

Analyse des incertitudes associées aux prédictions de la teneur en argile obtenues par imagerie hyperspectrale Vis-NIR aéroportée (0.4-2.5 μ m)

Gomez C. ¹, Drost A. ^{1 2}, Roger J-M³

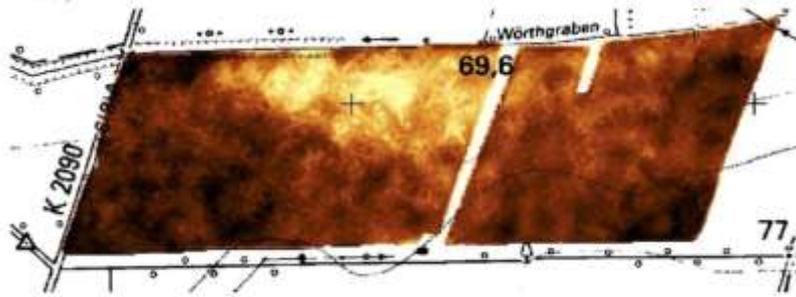
¹ IRD, UMR LISAH (INRA-IRD-SupAgro), F-34060 Montpellier, France (cecile.gomez@ird.fr)

² Centre for Geo-Information, Wageningen University, 6708 PB Wageningen, The Netherlands

³ IRSTEA, UMR ITAP, Montpellier, France

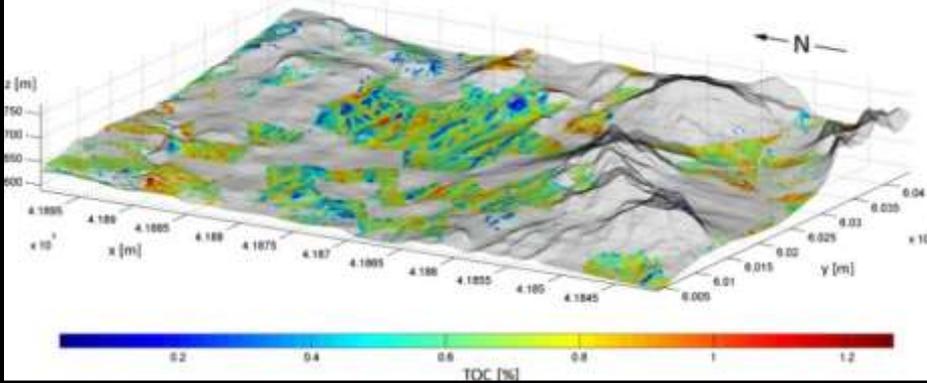
Vis-NIR spectroscopy for soil surface property mapping

Sand content with HyMap data (Selige et al., 2006)



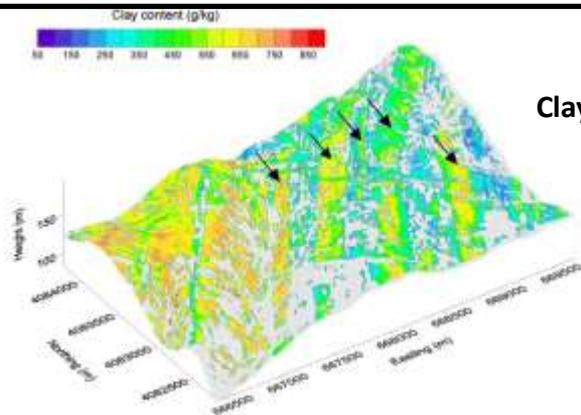
7km²

Organic carbon with Hymap data (Schwanghart & Jarmer, 2011)



50km²

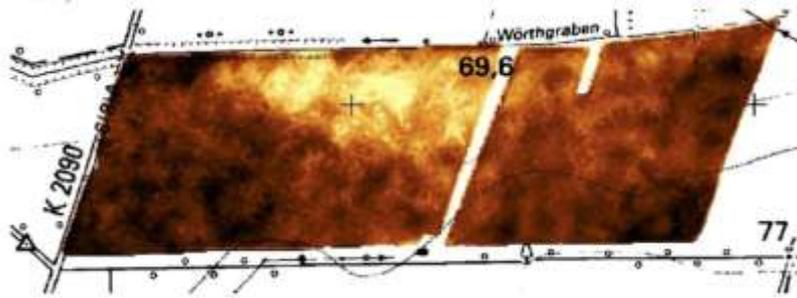
Clay content with AISA-Dual data (Gomez et al, 2012)



300 km²

Vis-NIR spectroscopy for soil surface property mapping

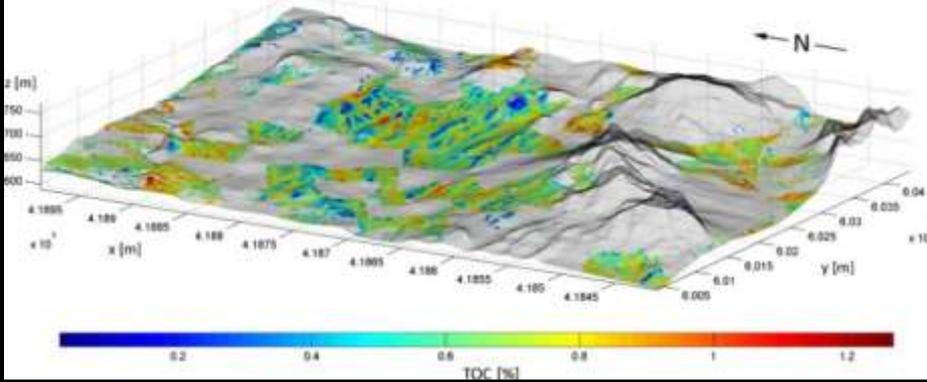
Sand content with HyMap data (Selige et al., 2006)



72 échantillons
dans
BD_calibration

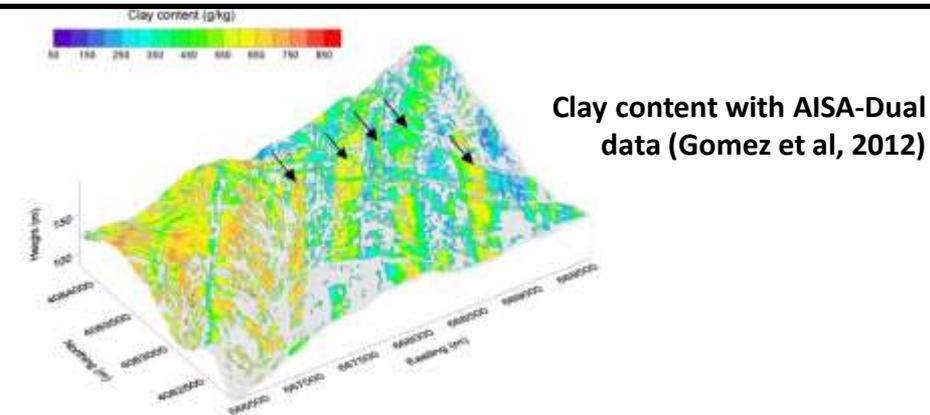
7km²

Organic carbon with Hymap data (Schwanghart & Jarmer, 2011)



61 échantillons
dans
BD_calibration

50km²



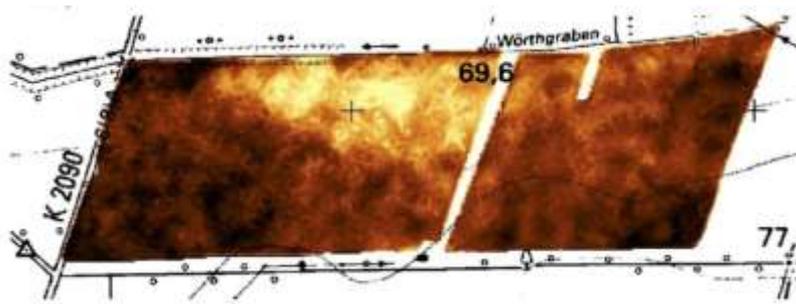
Clay content with AISA-Dual
data (Gomez et al, 2012)

96 échantillons dans
BD_calibration

300 km²

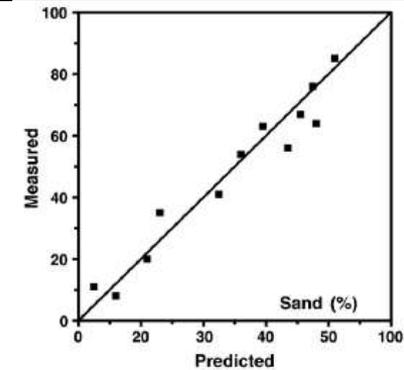
Vis-NIR spectroscopy for soil surface property mapping

Sand content with HyMap data (Selige et al., 2006)

