

# Estimation du taux d'Argile en condition semivégétalisée par traitement d'images Hyperspectrales

Ouerghemmi <sup>1</sup> W., Gomez <sup>2</sup> C., Nacer <sup>1</sup> S., Lagacherie <sup>3</sup> P.

- <sup>1</sup> LTSIRS ENIT, Laboratoire de Télédétection et Systèmes d'Information à Référence Spatiale, Tunis, **Tunisie**
- <sup>2</sup> IRD UMR LISAH, Laboratoire d'étude des Interactions Sols Agrosystèmes Hydrosystèmes, Montpellier, France
- <sup>3</sup> INRA UMR LISAH, Laboratoire d'étude des Interactions Sols Agrosystèmes Hydrosystèmes, Montpellier, France



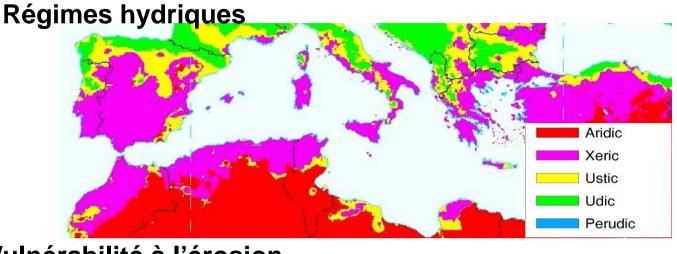




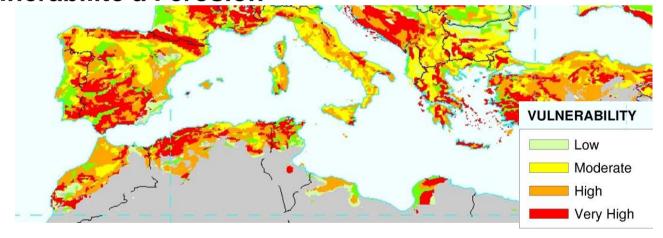




# Vers des cartes mondiales de propriétés fonctionnelles des sols



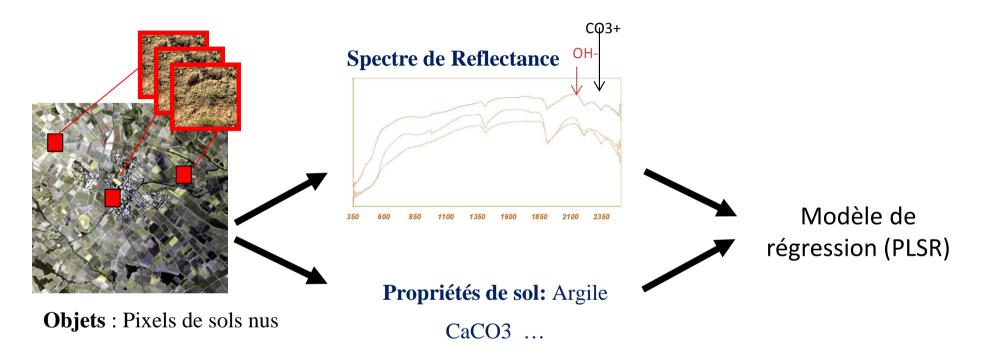




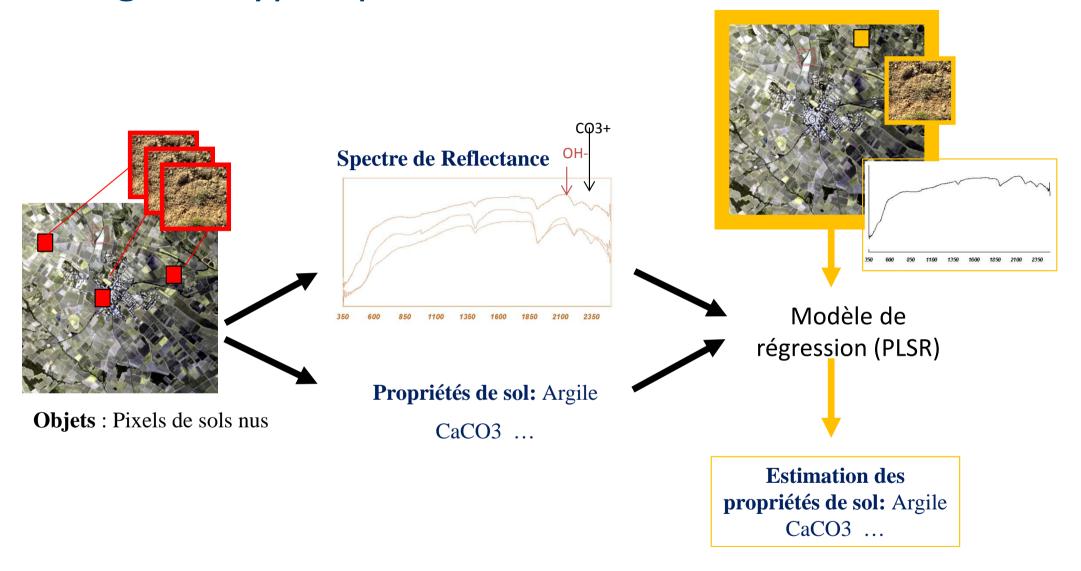
Sources USDA-NRCS, http://soils.usda.gov/use/worldsoils/mapindex

- → Une limite majeure : disponibilités en données sol « de base »
- → Un enjeu scientifique: développer de nouvelles méthodes de spatialisation des propriétés de sol

