



Analyse statistique de l'information spatiale

David Causeur

Laboratoire de Mathématiques Appliquées

Agrocampus Ouest

IRMAR CNRS UMR 6625

<http://www.agrocampus-ouest.fr/math/causeur/>

Plan du cours

1 Introduction

Problématiques

Modélisation spatiale

2 Description statistique de la variabilité spatiale

Données géostatistiques

Lissage de données spatiales

3 Modélisation géostatistique

Modèles de dépendance spatiale

Analyse variographique

Prédiction

4 Modélisation multivariée

5 Perspectives



Domaines d'application

Domaine d'application historique

Domaines d'application privilégiés

Domaines d'application plus récents

Domaines d'application

Domaine d'application historique

- **Prospection minière**

évaluation des ressources

Domaines d'application privilégiés

Domaines d'application plus récents

Domaines d'application

Domaine d'application historique

- Prospection minière

évaluation des ressources

Domaines d'application privilégiés

- Environnement

cartographie d'un polluant

Domaines d'application plus récents

Domaines d'application

Domaine d'application historique

- Prospection minière

évaluation des ressources

Domaines d'application privilégiés

- Environnement
- Épidémiologie

cartographie d'un polluant

diffusion d'une épidémie

Domaines d'application plus récents



Domaines d'application

Domaine d'application historique

- Prospection minière

évaluation des ressources

Domaines d'application privilégiés

- Environnement
- Épidémiologie

cartographie d'un polluant

diffusion d'une épidémie

Domaines d'application plus récents

- Économie

répartition géographique de populations



Domaines d'application

Domaine d'application historique

- Prospection minière

évaluation des ressources

Domaines d'application privilégiés

- Environnement
- Épidémiologie

cartographie d'un polluant

diffusion d'une épidémie

Domaines d'application plus récents

- Économie
- Agriculture

répartition géographique de populations

agriculture de précision



Domaines d'application

Domaine d'application historique

- Prospection minière

évaluation des ressources

Domaines d'application privilégiés

- Environnement
- Épidémiologie

cartographie d'un polluant

diffusion d'une épidémie

Domaines d'application plus récents

- Économie
- Agriculture
- Marketing

répartition géographique de populations

agriculture de précision

déplacement d'un consommateur

Domaines d'application

Domaine d'application historique

- Prospection minière évaluation des ressources

Domaines d'application privilégiés

- Environnement cartographie d'un polluant
- Épidémiologie diffusion d'une épidémie

Domaines d'application plus récents

- Économie répartition géographique de populations
- Agriculture agriculture de précision
- Marketing déplacement d'un consommateur
- Écologie dynamique des populations

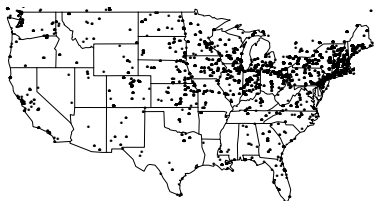


Typologie des problématiques

Variabilité spatiale

- Localisation
 - Modélisation des coordonnées
 - Problématiques : analyse de répartition spatiale, estimation de densité, ...

Lieux d'observations d'une espèce d'oiseaux



Typologie des problématiques

Variabilité spatiale

- Localisation
 - Modélisation des coordonnées
 - Problématiques : analyse de répartition spatiale, estimation de densité, ...
- Mesures localisées
 - Modélisation des variations dans l'espace
 - Problématiques : cartographie, prédiction, ...



Principes généraux de modélisation

Champ aléatoire Z

$$Z = \{Z(S), S \in \mathcal{S}\}$$

Données spatiales

<i>Site</i>	LOCALISATION DU SITE			CARACTÉRISTIQUES			
S	X	Y	...	Z_1	Z_2	...	Z_k
s_1	x_1	y_1	...	$Z_1(s_1)$	$Z_2(s_1)$...	$Z_k(s_1)$
s_2	x_2	y_2	...	$Z_1(s_2)$?	...	$Z_k(s_2)$
\vdots	\vdots	\vdots		\vdots	\vdots	...	\vdots
s_n	x_1	y_1	...	$Z_1(s_n)$	$Z_2(s_n)$...	?



Méthodologie

Typologie des méthodes :

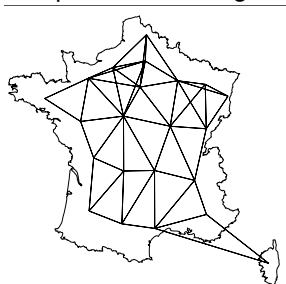
- Processus ponctuels
 - \mathcal{S} ensemble aléatoire de coordonnées
 - $Z(s)$ non-spécifié
 - Mots-clés : estimation de densité, répartition spatiale, ...



Méthodologie

Typologie des méthodes :

- Processus ponctuels
 - \mathcal{S} ensemble aléatoire de coordonnées
 - $Z(s)$ non-spécifié
 - Mots-clés : estimation de densité, répartition spatiale, ...
- Statistique latente
 - \mathcal{S} ensemble fixe de coordonnées sur une grille
 - $Z(s)$ champ aléatoire au site $s \in \mathcal{S}$
 - Mots-clés : modèles spatiaux auto-régressifs, ...





