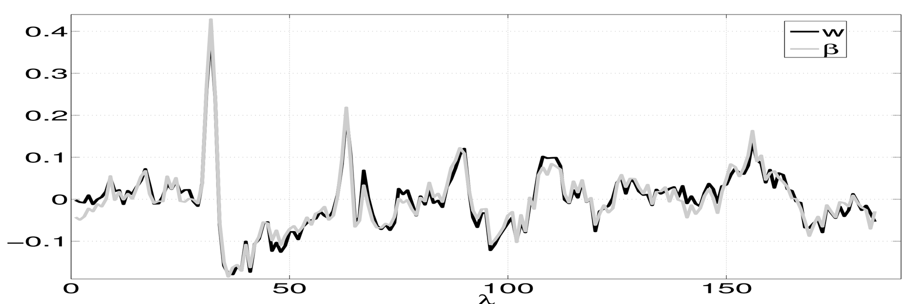
Comments

- SVM with Gaussian or polynomial kernel gives the best results in terms of Normalized Root Mean Square Errors (NRMSE) for all parameters
- Non-linear regression (GRSIR, Gaussian or polynomial SVM) performs better than linear regression (PLS and linear SVM).
- Training time is longer with SVM.
- Analysing the SVM solution: The Support Vectors $\leftrightarrow \alpha_i \neq 0$ in $f(\mathbf{x}) = \sum_{i=1}^{n} \alpha_i k(\mathbf{x}, \mathbf{x}_i) + b$



High number of SVs indicates that the estimation is difficult. The phenomenon is explained by the saturation of the physical model: different y generate very similar $\mathbf{x}$.

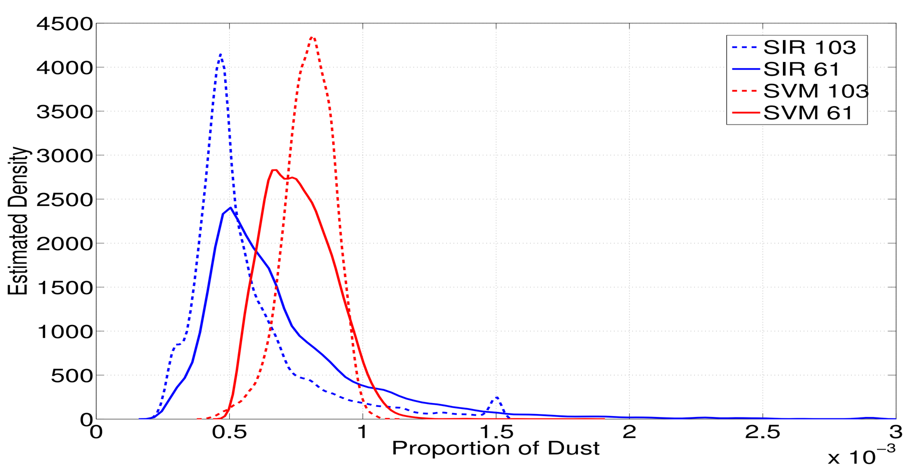
- Link between SVM with a linear kernel and GRSIR:



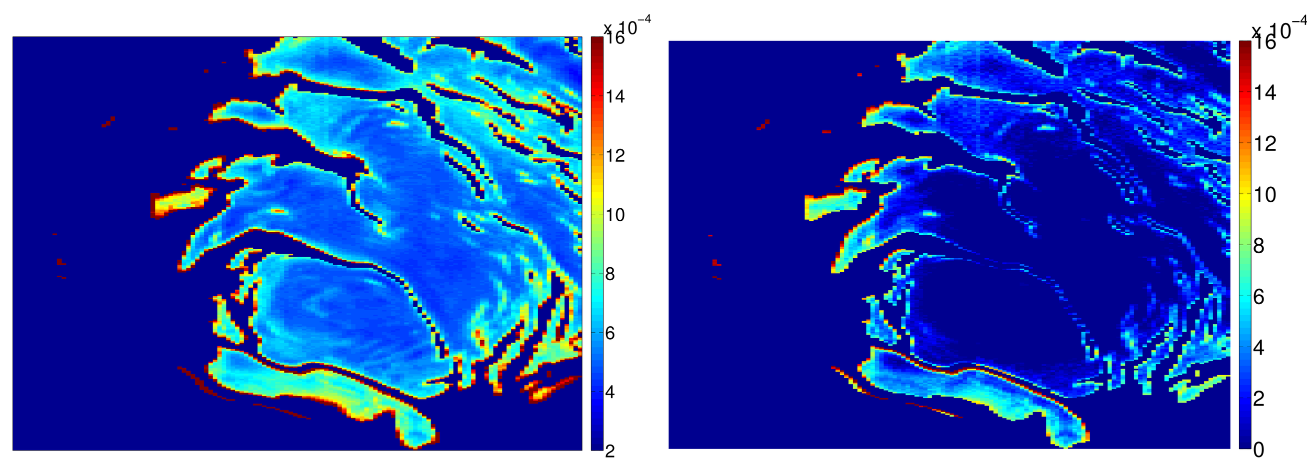
GRSIR axis β and SVM normal vector w as a function of λ

