Comparaison de procédures d'estimation dans le cadre des modèles non linéaires à paramètres aléatoires

Application à la modélisation de l'évolution temporelle de l'indice de surface foliaire de cultures observées par télédétection spatiale

Sébastien DÉJEAN

Université Paul Sabatier, Toulouse III

Unité de Biométrie et Intelligence Artificielle, INRA Toulouse

Enjeux du suivi de cultures

Prévision de production

$\Downarrow$

Gestion des marchés de

matières premières agro-alimentaires

Production = Surface $\times$ Rendement

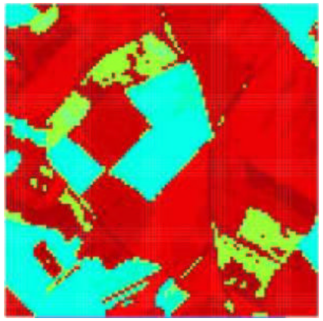
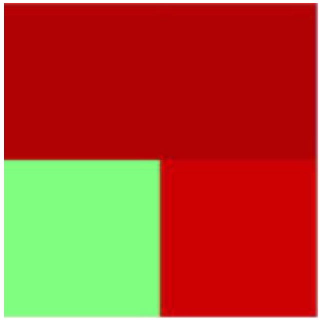
$\Downarrow$

Surface : détermination du POS

Rendement : suivi des cultures

Apports de la télédétection spatiale

Exemple : Satellite SPOT 4

Capteurs	HRVIR	VÉGÉTATION
Fréquence d'acquisition	26 jours	Quotidienne
Résolution spatiale	20 m	1 km
Fauchée	60 km	2250 km
Mesures physiques	Réflectances R et PIR	
Image		
Utilisation	POS	Dynamique temporelle

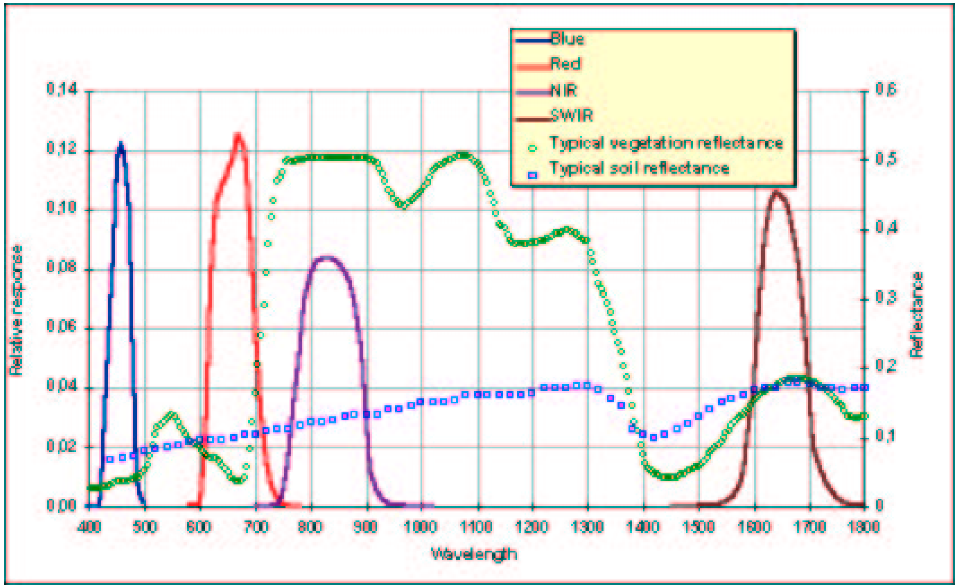
Désagrégation de pixels mixtes

$$\boxed{R_i = \sum_{k=1}^{N_c} \pi_i^k R_i^k}$$

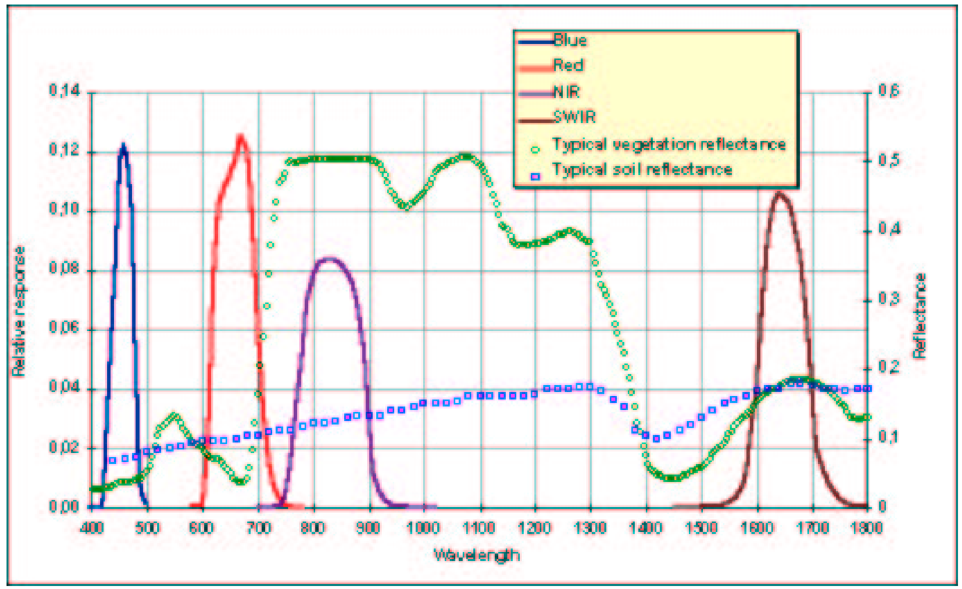
- R_i : réflectance mesurée sur le pixel i
- π_i^k : pourcentage du thème k sur le pixel i (POS)
- R_i^k : réflectance inconnue du thème k sur le pixel i