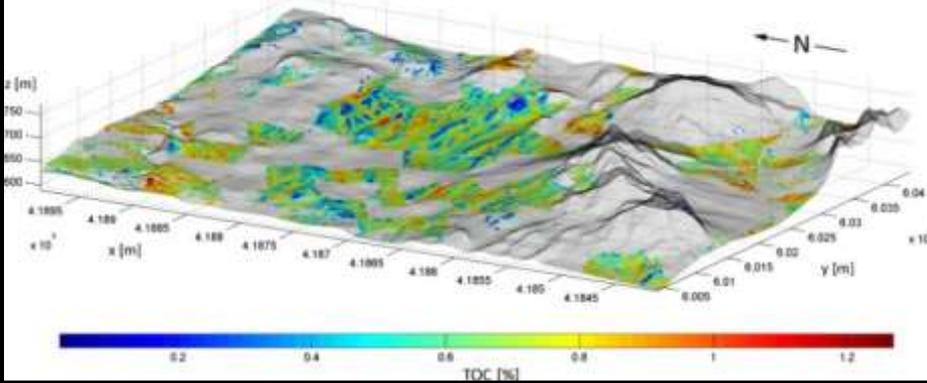


72 échantillons
dans
BD_calibration

7km²

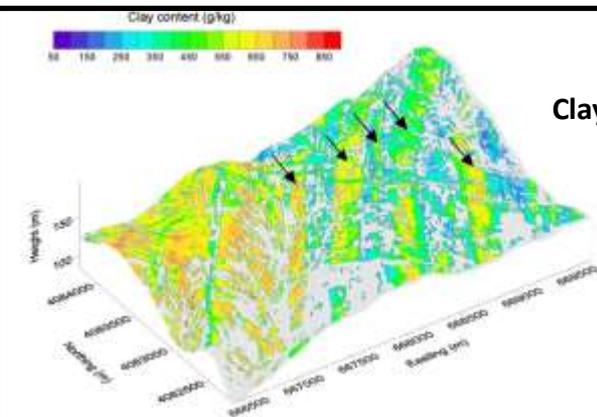
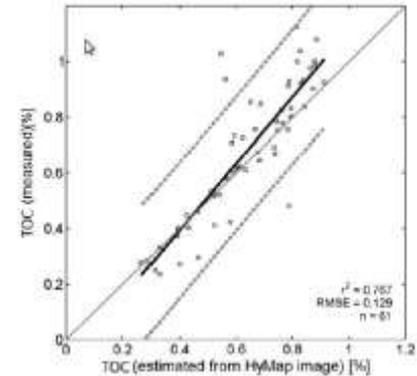


Organic carbon with Hymap data (Schwanghart & Jarmer, 2011)



61 échantillons
dans
BD_calibration

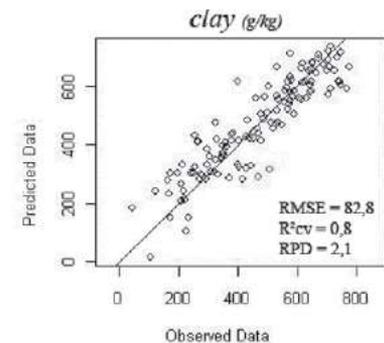
50km²



Clay content with AISA-Dual
data (Gomez et al, 2012)

96 échantillons dans
BD_calibration

300 km²



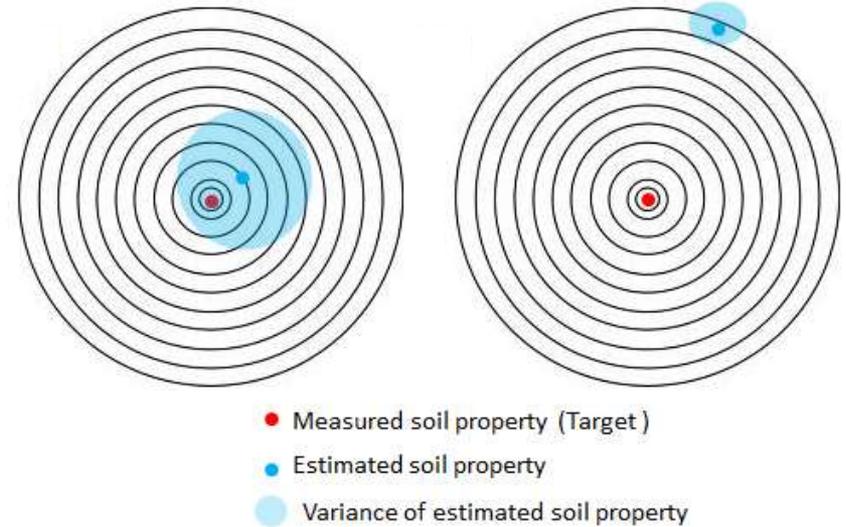
Erreur et incertitude

Erreur de prédiction (Le Graal !!!) :

« Distance entre la valeur prédite et la valeur vraie »

Incertitude de prédiction :

« Variance des prédictions »



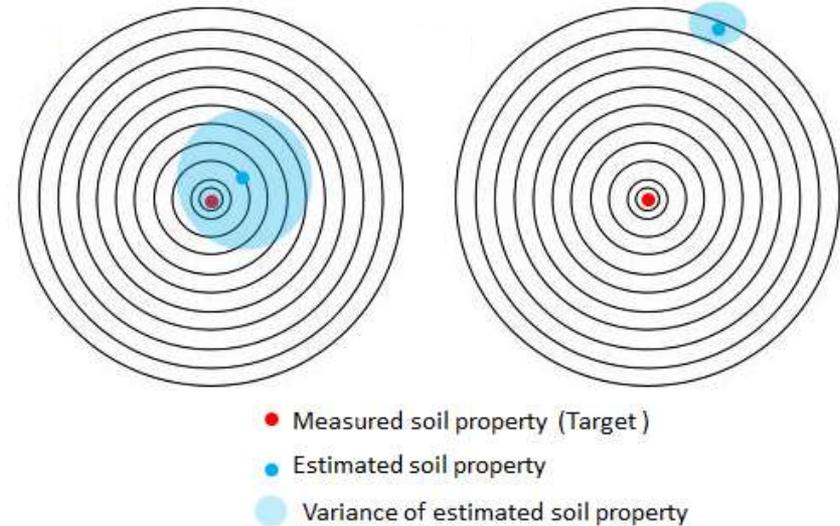
Erreur et incertitude

Erreur de prédiction (Le Graal !!!) :

« Distance entre la valeur prédite et la valeur vraie »

Incertitude de prédiction :

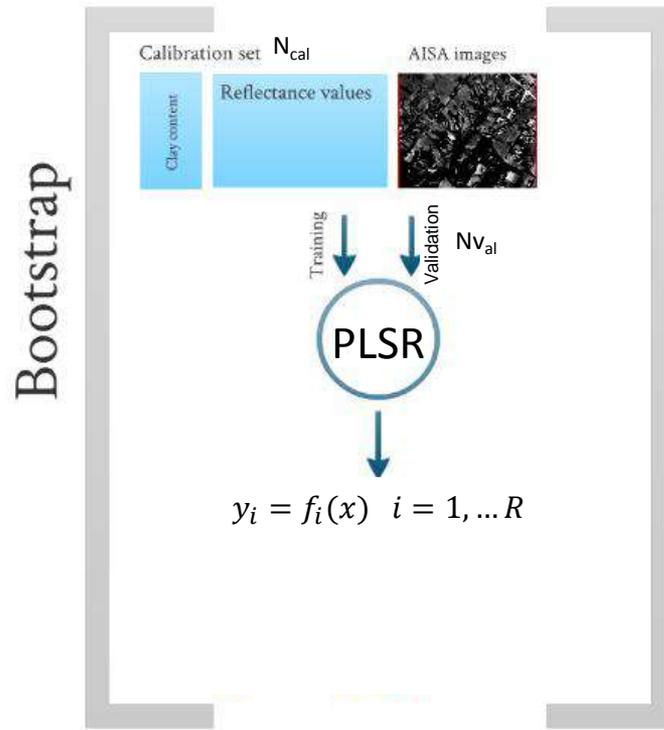
« Variance des prédictions »



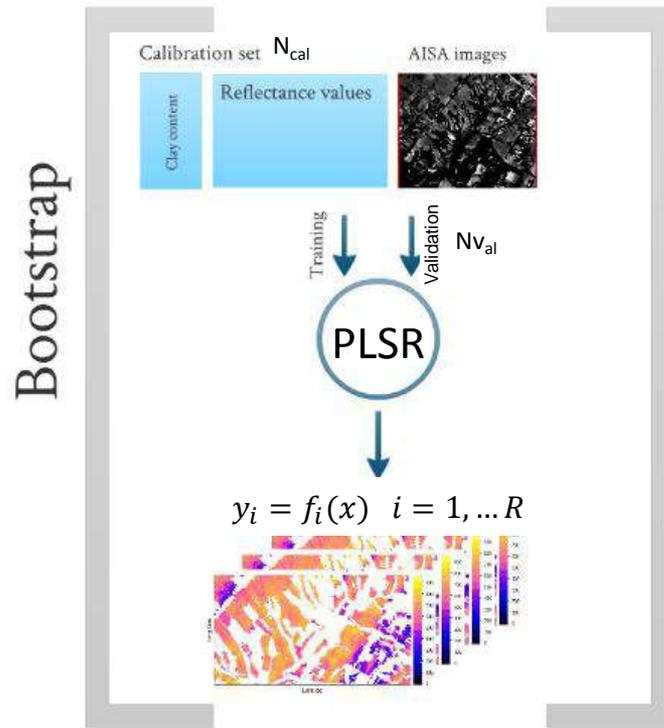
Objectifs :