Méthodologie

Typologie des méthodes :

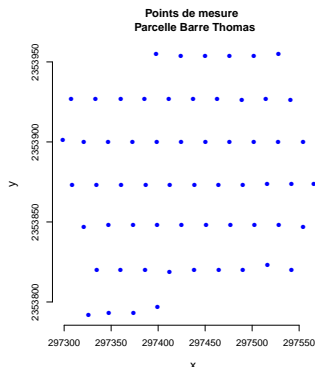
- Processus ponctuels
 - \mathcal{S} ensemble aléatoire de coordonnées
 - $Z(s)$ non-spécifié
 - Mots-clés : estimation de densité, répartition spatiale, ...
- Statistique latente
 - \mathcal{S} ensemble fixe de coordonnées sur une grille
 - $Z(s)$ champ aléatoire au site $s \in \mathcal{S}$
 - Mots-clés : modèles spatiaux auto-régressifs, ...
- Géostatistique
 - \mathcal{S} ensemble fixe de coordonnées géographiques
 - $Z(s)$ champ aléatoire au site $s \in \mathcal{S}$
 - Mots-clés : tendance spatiale, analyse variographique, krigeage, ...



Variabilité intra-parcellaire

Contexte : parcelle Barre Thomas

- Surface : 150 m \times 250 m
- Analyses pédologiques : pH, Carbone organique, Azote total, CEC, Argile

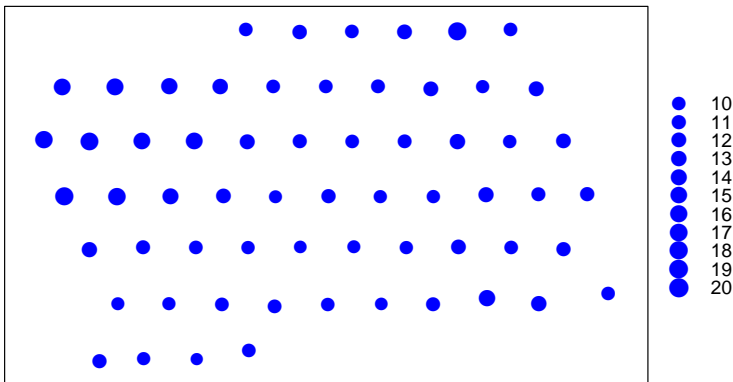


Questions

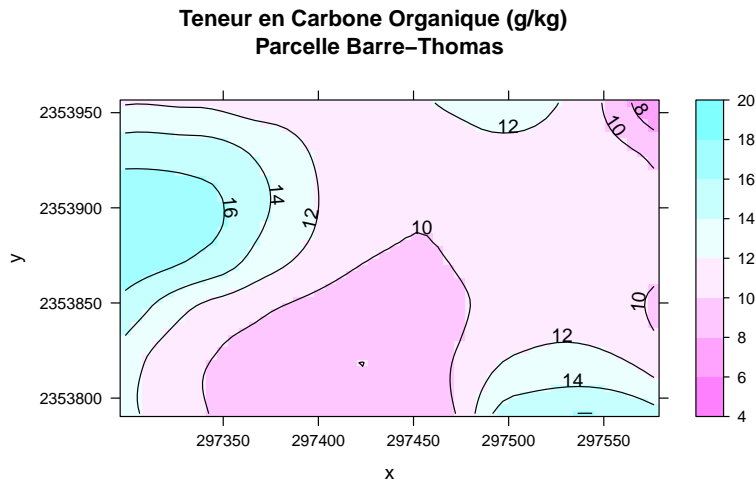
- Cartographie des indicateurs pédologiques
- Etude des liens entre indicateurs pédologiques
- Optimisation des ressources expérimentales

Visualisation de la variabilité intra-parcellaire

Teneur en Carbone Organique (g/kg)
Parcelle barre Thomas



Visualisation de la variabilité intra-parcellaire





Dépendance spatiale

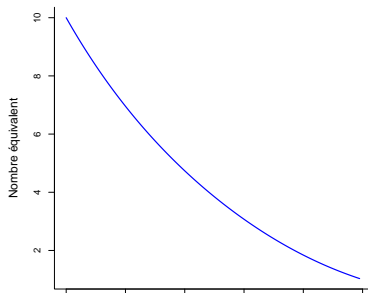
- Modèle de dépendance spatiale 1D

$$\mathbb{E}[Z(s)] = \mu, \quad \text{Cov}[Z(s_1), Z(s_2)] = \sigma^2 \rho^{|s_2 - s_1|}, \quad 0 < \rho < 1$$

- Échantillonnage : $\{Z(1), Z(2), \dots, Z(n)\}$
- Estimateur de μ : $\hat{\mu} = \bar{Z}$

$$\text{Var}(\hat{\mu}) = \frac{\sigma^2}{n} \left[1 + 2\alpha \frac{n-1}{n} - 2\alpha^2 \frac{1 - \rho^{n-1}}{n} \right] = \frac{\sigma^2}{n'}$$

où $\alpha = \rho/(1 - \rho)$.