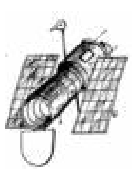
➢ Estimation des R_i^k dans un modèle linéaire à paramètres aléatoires (Faivre et Fischer, 1997).

Longueur d'onde et structure de la plante



Longueur d'onde et structure de la plante





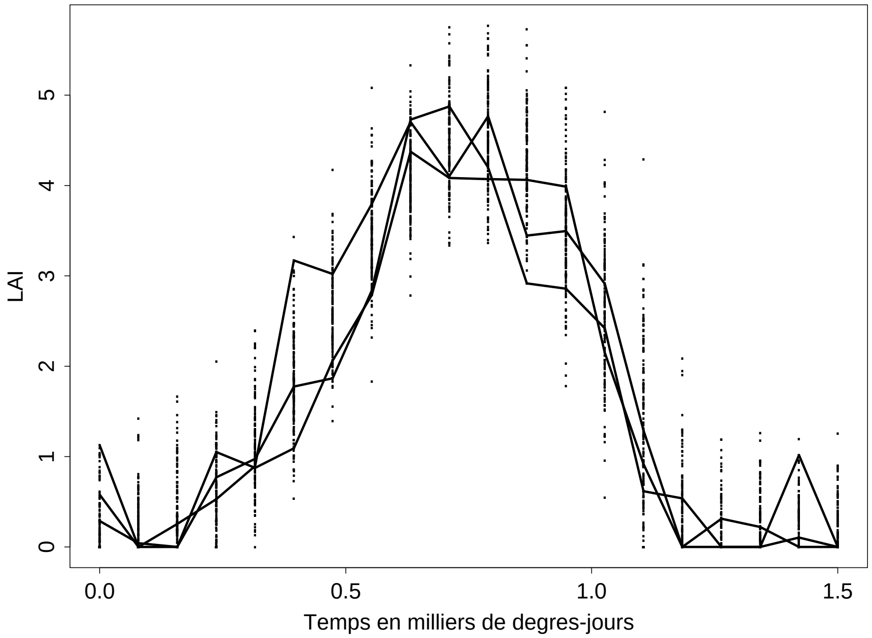
$$\boxed{\left.\begin{array}{l}\text{R}\\ \text{PIR}\end{array}\right\} \xrightarrow{(*)} \text{NDVI} \xrightarrow{(**)} \text{LAI}}$$

(*) $\text{NDVI} = \frac{\text{PIR}-\text{R}}{\text{PIR}+\text{R}}$: distinguer au mieux la végétation.

(**) $\text{LAI} = -\frac{1}{\alpha}\log\left(1 - \frac{\text{NDVI}-\text{NDVI}_s}{\text{NDVI}_m-\text{NDVI}_s}\right)$: information supplémentaire pour les cultures.

Les données

Données : valeurs de LAI de la culture d'intérêt pour chaque pixel à chaque date d'acquisition.



➥ Région de 30 à 40 km de côté : 1500 pixels de résolution kilométrique,

➥ Synthèse décadaire ou acquisition quotidienne d'images : de 20 à 200 dates sur 6 mois.

Question de recherche

➢ Dans le cadre du suivi de cultures :