## Imagerie hyperspectrale Vis-NIR en science du sol

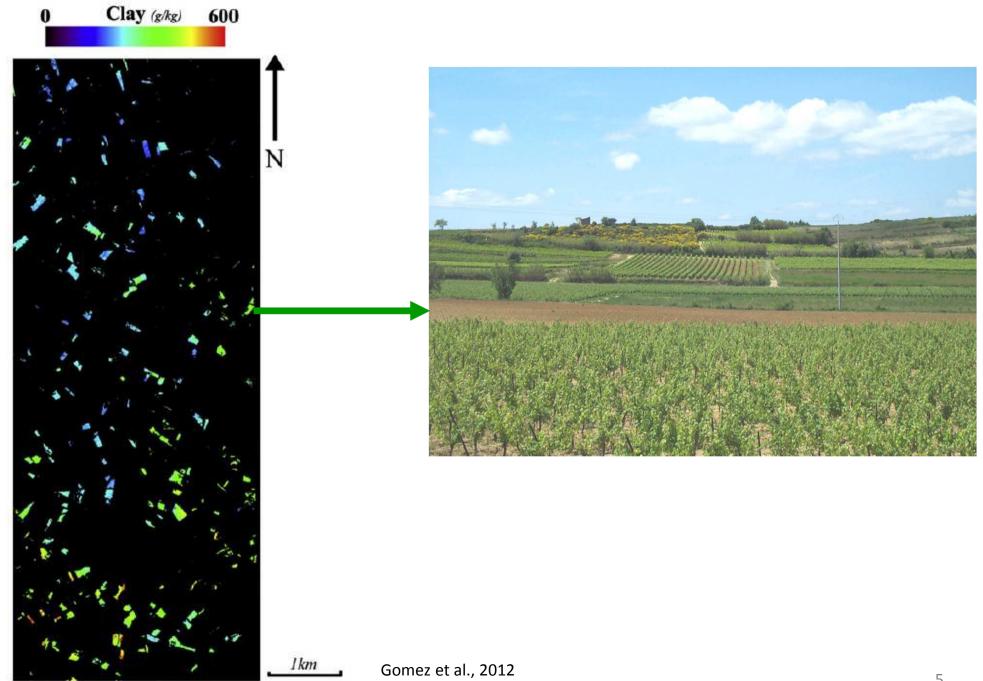


# Imagerie hyperspectrale Vis-NIR en science du sol

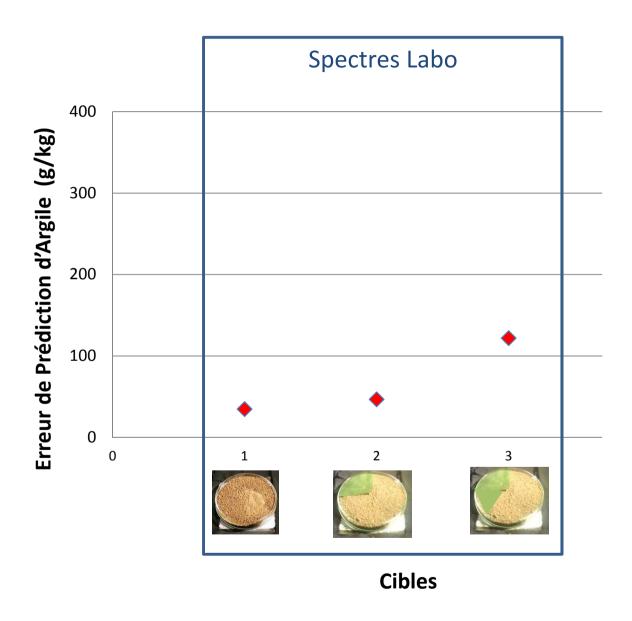


Stevens et al. 2007, Gomez et al. 2008, Selige et al. 2008, Gomez et al. 2012 etc ...

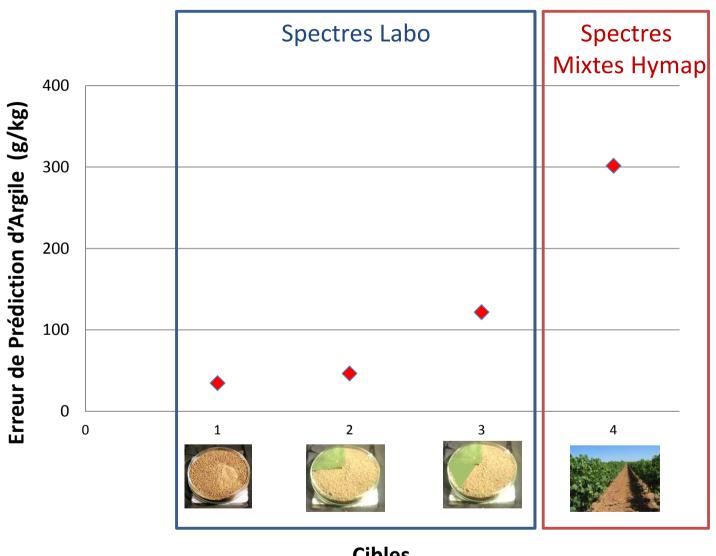
# Imagerie hyperspectrale Vis-NIR en science du sol