Plan du cours

- 1 Introduction
 - Problématiques
 - Modélisation spatiale
- 2 Description statistique de la variabilité spatiale
 - Données géostatistiques
 - Lissage de données spatiales
- 3 Modélisation géostatistique
 - Modèles de dépendance spatiale
 - Analyse variographique
 - Prédiction
- 4 Modélisation multivariée
- 5 Perspectives



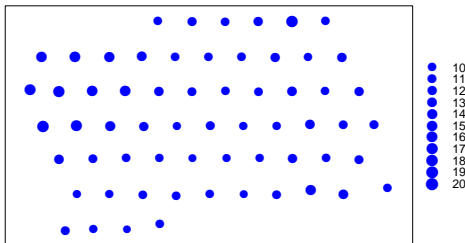
Données géostatistiques

- \mathcal{S} ensemble des sites possibles de mesure,
- $Z(s)$ la mesure au site s ,
- $\{s_1, s_2, \dots, s_n\}$, $n \geq 1$, ensemble des sites de mesure,



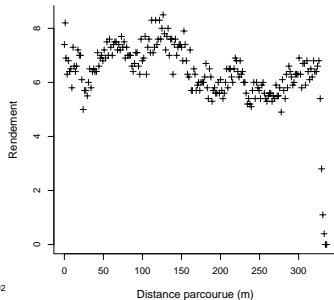
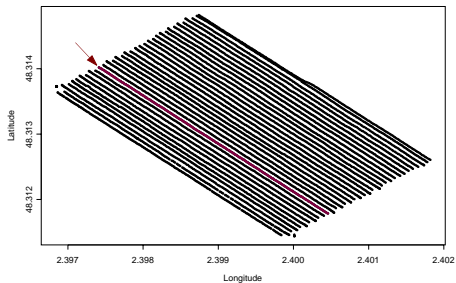
$$Z = \begin{pmatrix} Z(s_1) \\ Z(s_2) \\ \vdots \\ Z(s_n) \end{pmatrix}$$

Teneur en Carbone Organique (g/kg)
Parcelle barre Thomas



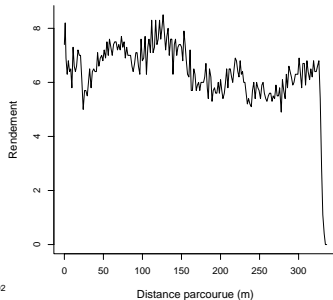
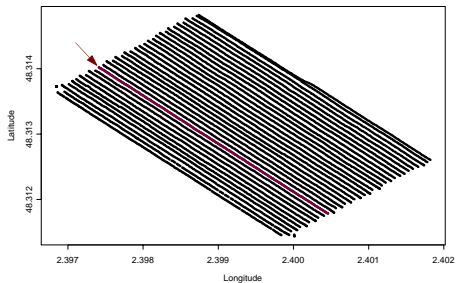


Modélisation géostatistique



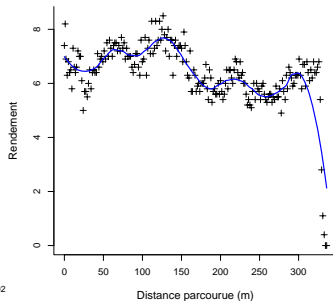
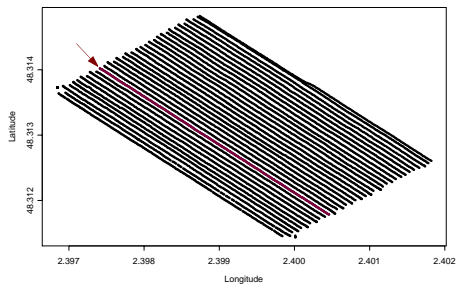


Modélisation géostatistique



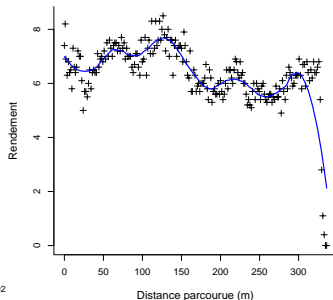
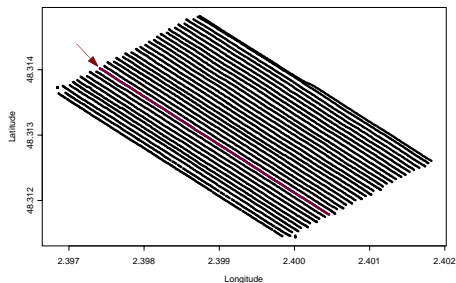


Modélisation géostatistique





Modélisation géostatistique



Modèle géostatistique

$$Z(s) = \mu(s) + \varepsilon(s)$$

$\mu(s)$: variations à grande échelle

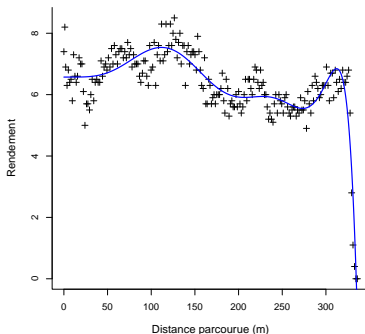
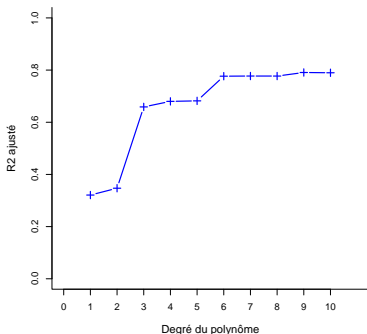
$\varepsilon(s)$: variations à faible échelle