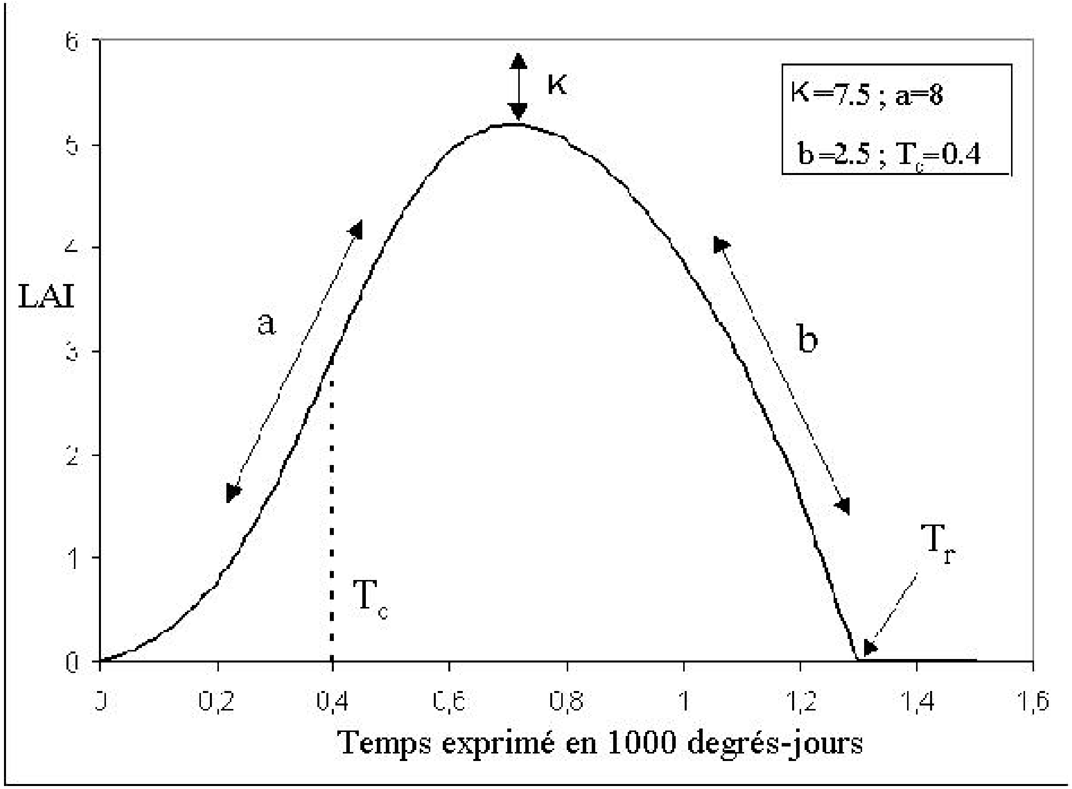
Comment améliorer le couplage de données de télédétection avec un modèle d'évolution temporelle d'un indice traduisant le développement de la culture ?

➢ Question statistique :

Comment améliorer l'estimation des paramètres d'un modèle non linéaire pour l'ajustement à un ensemble de courbes ?

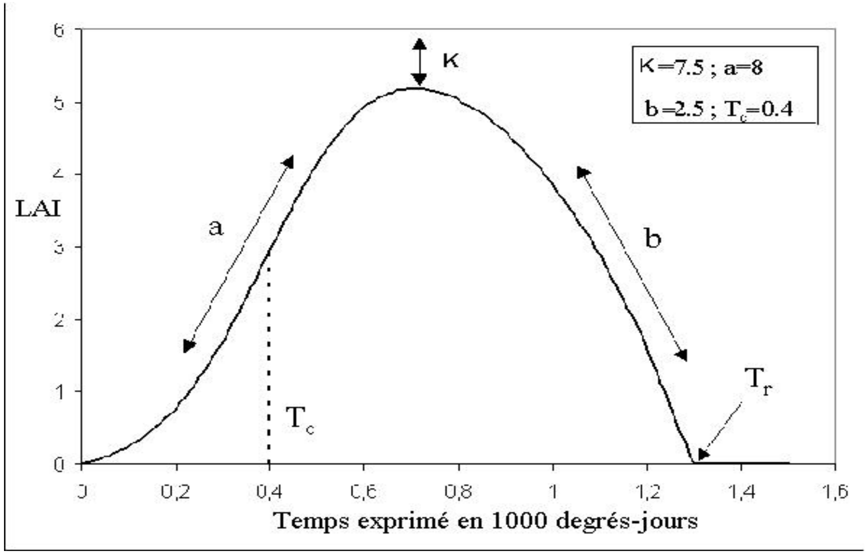
Modèle d'évolution du LAI *(Baret, 1986)*

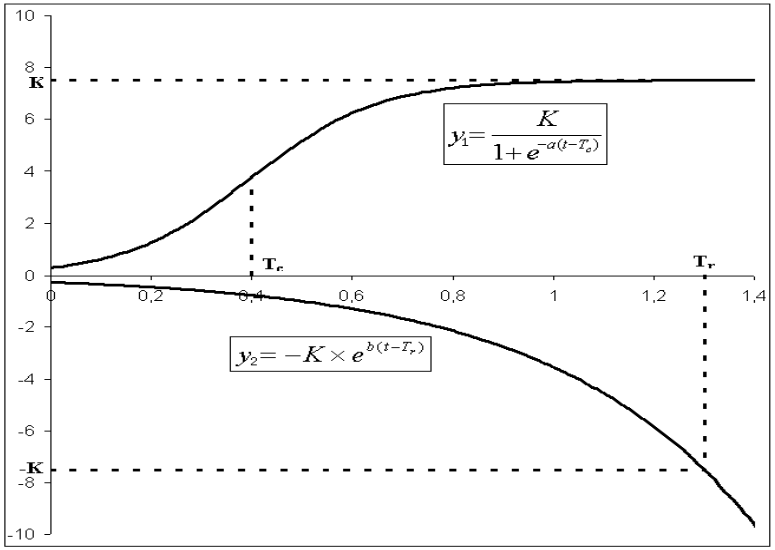
$$LAI(t,\beta) = f(t,\beta) = K\left(\frac{1}{1+e^{-a(t-T_c)}} - e^{b(t-T_r)}\right)$$