# Impact de la végétation

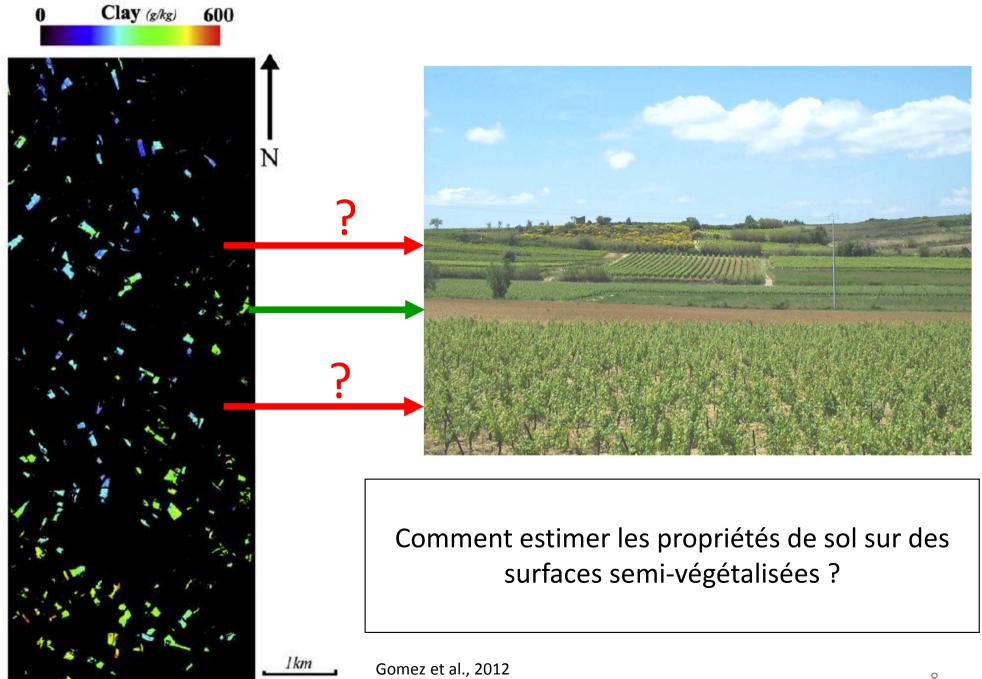


## Impact de la végétation

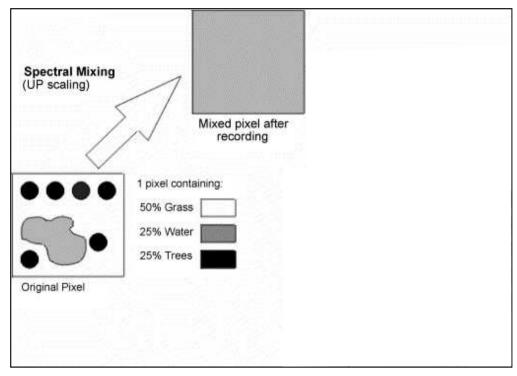


**Cibles** 

## Problématique

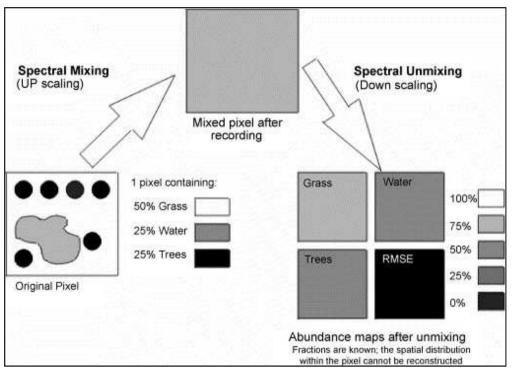


# Etat de l'art en contexte « mixte » : Unmixing ?



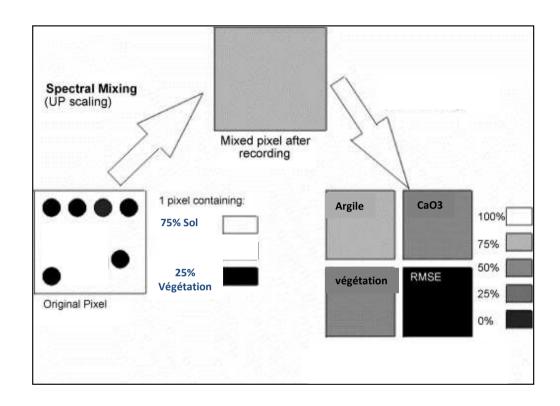
Smith et al. (1985)

## Etat de l'art en contexte « mixte » : Unmixing ?



Smith et al. (1985)

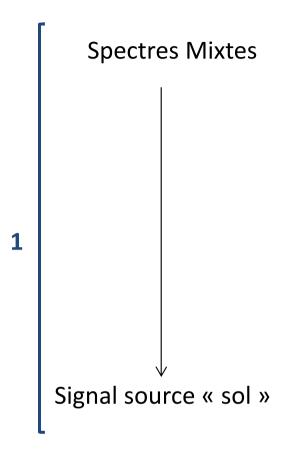
## Etat de l'art en contexte « mixte » : Unmixing ?

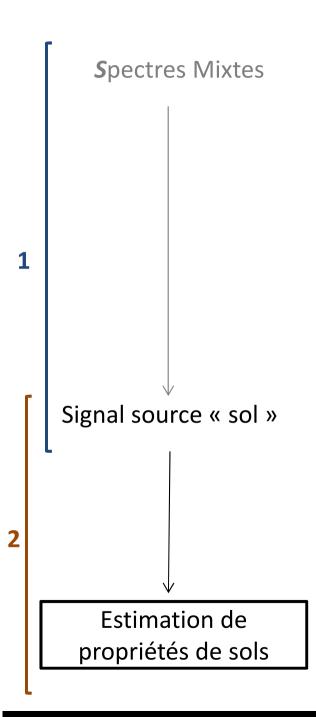


Rechercher une propriété du sol ⇔ Rechercher un sous-composant du sol

Or Impossible d'avoir un spectre référence (« endmembers ») pour l'Argile, le CaCO3, le Carbone organique ...

⇒ Impossible d'utiliser les méthodes d'unmixing, nécessitant des endmembers.