- Analyser les performances des modèles PLSR, en étudiant l'incertitude de prédiction
- Observer le lien éventuel entre l'incertitude et l'erreur

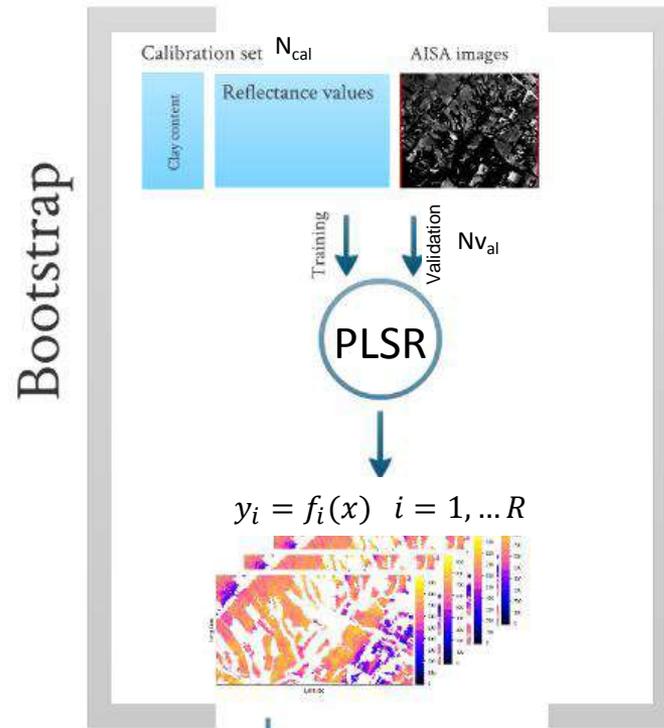
Calcul d'Incertitudes - Démarche



Calcul d'Incertitudes - Démarche

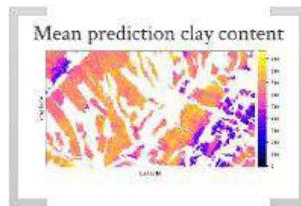


Calcul d'Incertitudes - Démarche

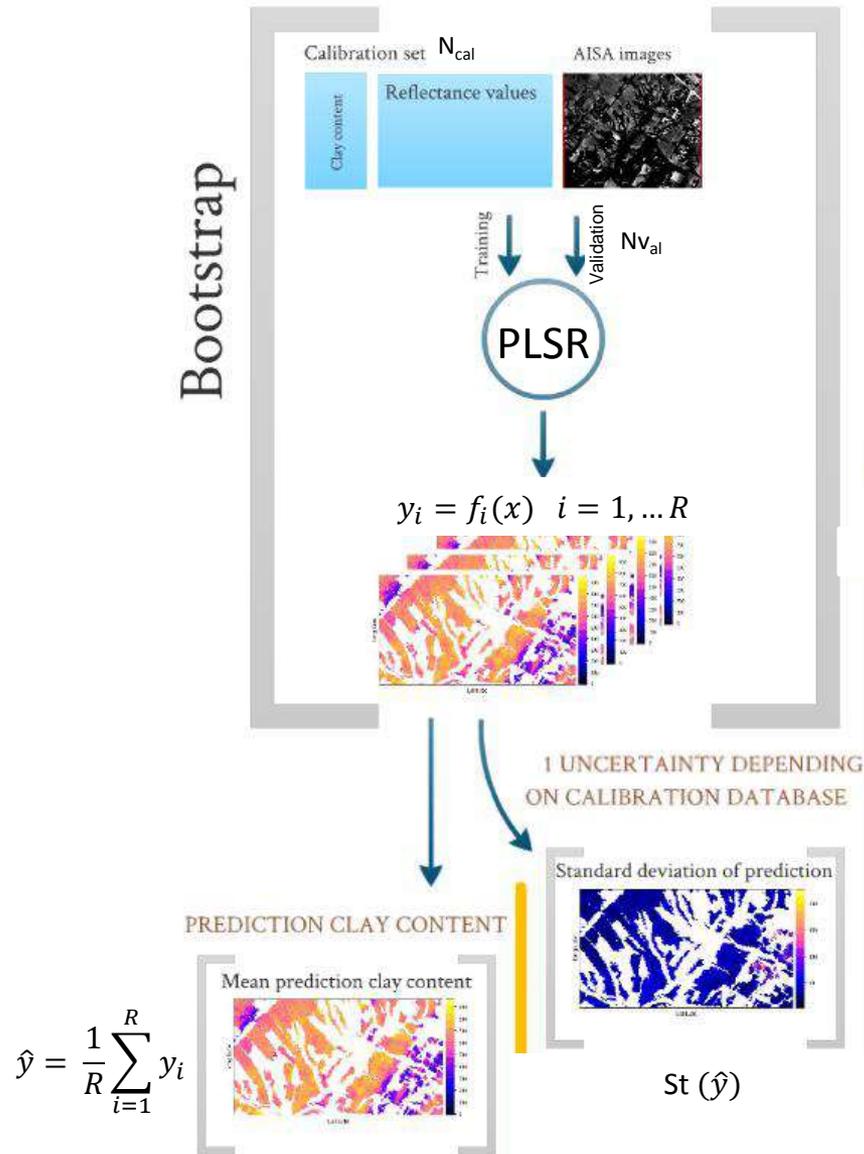


PREDICTION CLAY CONTENT

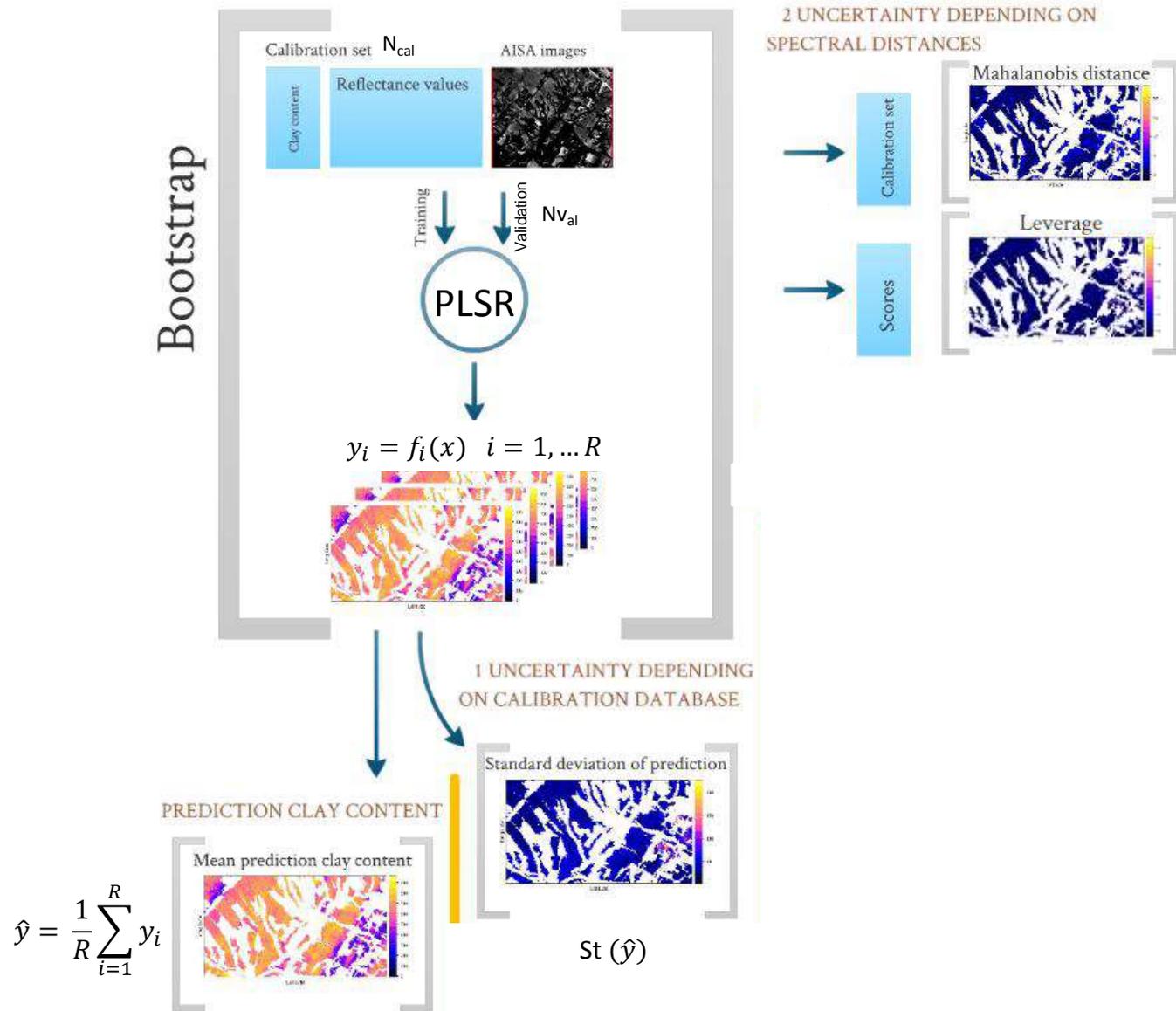
$$\hat{y} = \frac{1}{R} \sum_{i=1}^R y_i$$