Modèle d'évolution du LAI *(Baret, 1986)*

$$LAI(t,\beta) = f(t,\beta) = K\left(\frac{1}{1+e^{-a(t-T_c)}} - e^{b(t-T_r)}\right)$$





- K : LAI maximal théoriquement atteint sans sénescense
- a, b : indicateurs des vitesses de croissance et de sénescence
- T_c, T_r : dates significatives du cycle de vie de la plante

Proposition

➪ *Modélisation et estimation dans le cadre des modèles non linéaires à paramètres aléatoires*

➥ traite l'ensemble des pixels simultanément,

➥ crée un lien entre les pixels par l'introduction d'une distribution de probabilités sur les paramètres de chaque pixel,

➥ permet l'estimation de paramètres locaux (à l'échelle du pixel) et globaux (à l'échelle de l'image).

Modèle non linéaire à paramètres aléatoires

- **Origine** *(Sheiner et al., 1972)* : pharmacologie, suivi de patients
- **But** : modéliser les variabilités intra- et inter-individus
- **Écriture** : Structure hiérarchique
  - ➢ Niveau 1 : variabilité intra-individu

$$\boxed{\left\{\begin{array}{rcl} Y_i|\beta_i & = & f(t,\beta_i)+\varepsilon_i \\ E(\varepsilon_i|\beta_i) & = & 0 \\ \text{cov}(\varepsilon_i|\beta_i) & = & R_i(\beta_i,\xi) \end{array}\right.}$$

  - ➢ Niveau 2 : variabilité inter-individus

$$\boxed{\beta_i \sim H}$$

Modèle paramétrique

$$\begin{cases} Y_{ij} & = & f(t_j, \beta_i) + \varepsilon_{ij} \\ \beta_i & = & \beta + B_i \\ B_i & \sim & \mathcal{N}_p(0, \Gamma) \qquad \textit{i.i.d.} \\ \varepsilon_{ij} & \sim & \mathcal{N}(0, \sigma_\varepsilon^2) \qquad \textit{i.i.d.} \end{cases}$$

- Paramètres globaux : β, Γ et σ_ε^2
- Paramètres locaux : B_i ou β_i

Vraisemblance :

$$\log[V(Y|\beta,\Gamma,\sigma_\varepsilon^2)] = -\frac{IJ+Ip}{2}\log(2\pi) - IJ\log\sigma - \frac{I}{2}\log(|\Gamma|) + \sum_{i=1}^{I}\left(\log\left[\int \exp\left(-\frac{1}{2\sigma_\varepsilon^2}\sum_{j=1}^{J}[Y_{ij} - f(t_j, \beta + B_i)]^2 - \frac{1}{2}B_i^T\Gamma^{-1}B_i\right) dB_i\right]\right)$$

Méthodes d'estimation

- ➢ Méthodes de type Standard Two Stage (STS), basées sur les estimations individuelles *(Steimer et al., 1984)*.
- ➢ Maximum de vraisemblance

V. directe	V. directe appr.	V. indirecte
EM	Approx. laplacienne	Pseudo-vrais. simulée
	Monte Carlo	Linéarisation ($\leftarrow$ itératif)
	Quad. gaussienne	

Méthode de type STS (Standard Two Stage)

① Estimation des paramètres individuels

$$\tilde{\beta}_i = \arg\min_{\theta} \sum_{j=1}^{J} [Y_{ij} - f(t_j, \theta)]^2$$

Estimation de la variance asymptotique de $\tilde{\beta}_i$: $\tilde{\Sigma}_i = \frac{\tilde{\sigma}_\varepsilon^2}{J} \left[\sum_{j=1}^{J} \frac{\partial \tilde{f}_j}{\partial \beta_i} \left(\frac{\partial \tilde{f}_j}{\partial \beta_i} \right)^T \right]^{-1}$

② Estimation des paramètres globaux

$$\begin{cases} \tilde{\beta}_{STS} & = & \frac{1}{I} \sum_{i=1}^{I} \tilde{\beta}_i \\ \tilde{\Gamma}_{STS} & = & \frac{1}{I-1} \sum_{i=1}^{I} (\tilde{\beta}_i - \tilde{\beta}_{STS})(\tilde{\beta}_i - \tilde{\beta}_{STS})^T \end{cases}$$

Correction : $\tilde{\Gamma}_{STS} - \frac{1}{I} \sum_{i=1}^{I} \tilde{\Sigma}_i$

Méthode directe approchée

Laplace	$\int e^{n\psi(\theta)}d\theta \simeq e^{n\psi(\hat{\theta})}\left(\frac{2\pi}{n}\right)^{\frac{1}{2}}\lvert\psi"(\hat{\theta})\rvert^{-\frac{1}{2}}$
Monte-Carlo	$\int \psi(x)dx = E\left[\frac{\psi(X)}{\varphi(X)}\right] \simeq \frac{1}{n}\sum_{k=1}^{n}\frac{\psi(x_k)}{\phi(x_k)}$
Quadrature	$\int_{-\infty}^{+\infty} e^{-x^2}\psi(x)dx \simeq \sum_{k=1}^{n}\omega_k\psi(x_k)$