Estimation de la tendance spatiale

Modèles paramétriques

$$\mu(x, y) = \beta_0 + \beta_1 x + \beta_2 y,$$

$$\mu(x, y) = \beta_0 + \beta_1 x + \beta_2 y + \beta_3 xy + \beta_4 x^2 + \dots,$$





Estimation de la tendance spatiale

Modèles paramétriques

$$\mu(x, y) = \beta_0 + \beta_1 x + \beta_2 y,$$

$$\mu(x, y) = \beta_0 + \beta_1 x + \beta_2 y + \beta_3 xy + \beta_4 x^2 + \dots,$$

Modèles non-paramétriques

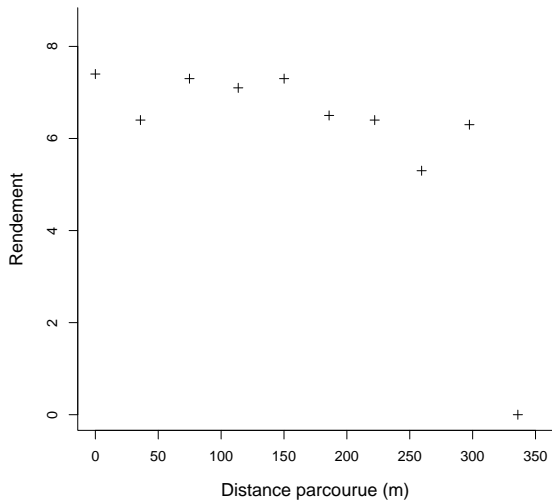
- Interpolation : s_0 site sans mesure

$$\hat{\mu}(s_0)?$$

- Approximation locale de $\mu(s)$