Calcul d'Incertitudes - Démarche

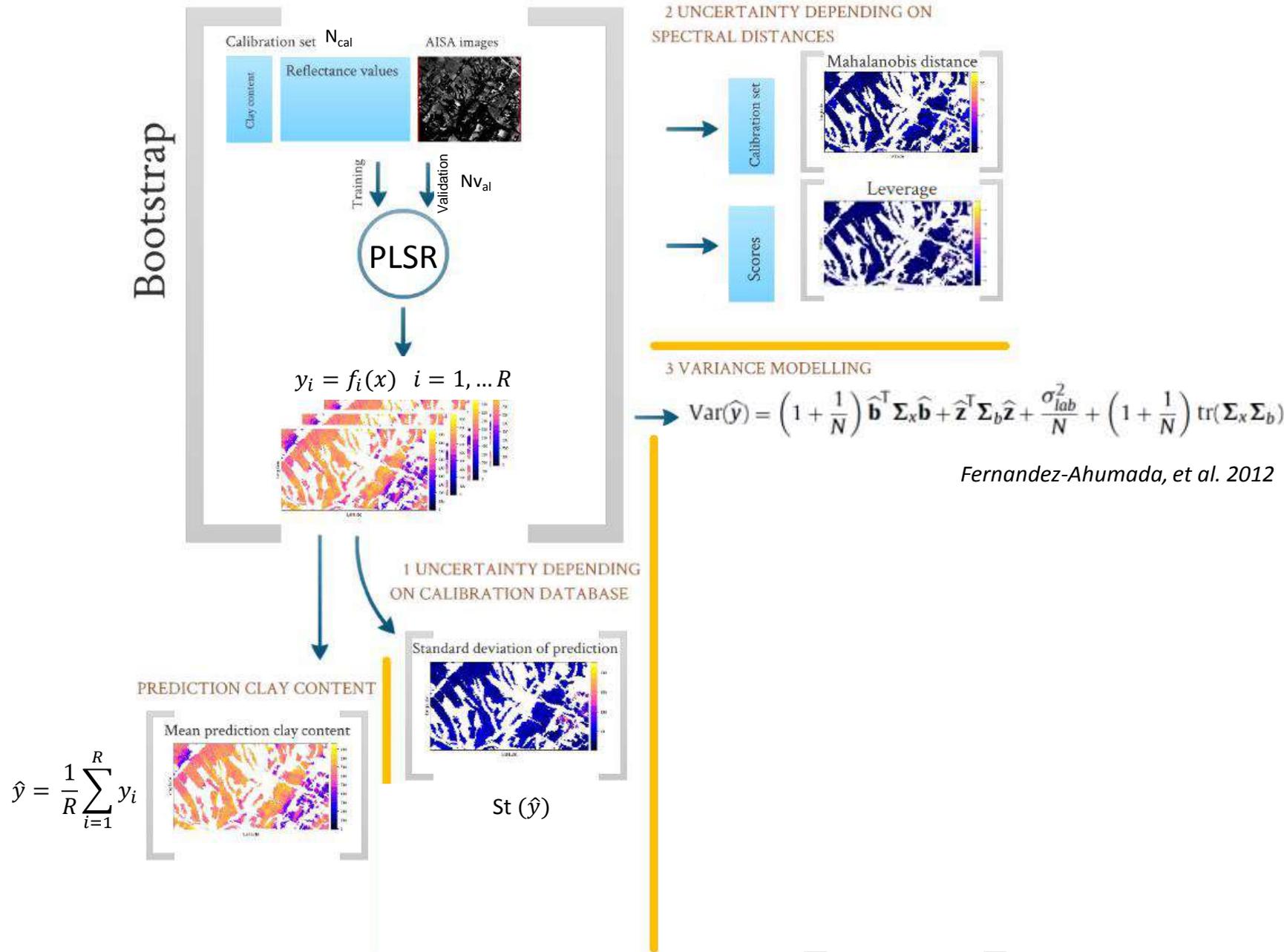


Calcul d'Incertitudes - Démarche

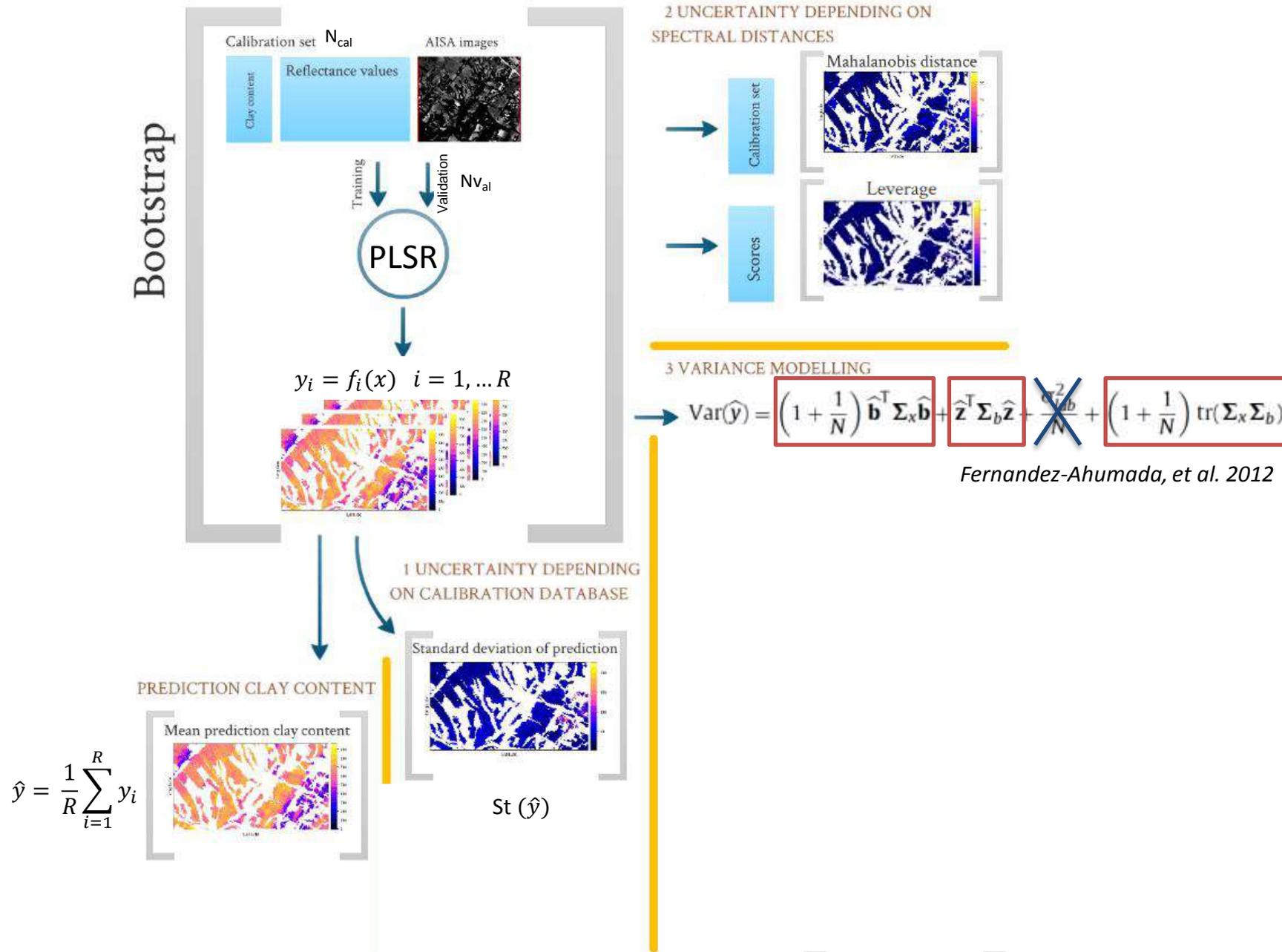


$$\hat{y} = \frac{1}{R} \sum_{i=1}^R y_i$$

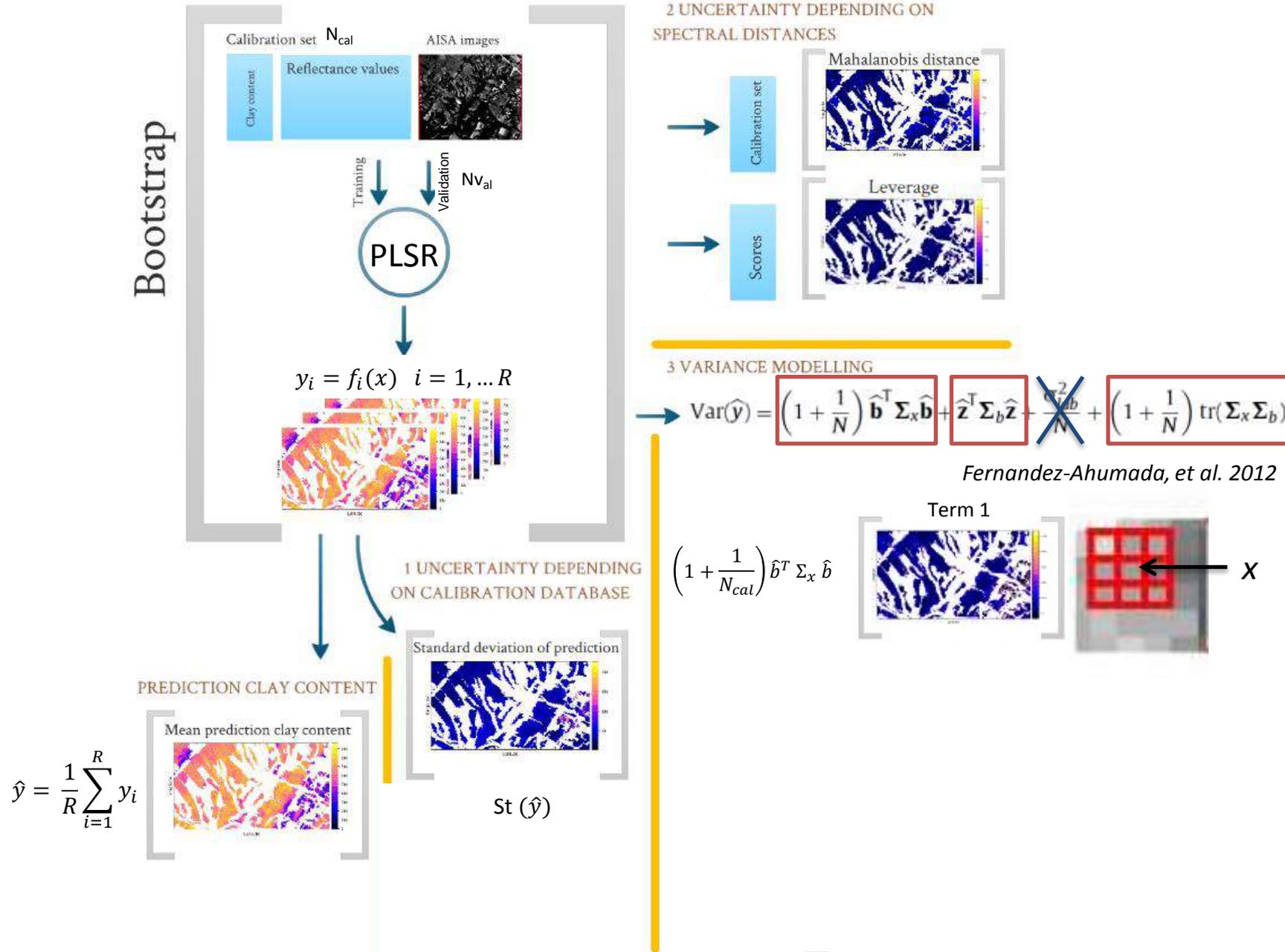
Calcul d'Incertitudes - Démarche



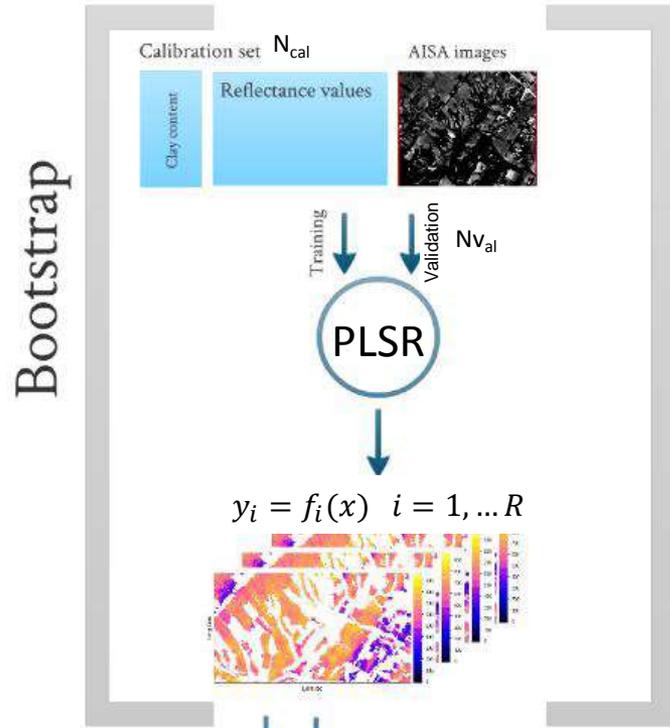
Calcul d'Incertitudes - Démarche



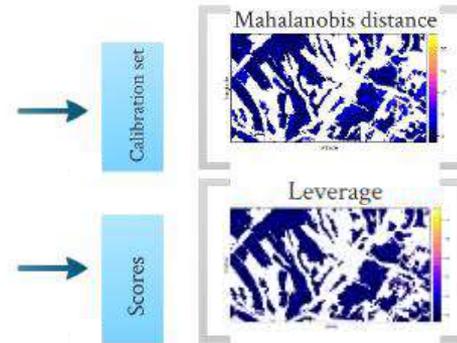
Calcul d'Incertitudes - Démarche



Calcul d'Incertitudes - Démarche