Application au modèle non linéaire à paramètres aléatoires *(Pinheiro et Bates, 1995)*

Méthode indirecte

➢ **Linéarisation** *(Lindstrom et Bates, 1990)*

Principe : développement de Taylor à l'ordre 1 de la fonction f

➠ estimation dans un modèle linéaire

$$\boxed{Y_i|B_i \dot{\sim} \mathcal{N}\left(f(t_j,\hat{\beta},\hat{B}_i)+\hat{Z}(B_i-\hat{B}_i)+\hat{X}(\beta-\hat{\beta}),\sigma_\varepsilon^2\right)}$$

$\Rightarrow W_i \dot{\sim} \mathcal{N}\left(\hat{X}\beta,\sigma_\varepsilon^2+\hat{Z}^T\Gamma\hat{Z}\right)$ avec $W_{ij}=Y_{ij}-f(t_j,\hat{\beta},\hat{B}_i)+\hat{Z}\hat{B}_i+\hat{X}\hat{\beta}$

Mise en œuvre : algorithme itératif 2 étapes : estimations globales / estimations locales

Convergence des estimateurs

Méthode	Convergence	Réf.
STS	❷	Demidenko, 1998
EM	❶	Walker, 1996
Laplace	❷ ?	Vonesh, 1996
Monte Carlo	?	
Quad. gaus.	?	
PVS	❶	Nuñez et Concordet, 2002
Linéarisation	❷	Demidenko, 1998

❶ : $I \to \infty$ ❷ $I, J \to \infty$

I : nombre d'individus J : nombre d'observations par individu

Plan d'expérience pour comparaison

A. Comparaison linéarisation / approximation numérique (laplacienne, Monte Carlo, quadrature gaussienne).

B. Comparaison linéarisation / STS (classiquement utilisée en télédétection) ?

C. Influence de structures spatiales (sur le bruit ou les paramètres) sur l'estimation par linéarisation ?

D. Sur données réelles, étude des effets croisés de la procédure d'estimation et de la mise en œuvre de la désagrégation.

A. Plan d'expérience

➢ **Motivations** : Meilleure méthode (critère d'EQM) pour l'estimation des paramètres globaux β,Γ ?

➢ **Cas de figures** :

a. zone d'étude hétérogène : paramètres aléatoires (Cfg. 1-2),

b. zone d'étude homogène : paramètres constants (Cfg. 3-4).

Simulation : $\beta = (7, 3, 0.5)$, $\Gamma =$ diag$(\sigma_a, \sigma_b, \sigma_{T_c})$

Config.	σ_a	σ_b	σ_{T_c}	σ_ε
1	0.1	0.1	0.01	0.05
2	0.1	0.1	0.01	0.5
3	0	0	0	0.05
4	0	0	0	0.5

100 répétitions par configuration

Nombre d'évaluations de la fonction f jusqu'à convergence :

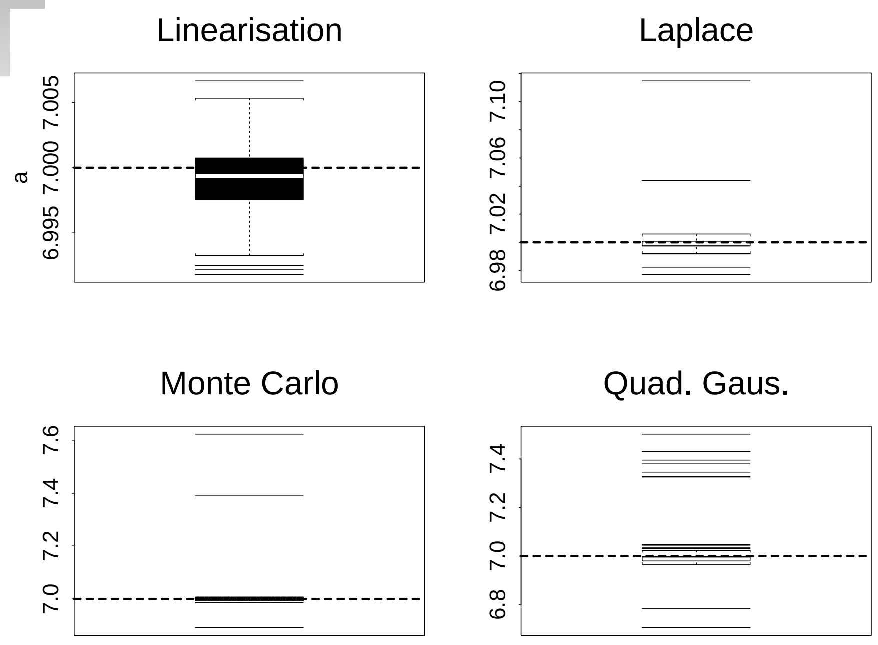
Config.	Laplace	MC	QG
1	80 000	8 500 000	41 500 000
2	100 000	8 300 000	52 500 000
3	92 000	8 700 000	55 200 000
4	96 000	8 200 000	30 300 000

Temps réel : Laplace 20 s, MC 600 s, QG 3600 s

Linéarisation : 1500 s (implémentation différente)