**N** spectres Mixtes Séparation Aveugle des sources (BSS) 1 **M** signaux sources Identification du signal source « sol » Signal source « sol »

- x(t)=vecteur des **N** signaux mélangés
- s(t)=vecteur des M signaux sourcesx(t)=A.s(t)

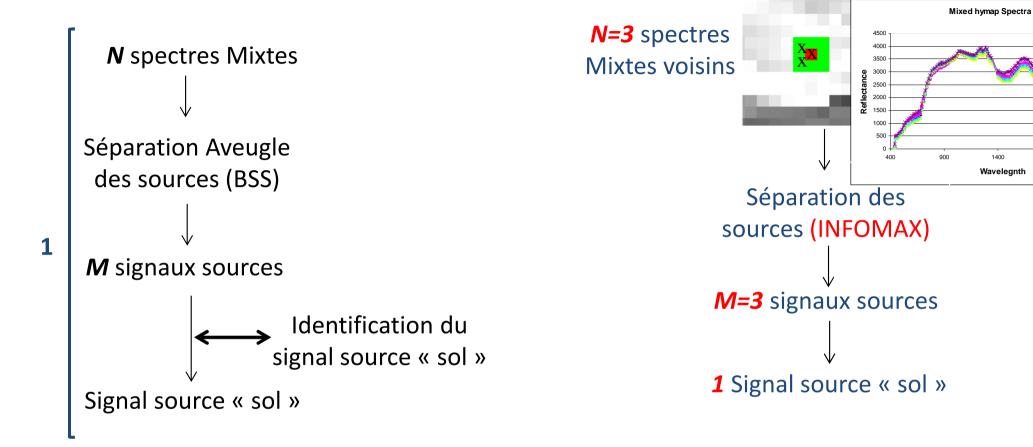
L'objectif de la BSS est de retrouver une estimation des signaux sources Ŝ, sans connaissance a priori des abondances A, tel que :

$$\hat{S}(t) = W \cdot x(t)$$

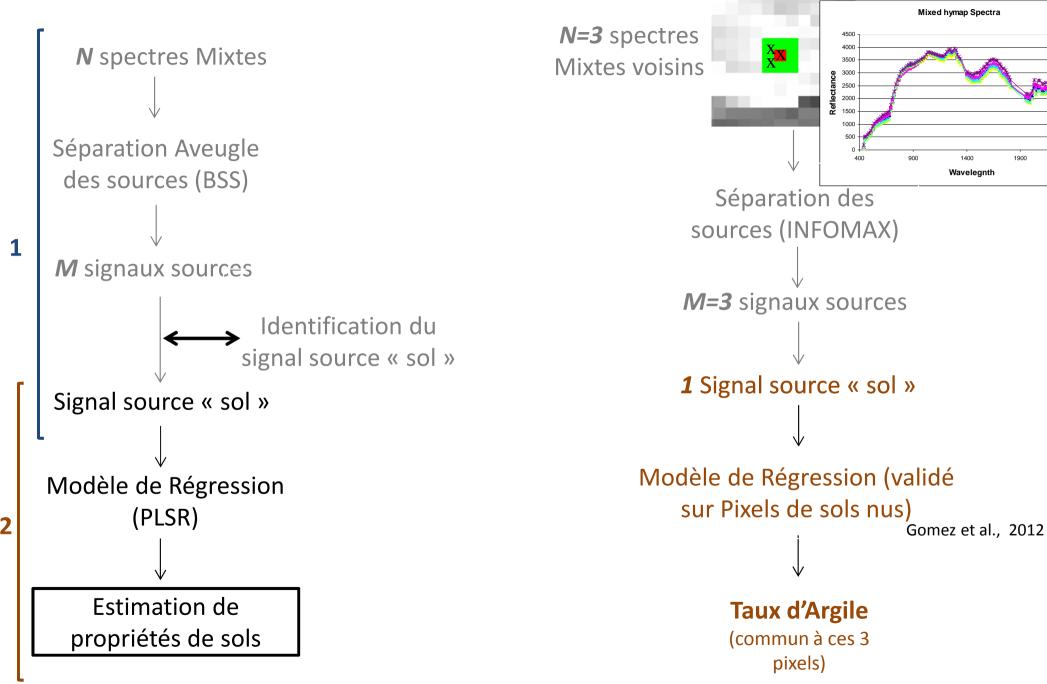
Herault and Jutten, 1982

Algorithme BSS utilisé: INFOMAX, basé sur un réseau de neurones

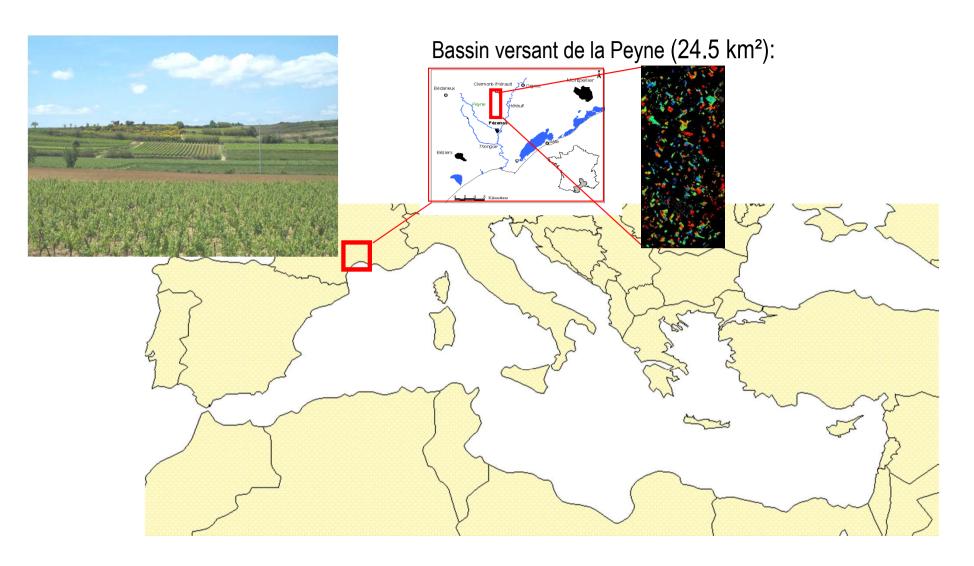
(Bell and Sejnowski, 1995)