2 UNCERTAINTY DEPENDING ON SPECTRAL DISTANCES

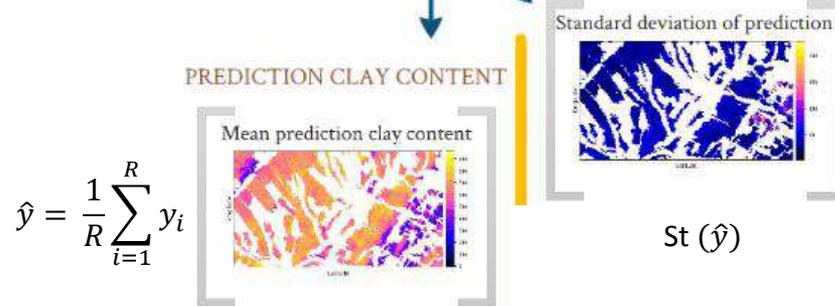


3 VARIANCE MODELLING

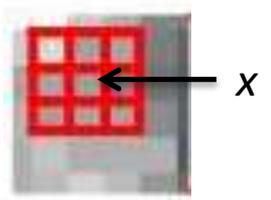
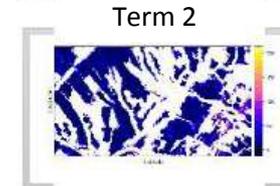
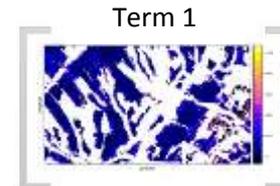
$$\text{Var}(\hat{y}) = \left(1 + \frac{1}{N}\right) \hat{\mathbf{b}}^T \Sigma_x \hat{\mathbf{b}} + \hat{\mathbf{z}}^T \Sigma_b \hat{\mathbf{z}} + \frac{\sigma_b^2}{N} + \left(1 + \frac{1}{N}\right) \text{tr}(\Sigma_x \Sigma_b)$$