A. Résultats

Exemple : config. 2



➥ linéarisation : meilleure méthode sur un critère d'erreur quadratique moyenne,

➥ aucune méthode n'améliore significativement les résultats de la linéarisation

➥ faible taille d'échantillon des méthodes Monte Carlo et quadrature imposée par la contrainte opérationnelle

➪ Validation pour notre problème de l'estimation par linéarisation

B. Plan d'expérience

➢ Motivations :

a. Influence du nombre de paramètres sur la précision des estimateurs ? (*Plan 1*)

b. Influence du nombre de dates sur la précision des estimateurs ? (Plan 2)

c. Précision des estimations en fonction du nombre de dates (bruit élevé) ? (Plan 3)

Configuration	Nombre de dates	Paramètres à estimer	σ_ε^2
1	*20*	(K, a, b, T_c)	0.05^2
2	*20*	(K, a, b, T_c)	0.5^2
3	*20*	(a, b, T_c)	0.05^2
4	*20*	(a, b, T_c)	0.5^2
5	10	(a, b, T_c)	0.05^2
6	10	(a, b, T_c)	0.5^2
7	40	(a, b, T_c)	0.5^2

B. Plan d'expérience

➢ **Motivations :**

a. Influence du nombre de paramètres sur la précision des estimateurs ? (*Plan 1*)

b. Influence du nombre de dates sur la précision des estimateurs ? (Plan 2)

c. Précision des estimations en fonction du nombre de dates (bruit élevé) ? (Plan 3)

Configuration	Nombre de dates	Paramètres à estimer	σ_ε^2
1	*20*	(K, a, b, T_c)	0.05^2
2	*20*	(K, a, b, T_c)	0.5^2
3	*20*	(a, b, T_c)	0.05^2
4	*20*	(a, b, T_c)	0.5^2
5	10	(a, b, T_c)	0.05^2
6	10	(a, b, T_c)	0.5^2
7	40	(a, b, T_c)	0.5^2

B. Plan d'expérience

➢ **Motivations :**

a. Influence du nombre de paramètres sur la précision des estimateurs ? (*Plan 1*)

b. Influence du nombre de dates sur la précision des estimateurs ? (Plan 2)

c. Précision des estimations en fonction du nombre de dates (bruit élevé) ? (Plan 3)

Configuration	Nombre de dates	Paramètres à estimer	σ_ε^2
1	*20*	(K, a, b, T_c)	0.05^2
2	*20*	(K, a, b, T_c)	0.5^2
3	*20*	(a, b, T_c)	0.05^2
4	*20*	(a, b, T_c)	0.5^2
5	10	(a, b, T_c)	0.05^2
6	10	(a, b, T_c)	0.5^2
7	40	(a, b, T_c)	0.5^2

B. Plan d'expérience

➢ **Motivations** :

a. Influence du nombre de paramètres sur la précision des estimateurs ? (*Plan 1*)

b. Influence du nombre de dates sur la précision des estimateurs ? (Plan 2)

c. Précision des estimations en fonction du nombre de dates (bruit élevé) ? (Plan 3)

Configuration	Nombre de dates	Paramètres à estimer	σ_ε^2
1	*20*	(K, a, b, T_c)	0.05^2
2	*20*	(K, a, b, T_c)	0.5^2
3	*20*	(a, b, T_c)	0.05^2
4	*20*	(a, b, T_c)	0.5^2
5	10	(a, b, T_c)	0.05^2
6	10	(a, b, T_c)	0.5^2
7	40	(a, b, T_c)	0.5^2

B. Résultats

- ➢ Non convergence des régressions par pixel : 1% (resp. 10%) dans les configurations faiblement (resp. fortement) bruitées.
- ➢ Temps de calcul : linéarisation 10 fois plus long que STS.

B. Résultats

EQM(β_i)	LIN				STS			
Config.	K	a	b	T_c	K	a	b	T_c
1	*0.02*	*0.03*	*0.04*	*0.0004*	*0.1*	*0.06*	*0.007*	*0.00003*
2	*0.8*	*0.4*	*0.07*	*0.0004*	*80000*	*400*	*50*	*0.02*
3	.	*0.04*	*0.04*	*0.0004*	.	*0.002*	*0.001*	*0.00001*
4	.	*0.02*	*0.03*	*0.0002*	.	*0.4*	*0.2*	*0.001*
5	.	0.03	0.04	0.0003		0.005	0.004	0.00003
6	.	0.02	0.03	0.0002	.	14	1	0.01
7	.	0.02	0.03	0.0002	.	0.1	0.06	0.0004

Critère d'EQM sur les estimations individuelles :

$$\text{EQM}(\hat{\beta}_i) = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{1}{I}\sum_{i=1}^{I}\left(\hat{\beta}_i^{(n)} - \beta_i^{(n)}\right)^2\right)$$

B. Résultats

EQM(β_i)	LIN				STS			
Config.	K	a	b	T_c	K	a	b	T_c
1	0.02	0.03	0.04	0.0004	0.1	0.06	0.007	0.00003
2	0.8	0.4	0.07	0.0004	80000	400	50	0.02
3	.	0.04	0.04	0.0004	.	0.002	0.001	0.00001
4	.	0.02	0.03	0.0002	.	0.4	0.2	0.001
5	.	0.03	0.04	0.0003		0.005	0.004	0.00003
6	.	0.02	0.03	0.0002	.	14	1	0.01
7	.	0.02	0.03	0.0002	.	0.1	0.06	0.0004