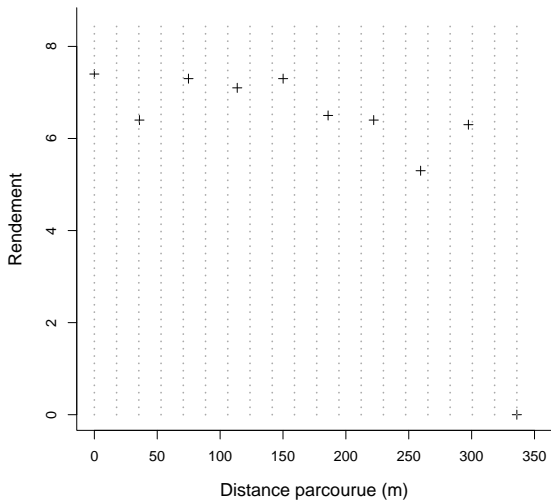


Interpolation linéaire



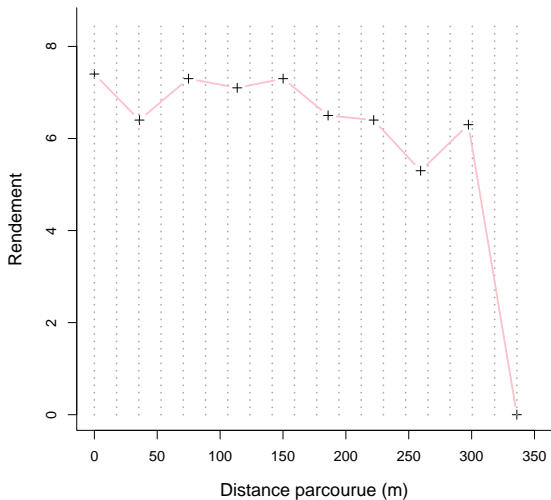


Interpolation linéaire



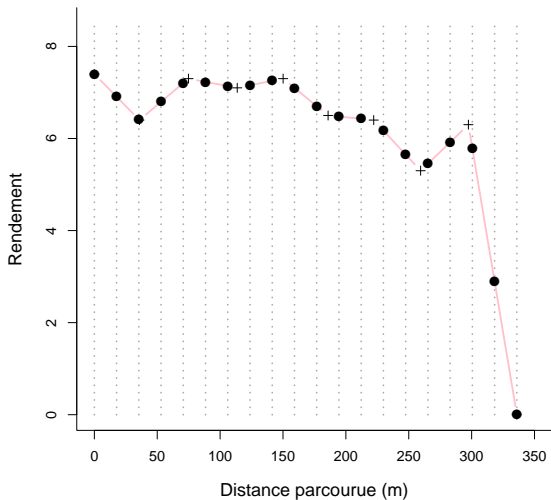


Interpolation linéaire



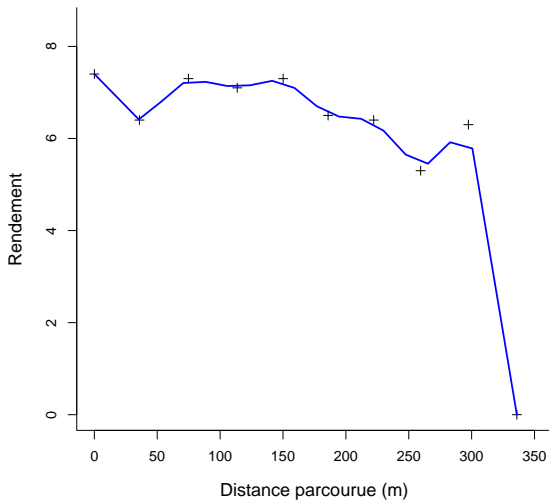


Interpolation linéaire



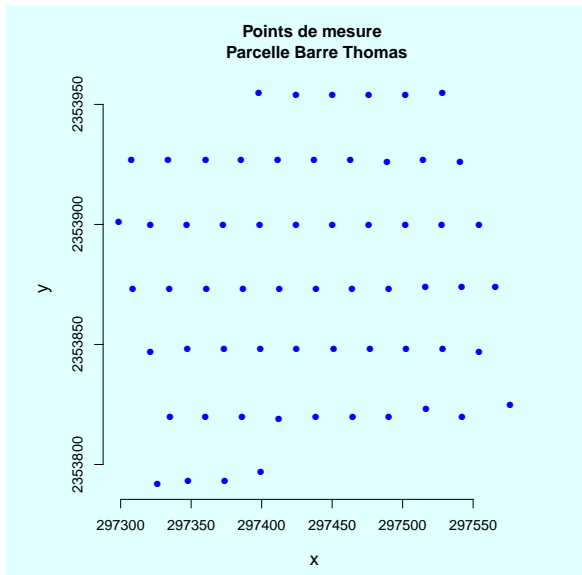


Interpolation linéaire



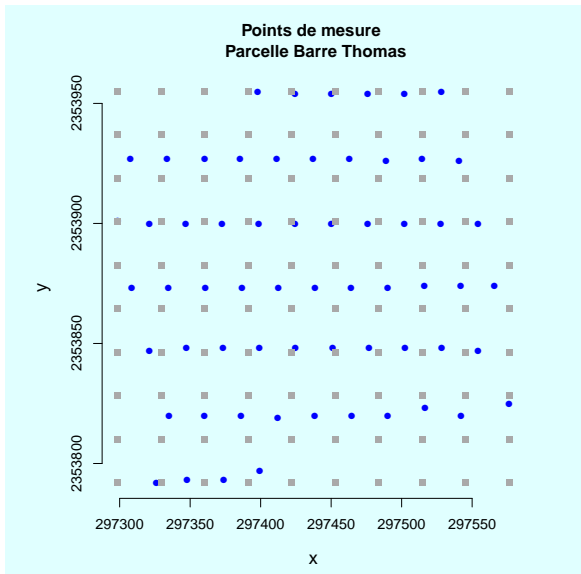


Interpolation linéaire



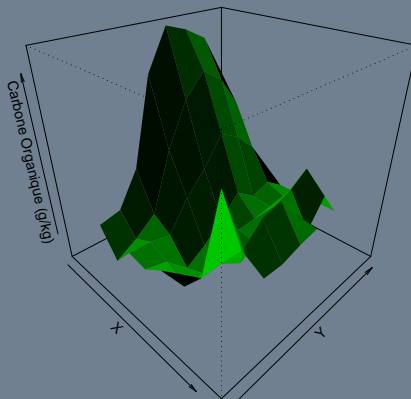


Interpolation linéaire



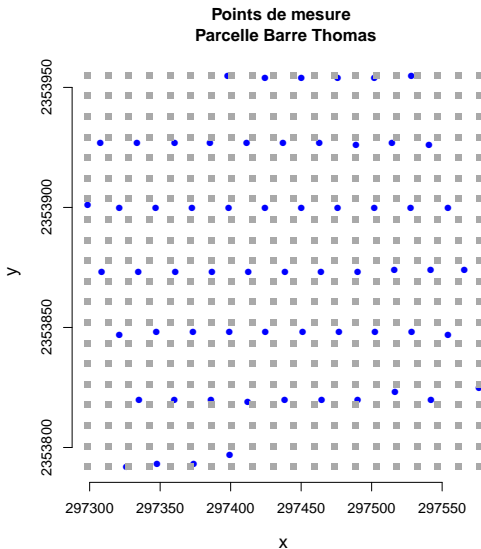


Interpolation linéaire



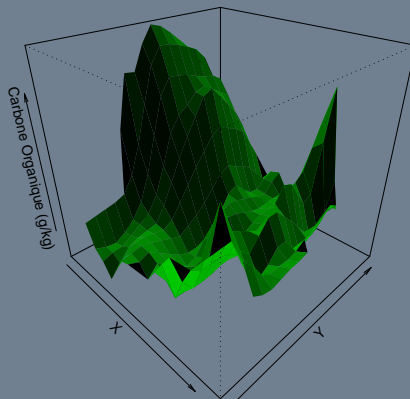