Fernandez-Ahumada, et al. 2012

1 UNCERTAINTY DEPENDING ON CALIBRATION DATABASE

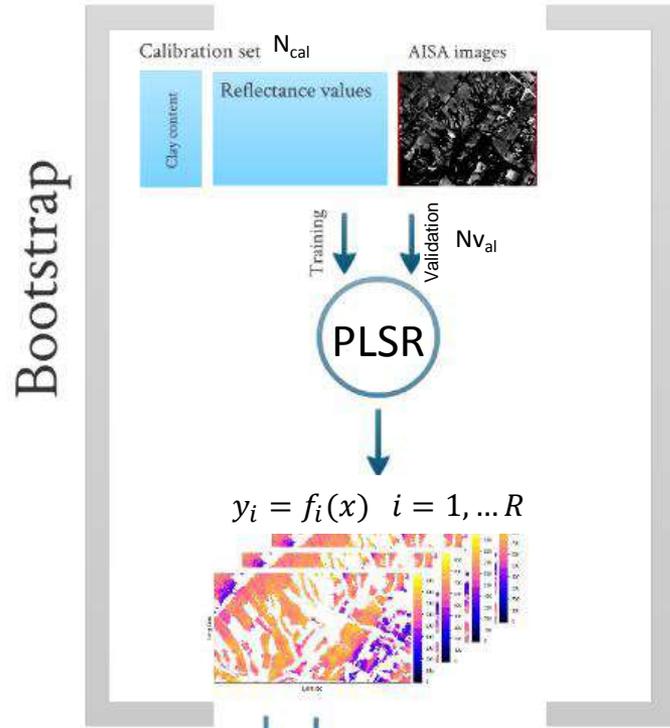


$$\left(1 + \frac{1}{N_{cal}}\right) \hat{\mathbf{b}}^T \Sigma_x \hat{\mathbf{b}}$$

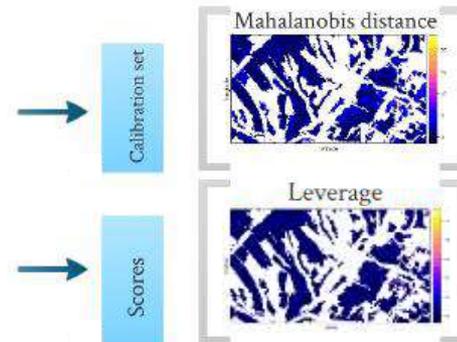


$$\hat{\mathbf{z}}^T \Sigma_b \hat{\mathbf{z}}$$

Calcul d'Incertitudes - Démarche



2 UNCERTAINTY DEPENDING ON SPECTRAL DISTANCES

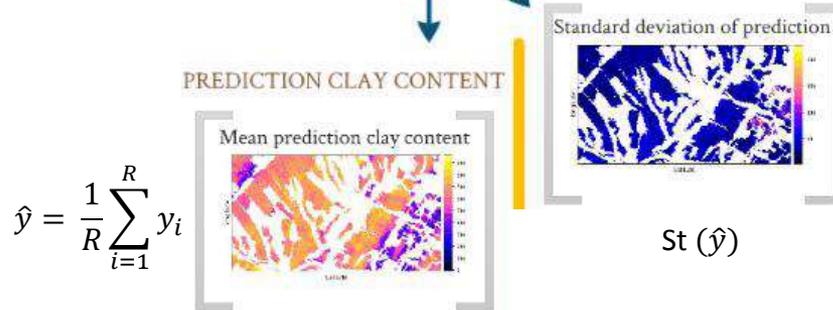


3 VARIANCE MODELLING

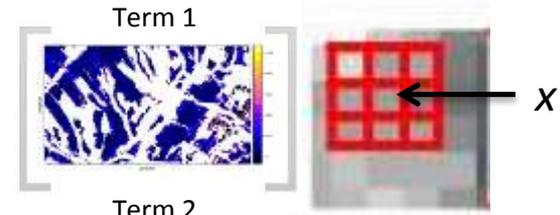
$$\text{Var}(\hat{y}) = \left(1 + \frac{1}{N}\right) \hat{\mathbf{b}}^T \Sigma_x \hat{\mathbf{b}} + \hat{\mathbf{z}}^T \Sigma_b \hat{\mathbf{z}} + \frac{\sigma_b^2}{N} + \left(1 + \frac{1}{N}\right) \text{tr}(\Sigma_x \Sigma_b)$$

Fernandez-Ahumada, et al. 2012

1 UNCERTAINTY DEPENDING ON CALIBRATION DATABASE



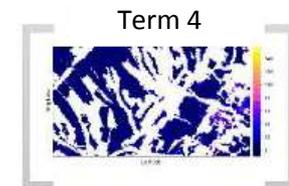
$$\left(1 + \frac{1}{N_{cal}}\right) \hat{\mathbf{b}}^T \Sigma_x \hat{\mathbf{b}}$$



$$\hat{\mathbf{z}}^T \Sigma_b \hat{\mathbf{z}}$$

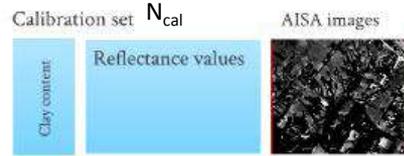


$$\left(1 + \frac{1}{N_{cal}}\right) \text{tr}(\Sigma_x \Sigma_b)$$

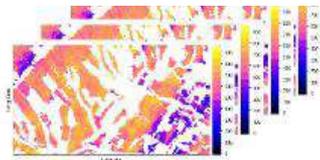


Calcul d'Incertitudes - Démarche

Bootstrap



$$y_i = f_i(x) \quad i = 1, \dots, R$$

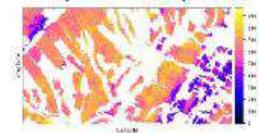


Tirage avec Remise
R=999

$$\hat{y} = \frac{1}{R} \sum_{i=1}^R y_i$$

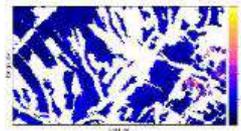
PREDICTION CLAY CONTENT

Mean prediction clay content



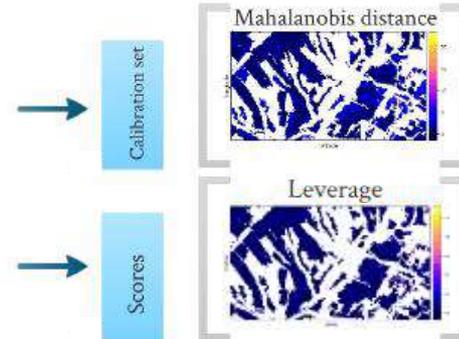
1 UNCERTAINTY DEPENDING ON CALIBRATION DATABASE

Standard deviation of prediction



St(\hat{y})

2 UNCERTAINTY DEPENDING ON SPECTRAL DISTANCES

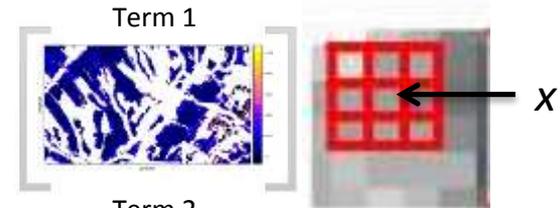


3 VARIANCE MODELLING

$$\text{Var}(\hat{y}) = \left(1 + \frac{1}{N}\right) \hat{\mathbf{b}}^T \Sigma_x \hat{\mathbf{b}} + \hat{\mathbf{z}}^T \Sigma_b \hat{\mathbf{z}} + \frac{\sigma_b^2}{N} + \left(1 + \frac{1}{N}\right) \text{tr}(\Sigma_x \Sigma_b)$$

Fernandez-Ahumada, et al. 2012

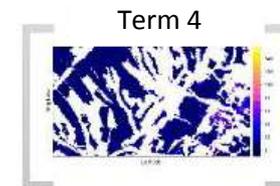
$$\left(1 + \frac{1}{N_{cal}}\right) \hat{\mathbf{b}}^T \Sigma_x \hat{\mathbf{b}}$$



$$\hat{\mathbf{z}}^T \Sigma_b \hat{\mathbf{z}}$$

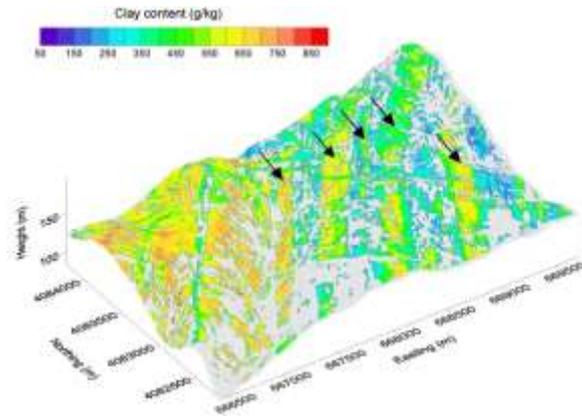


$$\left(1 + \frac{1}{N_{cal}}\right) \text{tr}(\Sigma_x \Sigma_b)$$



Données et sites d'étude

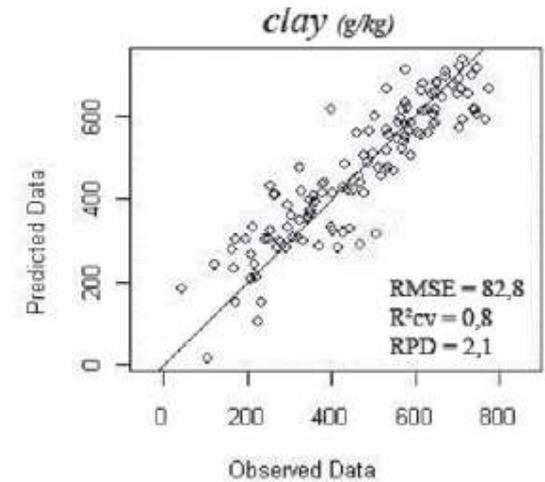
Spatialisation de la Teneur en Argile (*Gomez et al, 2012*)