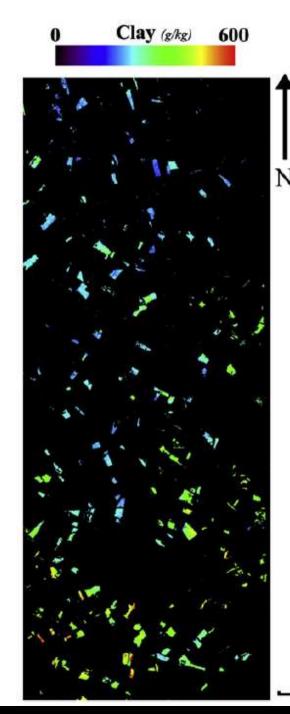
Wavelegnth



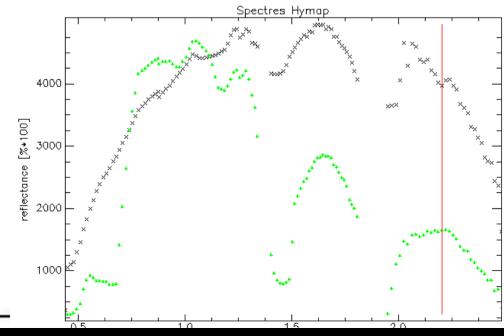
### Terrain d'étude



## Données hyperspectrales aéroportées HyMap



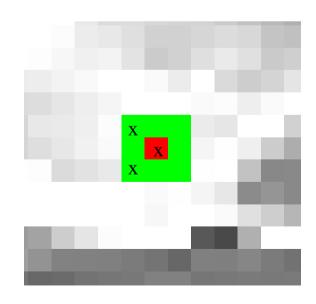
- Domaine spectral: 400-2500 nm
- Nombre de bandes spectrales utilisées : 114
- Résolution spatiale : 5 m
- 100:1 < SNR < 500:1
- Corrections atmosphériques avec ATCOR 4
- 4% de la zone couverte par des sols nus

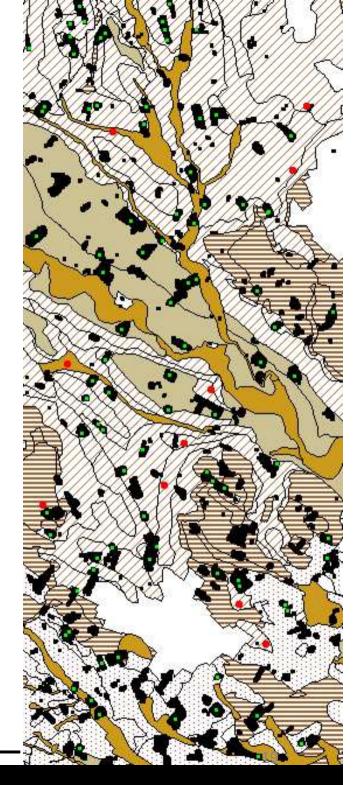


1km

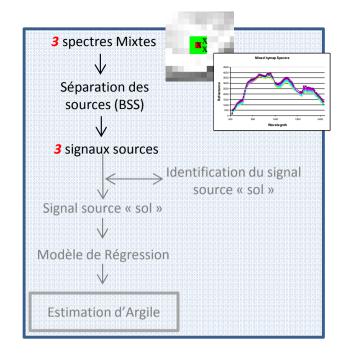
#### Données

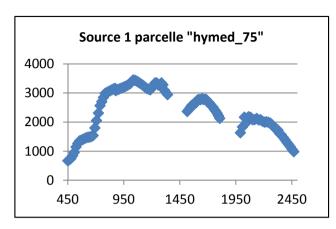
- <u>49 échantillons de sol</u> récoltés en 2009 sur des surfaces couvertes par un <u>mélange [sol & vigne]</u> au moment de l'acquisition de l'image. *0,2 < NDVI < 0,72*
- Analyse physico-chimique d'<u>Argile</u> (Labo ARRAS)