Critère d'EQM sur les estimations individuelles :

$$\text{EQM}(\hat{\beta}_i) = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{1}{I}\sum_{i=1}^{I}\left(\hat{\beta}_i^{(n)} - \beta_i^{(n)}\right)^2\right)$$

Configurations faiblement bruitées

EQM(β_i)	LIN				STS			
Config.	K	a	b	T_c	K	a	b	T_c
1	0.02	0.03	0.04	0.0004	0.1	0.06	0.007	0.00003
2	0.8	0.4	0.07	0.0004	80000	400	50	0.02
3	.	0.04	0.04	0.0004	.	0.002	0.001	0.00001
4	.	0.02	0.03	0.0002	.	0.4	0.2	0.001
5	.	0.03	0.04	0.0003		0.005	0.004	0.00003
6	.	0.02	0.03	0.0002	.	14	1	0.01
7	.	0.02	0.03	0.0002	.	0.1	0.06	0.0004

Critère d'EQM sur les estimations individuelles :

$$\text{EQM}(\hat{\beta}_i) = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{1}{I}\sum_{i=1}^{I}\left(\hat{\beta}_i^{(n)} - \beta_i^{(n)}\right)^2\right)$$

B. Résultats

Configurations fortement bruitées

EQM(β_i)	LIN				STS			
Config.	K	a	b	T_c	K	a	b	T_c
1	0.02	0.03	0.04	0.0004	0.1	0.06	0.007	0.00003
2	0.8	0.4	0.07	0.0004	80000	400	50	0.02
3	.	0.04	0.04	0.0004	.	0.002	0.001	0.00001
4	.	0.02	0.03	0.0002	.	0.4	0.2	0.001
5	.	0.03	0.04	0.0003		0.005	0.004	0.00003
6	.	0.02	0.03	0.0002	.	14	1	0.01
7	.	0.02	0.03	0.0002	.	0.1	0.06	0.0004

Critère d'EQM sur les estimations individuelles :

$$\text{EQM}(\hat{\beta}_i) = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{1}{I}\sum_{i=1}^{I}\left(\hat{\beta}_i^{(n)} - \beta_i^{(n)}\right)^2\right)$$

B. Conclusion

➢ Cas favorables au traitement par pixel (STS) :

- ➥ nombre de paramètres à estimer faible
- ➥ nombre d’observations par individu élevé
- ➥ données faiblement bruitées

➪ L’estimation par linéarisation est meilleure quand ces conditions ne sont plus respectées.

C. Plan d'expérience

➢ **Motivations** : Étude du comportement de l'estimation par linéarisation face à des données structurées spatialement

① Structure spatiale sur le bruit

➥ traduit une dépendance des corrections apportées aux images vis-à-vis du POS

② Structure spatiale sur les paramètres

➥ traduit des comportements proches pour le développement de la culture sur des pixels voisins

C. Résultats

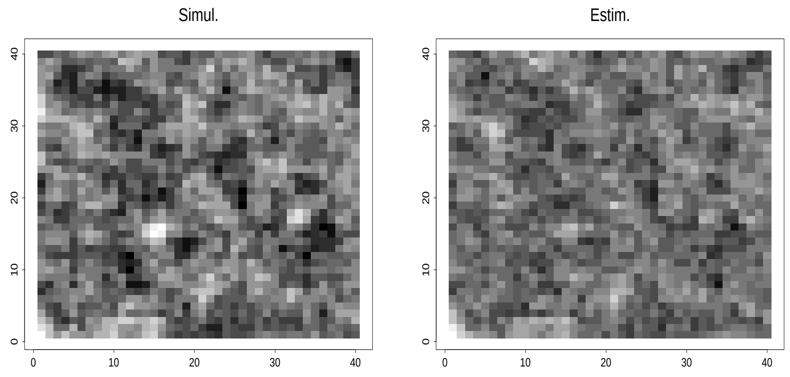
① Structure spatiale sur l'erreur

➥ L'estimation par linéarisation obtient de bons résultats (critère EQM), comparables au cas où l'erreur est non structurée.

② Structure spatiale sur les paramètres

➥ Bonne estimation des paramètres globaux,

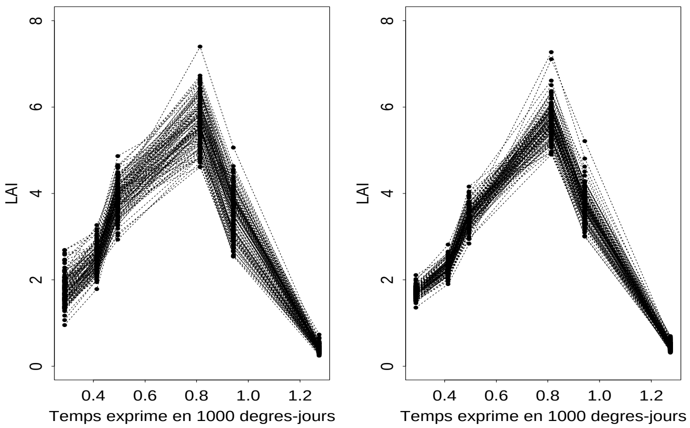
➥ l'estimation des paramètres par linéarisation retrouve les zones homogènes simulées (bruit faible).