96 échantillons dans
BD_calibration

32 échantillons dans
BD_Validation

300 km²



Spectral domain: 400-2500 nm

Number of Spectral Bands: 290

Spatial resolution: 5 m

Incertitudes sur BD_Validation

	St D of Predicted Values	Mean of Residues	MD	leverage	T1 (lié au spectres)	T2 (lié au modèle)	T4	Var Totale (T1 + T2 + T4)
St D of Predicted Values	1	0.32	<u>0.51</u>	<u>0.88</u>	0.31	<u>0.96</u>	0.49	<u>0.96</u>
Mean of Residues		1	0.09	0.36	0.2	0.32	-0.06	0.32
MD			1	0.4	0.39	<u>0.54</u>	0.5	<u>0.58</u>
leverage				1	0.05	<u>0.93</u>	0.19	<u>0.87</u>
T1 (lié au spectres)					1	0.28	<u>0.67</u>	0.46
T2 (lié au modèle)						1	0.46	<u>0.98</u>
T4							1	0.59

- Pas de relation entre « Incertitude » et « Erreur de prédiction » $R^2 < 0,36$

Incertitudes sur BD_Validation

	St D of Predicted Values	Mean of Residues	MD	leverage	T1 (lié au spectres)	T2 (lié au modèle)	T4	Var Totale (T1 + T2 + T4)
St D of Predicted Values	1	0.32	<u>0.51</u>	<u>0.88</u>	0.31	<u>0.96</u>	0.49	<u>0.96</u>
Mean of Residues		1	0.09	0.36	0.2	0.32	-0.06	0.32
MD			1	0.4	0.39	<u>0.54</u>	0.5	<u>0.58</u>
leverage				1	0.05	<u>0.93</u>	0.19	<u>0.87</u>
T1 (lié au spectres)					1	0.28	<u>0.67</u>	0.46
T2 (lié au modèle)						1	0.46	<u>0.98</u>
T4							1	0.59

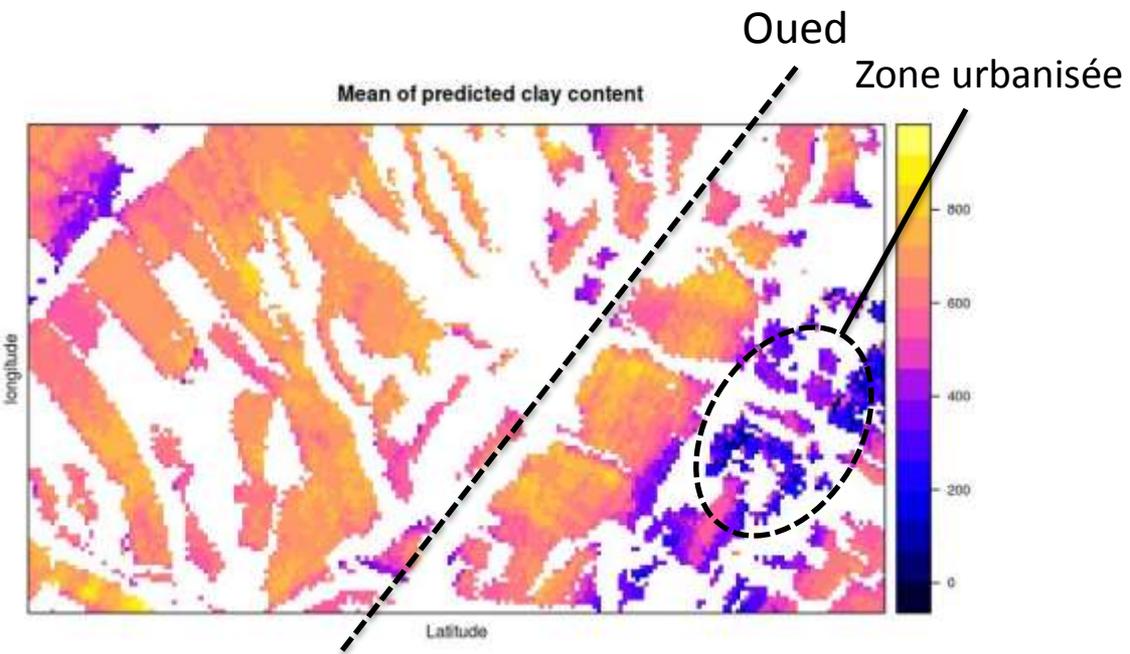
- Pas de relation entre « Incertitude » et « Erreur de prédiction » $R^2 < 0,36$
- Variance « vrai » (Sd des prédictions) \Leftrightarrow Distance dans l'espace du modèle (Leverage)
 \Leftrightarrow Variance liée au modèle (T2)

Incertitudes sur BD_Validation

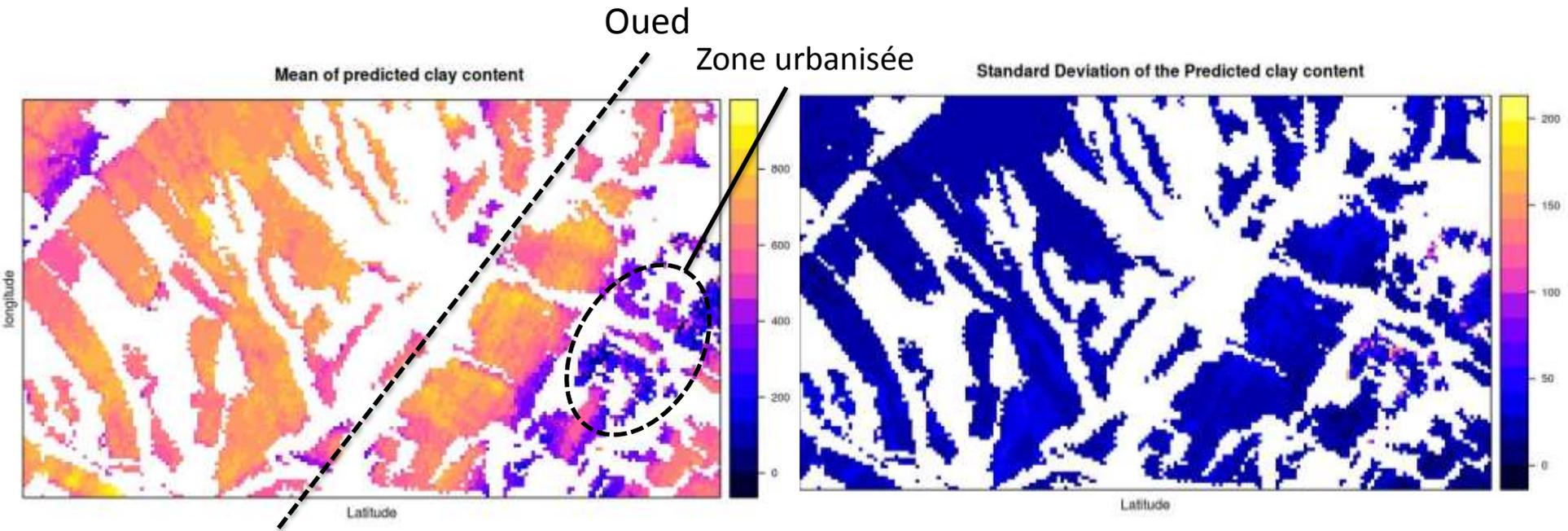
	St D of Predicted Values	Mean of Residues	MD	leverage	T1 (lié au spectres)	T2 (lié au modèle)	T4	Var Totale (T1 + T2 + T4)
St D of Predicted Values	1	0.32	<u>0.51</u>	<u>0.88</u>	0.31	<u>0.96</u>	0.49	<u>0.96</u>
Mean of Residues		1	0.09	0.36	0.2	0.32	-0.06	0.32
MD			1	0.4	0.39	<u>0.54</u>	0.5	<u>0.58</u>
leverage				1	0.05	<u>0.93</u>	0.19	<u>0.87</u>
T1 (lié au spectres)					1	0.28	<u>0.67</u>	0.46
T2 (lié au modèle)						1	0.46	<u>0.98</u>
T4							1	0.59