Interpolation linéaire





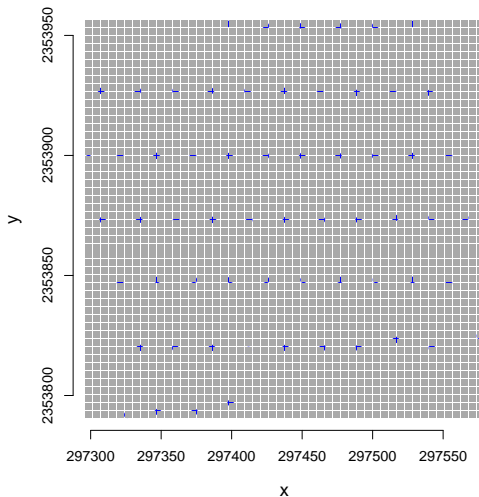
Interpolation linéaire





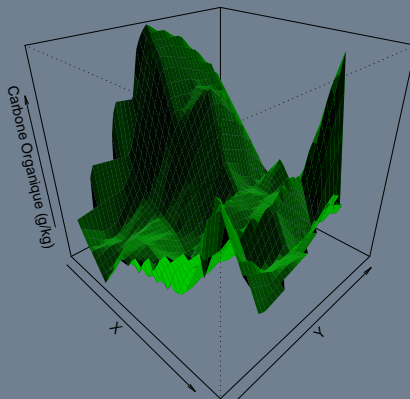
Interpolation linéaire

Points de mesure
Parcelle Barre Thomas



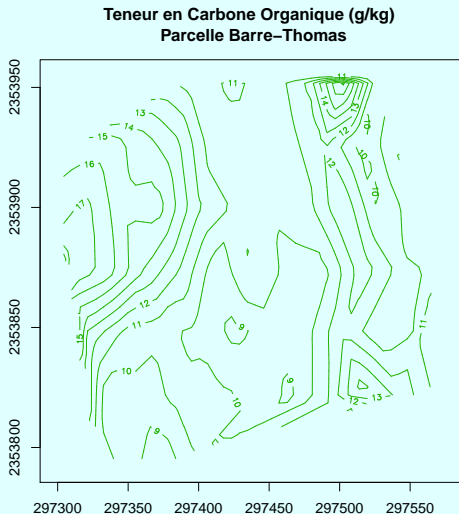


Interpolation linéaire



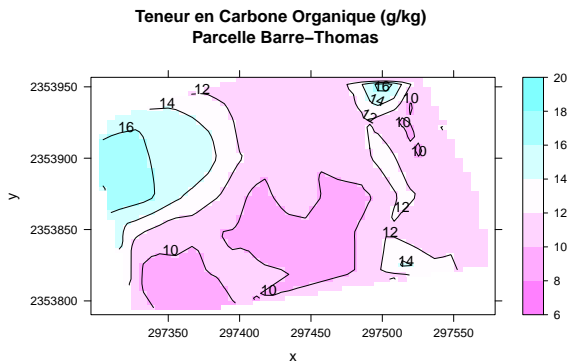


Interpolation linéaire





Interpolation linéaire





Pondération par la distance inverse

Estimation par moyenne pondérée mobile

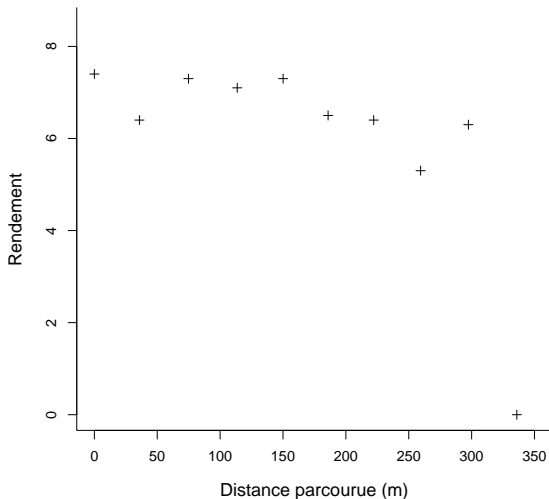
$$\hat{\mu}(s_0) = \frac{\sum_{i=1}^n p_i(s_0)Z(s_i)}{\sum_{i=1}^n p_i(s_0)},$$