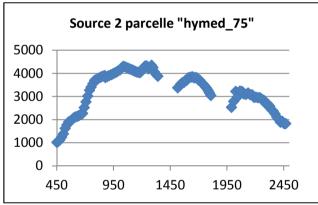


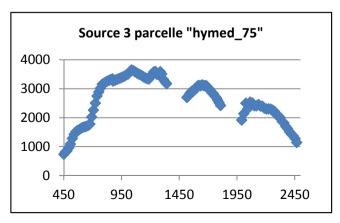


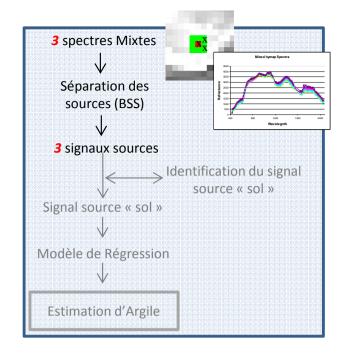
<u>1km</u>

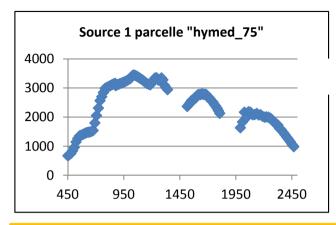




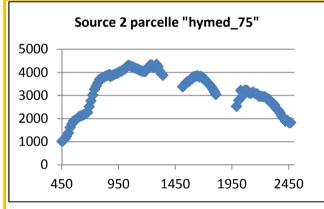


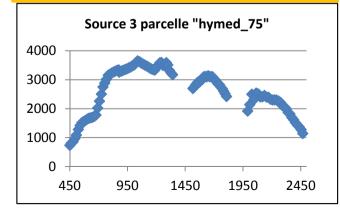




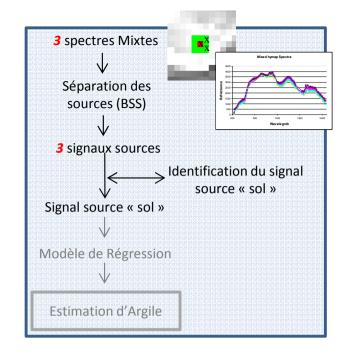


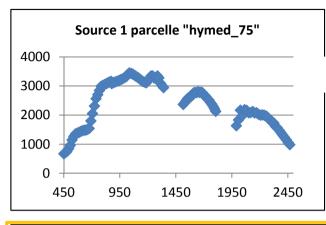
$$R^2_{avec « Sol moyen »} = 0.6$$



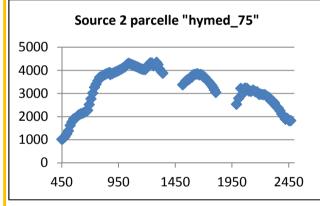


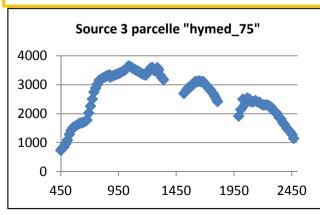
$$R^2_{avec « Sol moyen »} = 0.65$$



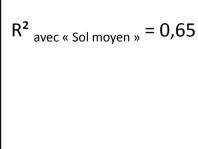


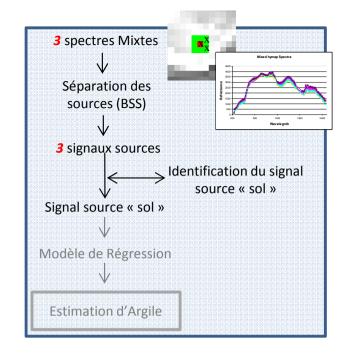
$$R^2_{avec « Sol moyen »} = 0.6$$

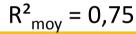


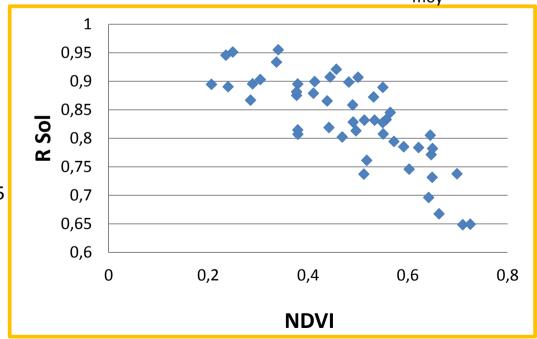


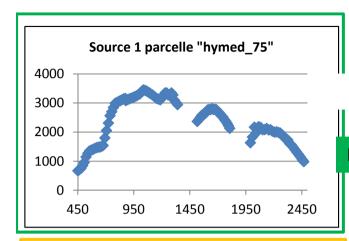






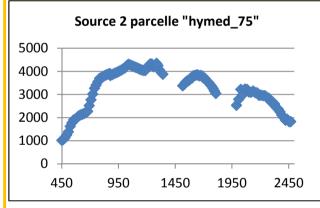




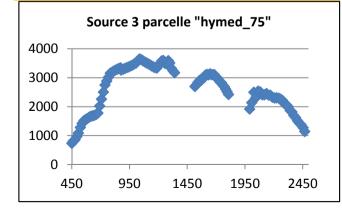




$$R^2$$
 avec « Vigne Verte » = 0,92

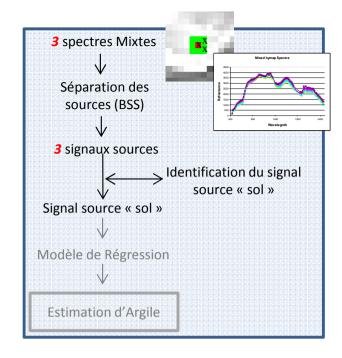


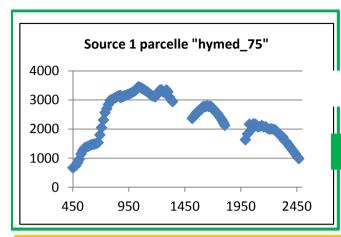
$$R^2_{avec ext{ w Vigne Verte }} = 0.84$$