D. Comparaison sur données réelles

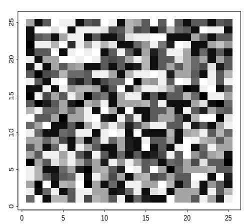
➢ **Motivations** : étude des effets croisés de la procédure d'estimation et de la mise en œuvre de la désagrégation.

- 6 scènes SPOT/HRV (36 km × 36 km, résolution 20 m) de la Beauce Chartraine prises entre le 19 mars et le 16 juillet 1996
- Culture d'intérêt : céréales d'hiver (53% du POS)
- Traitement de 1205 pixels kilométriques
- Création de 2 jeux de données

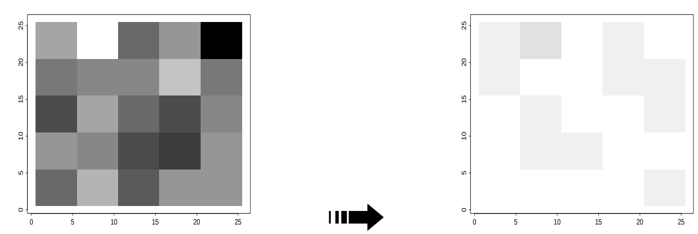


D. Création des 2 jeux de données

Haute résolution

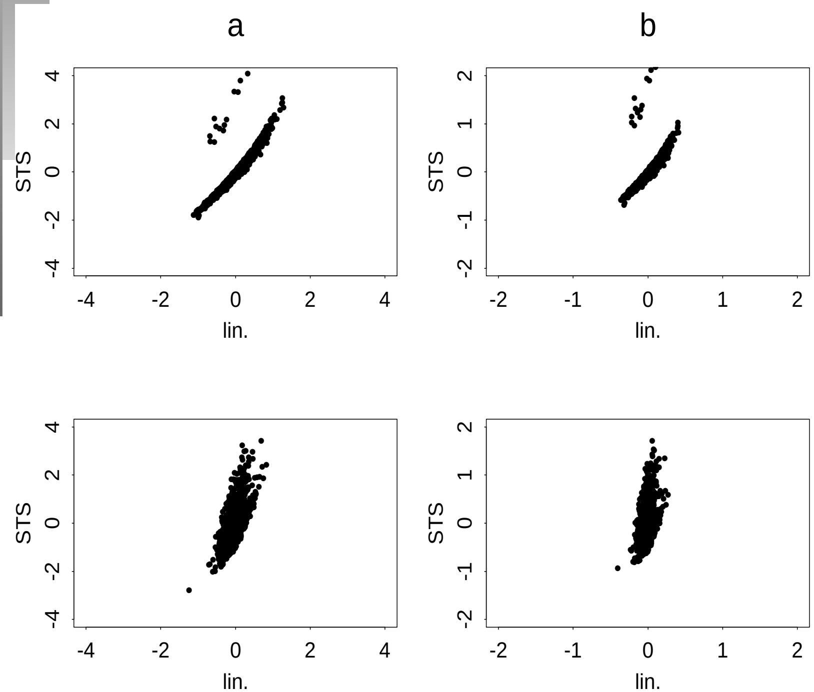


① Pixels purs moyennés



② Pixels mixtes désagrégés

D. Résultats



- Paramètres K et T_c fixés
- STS et LIN varient globalement dans le même sens
- Comportement rétrécisseur de l'estimation par image (/ par pixel)
- Étendue réduite des estimations sur la désagrégation

Écart-type résiduel estimé : ˜0.5

➥ erreur probable de modélisation ($\neq$ simulations)

➥ superposition et correction des images

Conclusion générale

➢ **Contexte :** télédétection spatiale - statistique - agronomie

➢ **Proposition :** modèle non linéaire à paramètres aléatoires

➢ **Originalité des travaux :**

➥ nouvelle méthodologie pour le traitement de données de télédétection alternative à une approche par pixel

➥ étude des modèles non linéaires à paramètres aléatoires dans un nouveau cadre d'étude

➥ cartographie de paramètres : dynamique temporelle

➢ **Résultats :**

➥ synthèse bibliographique

➥ efficacité de l'estimation par linéarisation

Perspectives

➢ Prise en compte d'une structure spatiale dans les modèles non linéaires à paramètres aléatoires,

➢ Coupler les phases de désagrégation et d'estimation des paramètres du modèle de LAI : $Y_i(t) = \sum_{k=1}^{p} \pi_i^k f^k(t, \beta_i^k) + \eta_i^k$ avec $\beta_i^k \sim \mathcal{N}(\beta^k, \sigma^k)$.

➢ Mise en œuvre sur un cas réel : couplage avec un modèle de fonctionnement.