$$\text{où } p_i(s_0) = \frac{1}{d(s_0, s_i)},$$

$$\hat{\mu}(s_i) = Z(s_i)$$

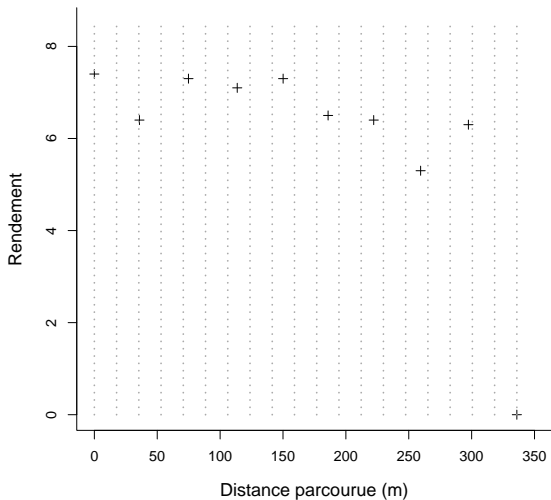


Pondération par la distance inverse



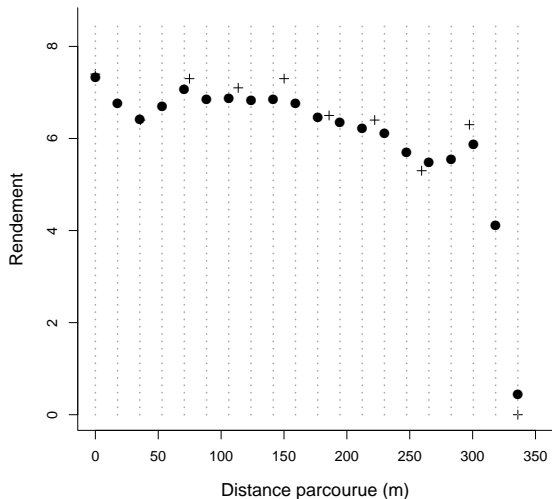


Pondération par la distance inverse



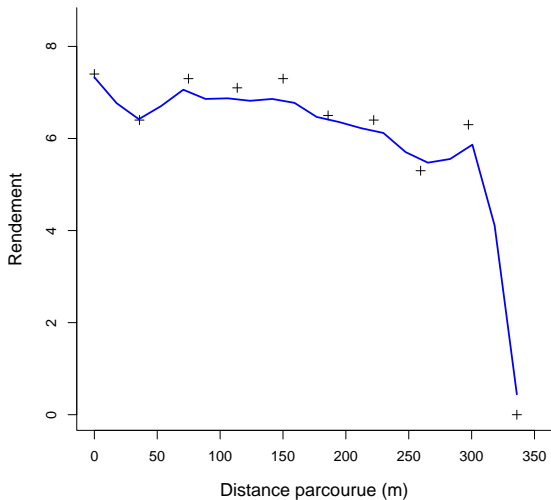


Pondération par la distance inverse



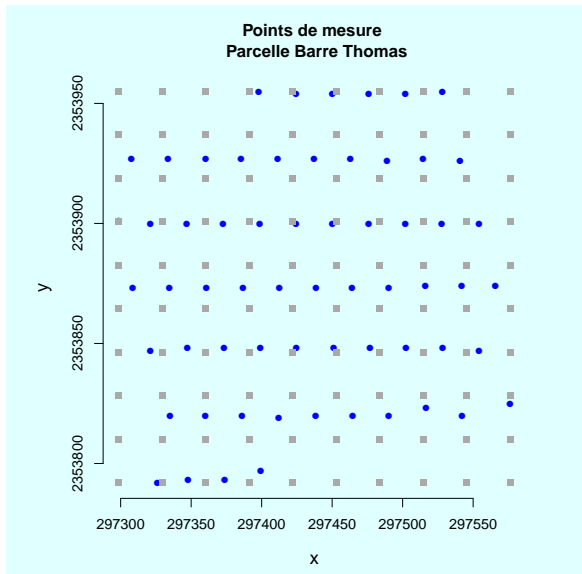


Pondération par la distance inverse



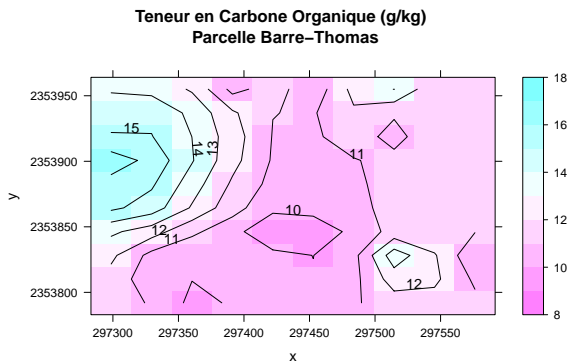


Pondération par la distance inverse



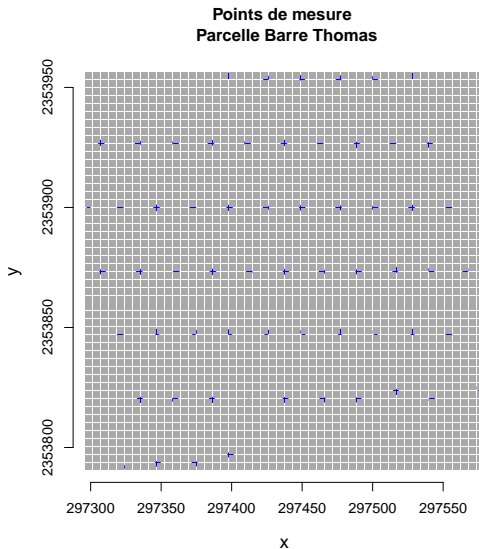


Pondération par la distance inverse



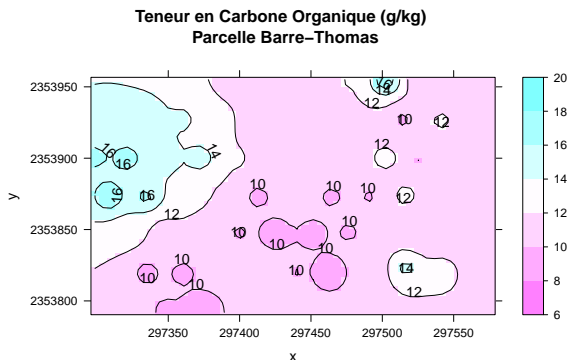


Pondération par la distance inverse





Pondération par la distance inverse



Régression locale pondérée

Estimation par régression pondérée mobile

$\mathcal{V}_0(\alpha)$ ensemble des 100α % des s_i les plus proches de s_0

$\lambda_0(\alpha)$ distance maximale entre un site de $\mathcal{V}_0(\alpha)$ et s_0