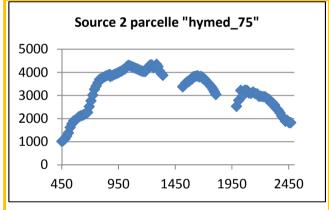
$$R^2_{avec \ll Sol moyen } = 0.65$$

$$R^2_{avec « Vigne Verte »} = 0.88$$



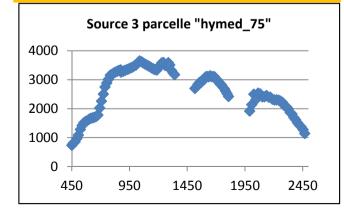


$$R^2_{avec « Sol moyen »} = 0,6$$

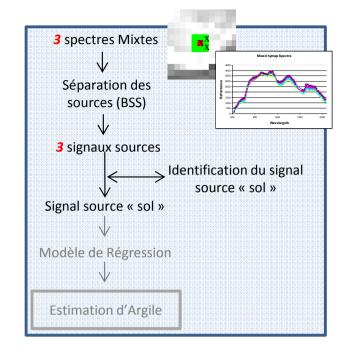


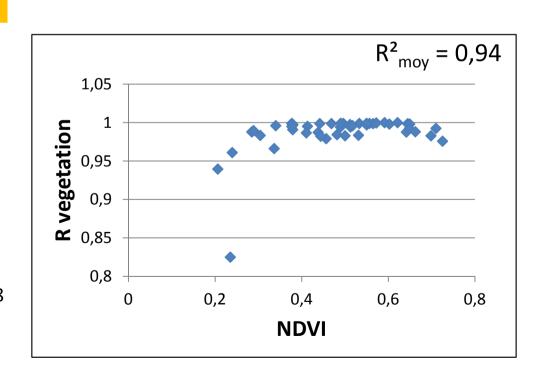


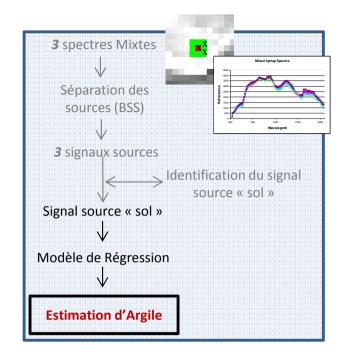


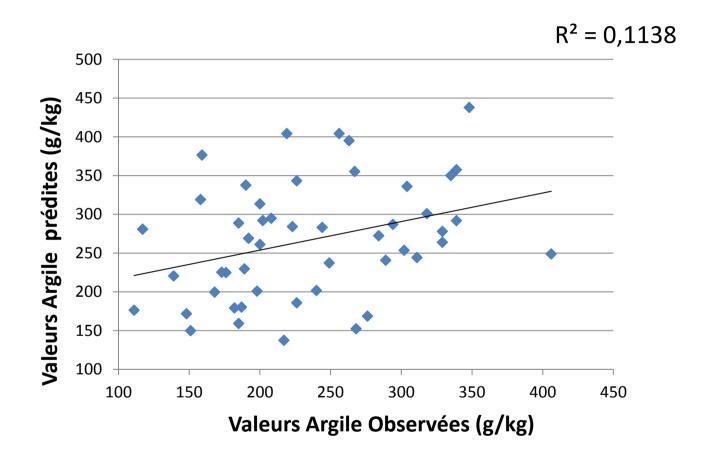


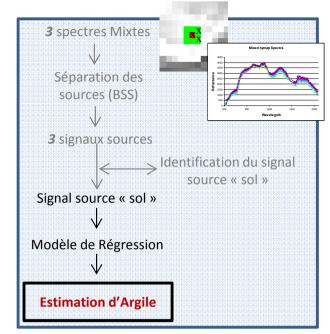
$$R^2_{avec \ll Sol moyen } = 0.65$$

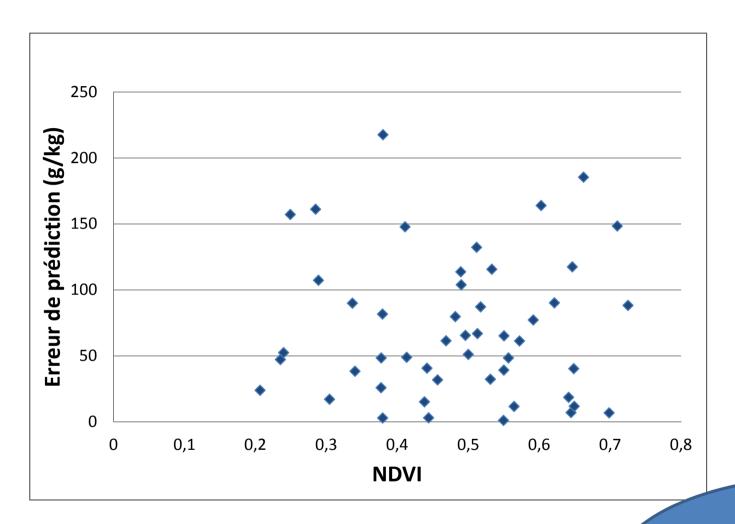


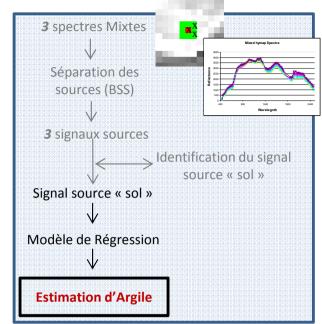




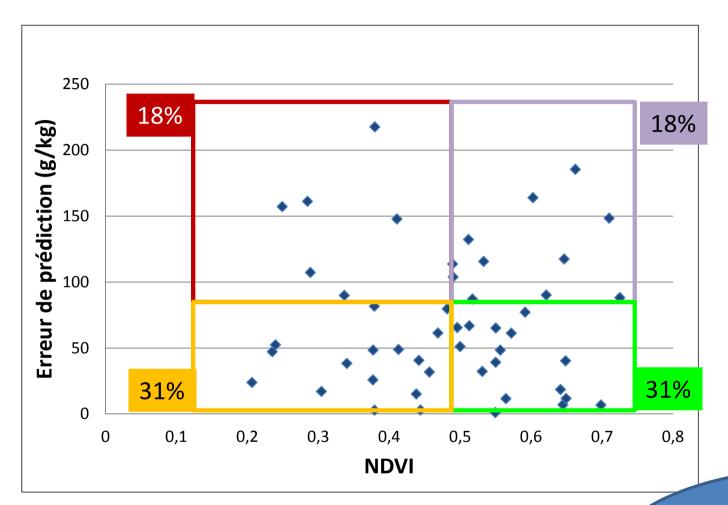


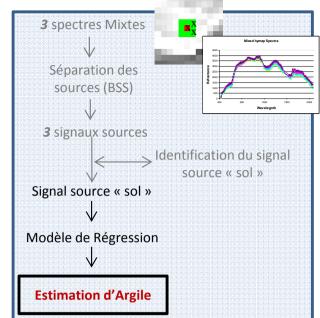






Erreur moyenne de prédiction sur les 49 parcelles : 70 g/kg





=> Pas de relation entre l'erreur de prédiction et le NDVI dans le mélange

Erreur moyenne de prédiction sur les 49 parcelles : 70 g/kg

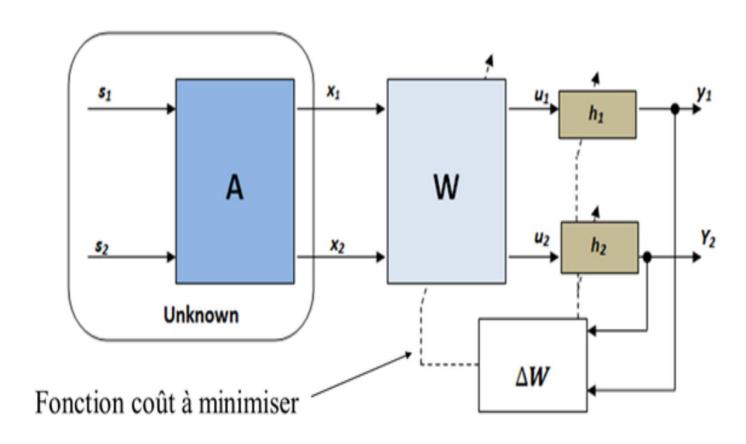
# Conclusion

- Extraction du signal sol prometteur avec le concept BSS :
  - Signal sol extrait après BSS avec un  $R^2 = 0.75$
  - Signal végétation extrait après BSS avec un R² = 0,94
- Résultats décevants pour la prédiction d'argile après BSS (R²=0.1)
- Néanmoins encourageants : 62% des 49 parcelles tests ont une erreur de prédiction < 80g/kg</li>

# Conclusion

- Extraction du signal sol prometteur avec le concept BSS :
  - Signal sol extrait après BSS avec un  $R^2 = 0.75$
  - Signal végétation extrait après BSS avec un R² = 0,94
- Résultats décevants pour la prédiction d'argile après BSS (R²=0.1)
- Néanmoins encourageants : 62% des 49 parcelles tests ont une erreur de prédiction < 80g/kg</li>
  - ⇒ Comprendre pourquoi certain spectres de pixels mixtes sont plus difficiles à séparer puis prédire.
    - » Hétérogénéité de la surface ?
    - » Végétation différente ?
    - » Sol particulier ?