IV. Inversions of real hyperspectral images

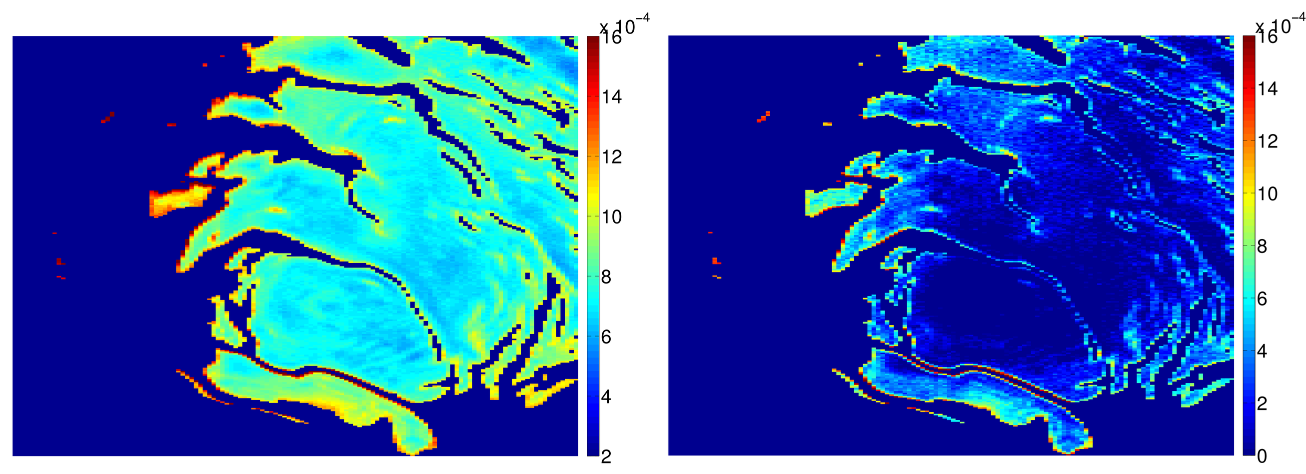
- Validation is difficult because no ground truth data is available.
- SVM estimations vary continuously and seem to be spatially coherent.
- SVM and GRSIR estimation are of different magnitude.
- Images from different orbits but analyzing the same portion of Mars does not give similar SVM estimates, unlike to GRSIR estimates.



Histogram of SVM and GRSIR estimates from two images of the same portion of Mars



GRSIR PLS



Gaussian SVM Linear SVM

Bibliography

[1] J-P. Bibring et al. (2004), Perennial water ice identified in the south polar cap of mars, *Nature*, 428:627–630.

[2] T. Hastie, R. Tibshirani, and J. Friedman (2001), *The elements of Statistical Learning*, Springer.

[3] C. Bernard-Michel, S. Douté, M. Fauvel, L. Gardes, and S. Girard (2007), Retrieval of Mars surface physical properties from OMEGA hyperspectral images using regularized sliced inverse regression, *Journal of Geophysical Research*, 2009.

Acknowledgment: *This work is supported by a contract with CNES through its Groupe Système Solaire Program and by INRIA and with the financial support of the "Agence Nationale de la Recherche" (French Research Agency) through its MDCO program ("Masse de Données et COnnaissances"). The Vahiné project was selected in 2007 under the reference ANR-07-MDCO-013.*