- Pas de relation entre « Incertitude » et « Erreur de prédiction » $R^2 < 0,36$
- Variance « vrai » (Sd des prédictions) \Leftrightarrow Distance dans l'espace du modèle (Leverage)
 \Leftrightarrow Variance liée au modèle (T2)
- Variance Totale selon *Fernandez-Ahumada et al* (2012) \Leftrightarrow Variance liée au modèle (T2)

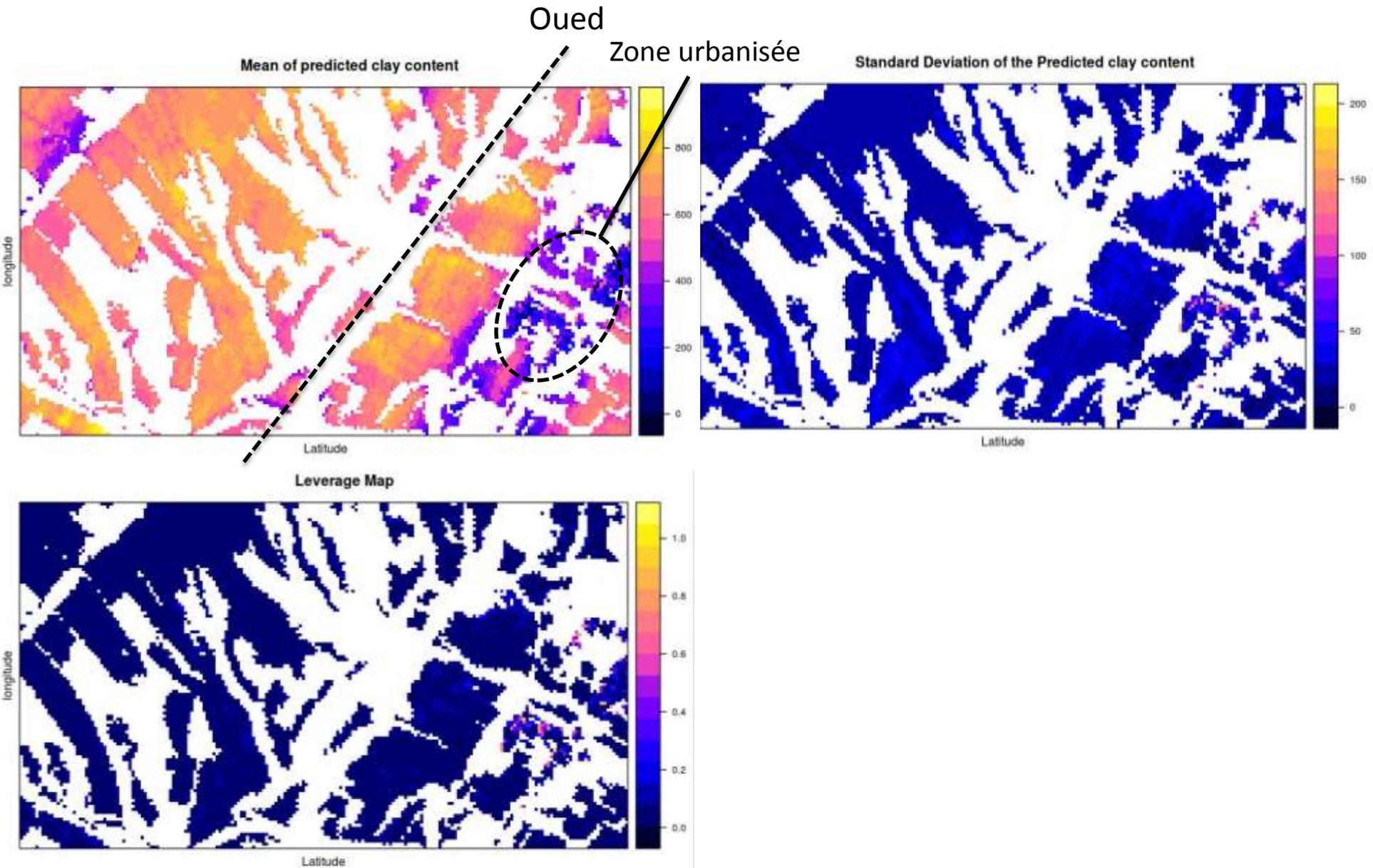
Spatialisation de Prédictions



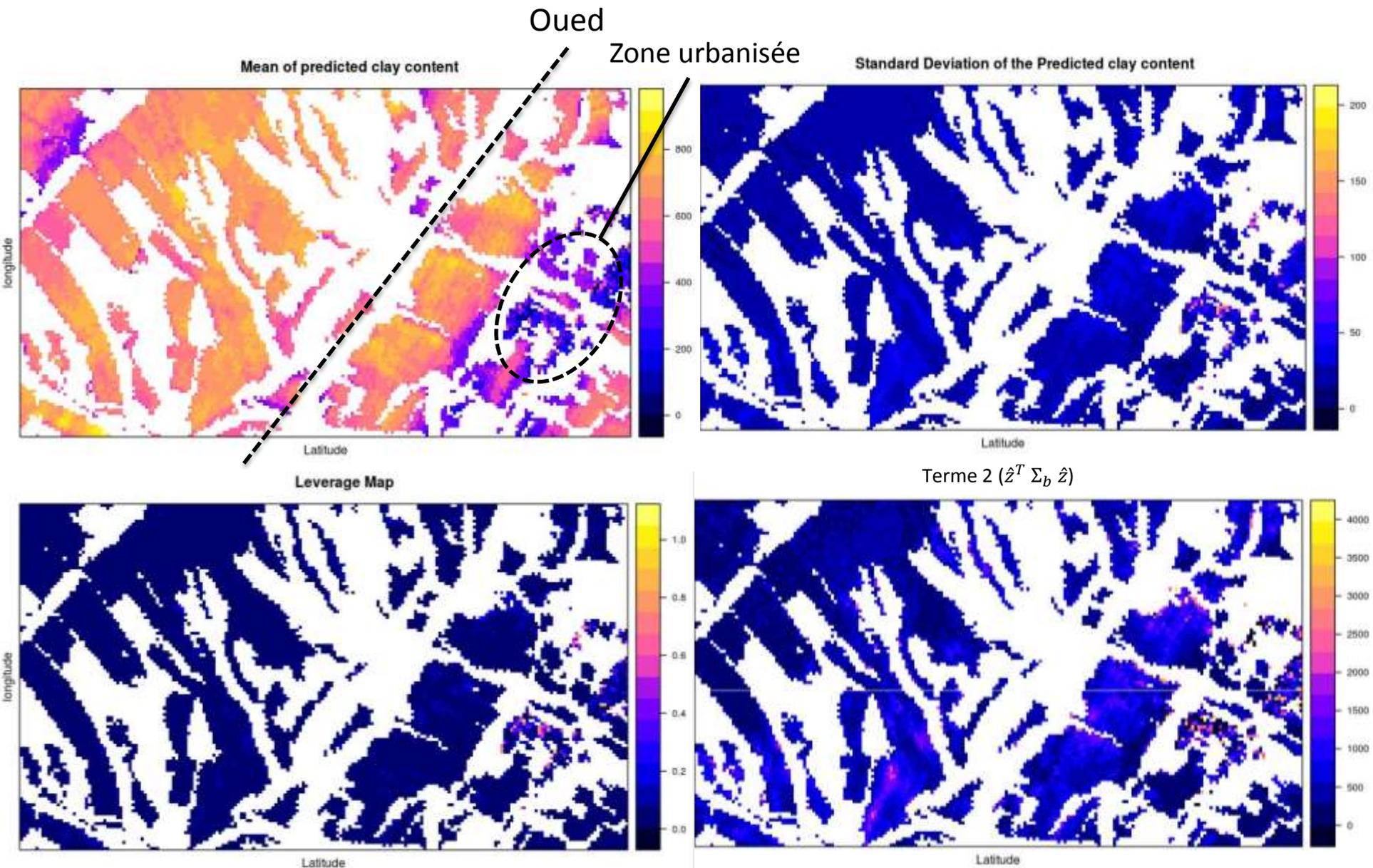
Spatialisation de Prédictions et d'Incertitudes



Spatialisation de Prédictions et d'Incertitudes



Spatialisation de Prédictions et d'Incertitudes



Conclusions

- Terme 2 $[\hat{z}^T \Sigma_b \hat{z}]$ plus d'important que Terme T1 $[(1 + \frac{1}{N_{cal}}) \hat{b}^T \Sigma_x \hat{b}]$
⇒ Modèles PLSR semble conduire à de plus fortes incertitudes que les spectres Vis-NIR aéroportés
- Pas de lien mis en évidence entre Incertitude et Erreur de prédiction
Mais vérifiée sur seulement 32 données, localisées au centre de parcelles de sol parfaitement nu.
- Permet d'identifier des zones à risques
Exemple des zones d'urbanisations => Amélioration des masques, avant spatialisation sur Sols « Nus »