  - ⇒ Importance du choix du signal «sol» en sortie de BSS.
  - ⇒ Amélioration de la paramétrisation d'INFOMAX
  - ⇒ Tests sur d'autres algorithmes : Non-negative matrix factorization ?

# Algorithme: Infomax

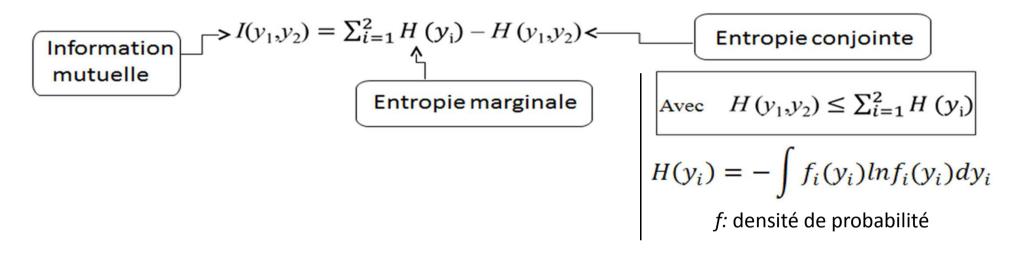


- •INFOMAX est un algorithme basé sur un réseau de neurones
- •La sortie u du réseau est transformée par une fonction non-linèaire h tel que y = h(u)
- •Une fonction coût mesurant l'indépendance statistique de la sortie **y** est optimisée

Objectif: minimiser l'information mutuelle entre les sorties y du réseau de neurones

#### Infomax

Le fait de minimiser l'information mutuelle entre les signaux à la sortie du réseau assurera l'indépendance de ces signaux



minimiser l'information mutuelle revient à maximiser l'entropie conjointe, l'optimisation sera réalisée en utilisant l'algorithme du gradient stochastique

Pour déduire l'équation d'apprentissage finale du réseau de neurones nous allons maximiser l'entropie conjointe

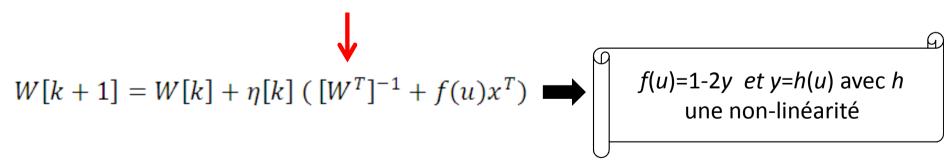
# Infomax

La maximisation (ou bien minimisation) d'une fonction coût  $\mathcal{L} \{ \Phi \}$  par rapport à un paramètre  $\Phi$  est obtenue en appliquant la méthode du gradient stochastique à l'itération (l+1)

$$\Phi(I+1) = \Phi(I) + \eta_{\Phi} \frac{\partial L \{\Phi(I)\}}{\partial \Phi} = \Phi(I) + \eta_{\Phi} \Delta \Phi(I)$$

ηφ est Le pas d'apprentissage

Dans notre cas la fonction coût à maximiser est l'entropie conjointe H(y) et le paramètre est W. Ceci nous amènera à l'équation d'apprentissage finale du réseau de neurones :



# Paramètres

 Learning rate (facteur d'actualisation ou pas d'apprentissage) 0.02

Nombre d'itérations maximales 50

 Fonction quadratique non-linéaire à intégrer dans l'équation d'apprentissage sigmoidale (f(x)=1/(1+e-x))