$\hat{D}_0(s)$ droite de régression autour de s_0

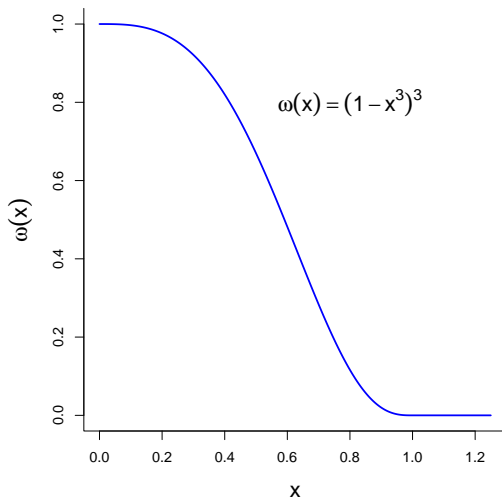
$$\sum_{i=1}^n p_i(s_0) \left[\hat{D}_0(s_i) - Z(s_i) \right]^2 \text{ est aussi petit que possible}$$

$$\text{où } p_i(s_0) = \omega \left(\frac{d(s_0, s_i)}{\lambda_0(\alpha)} \right)$$



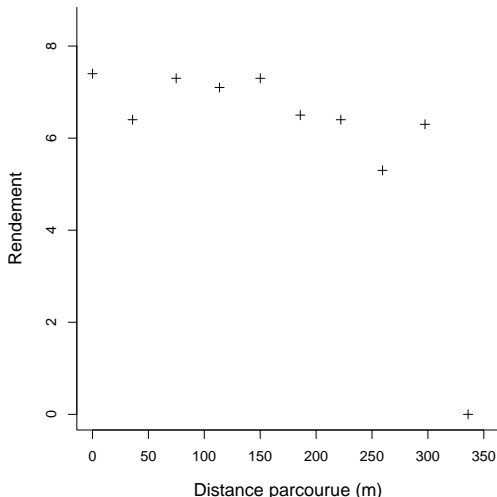
Régression locale pondérée

Fonction de poids tricube



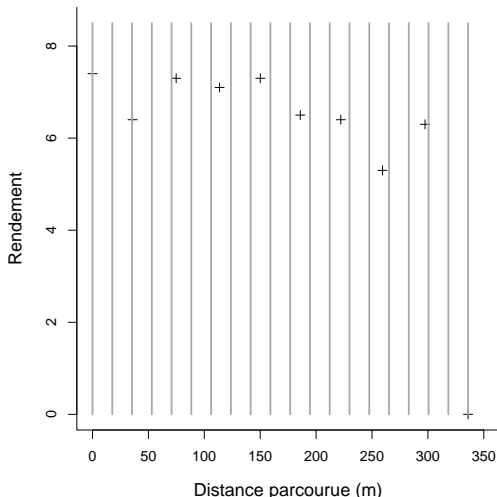


Régression locale pondérée



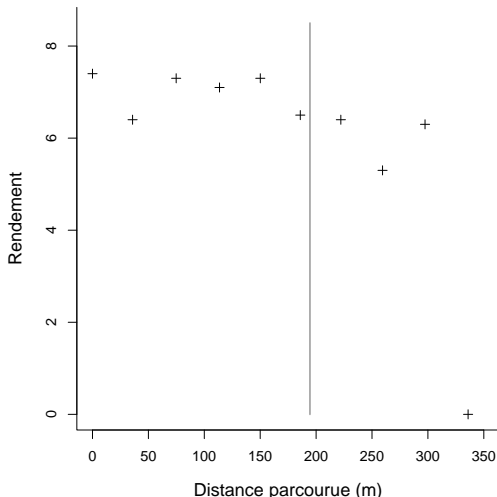


Régression locale pondérée



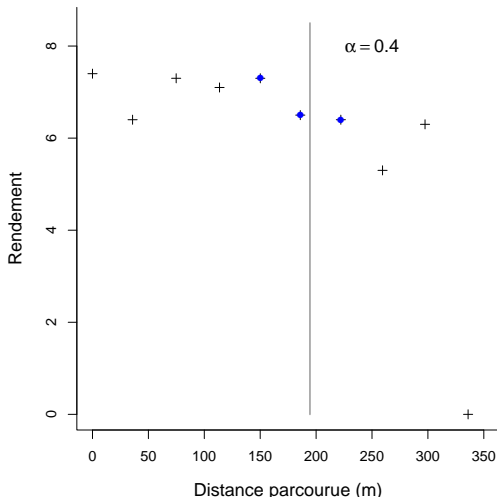


Régression locale pondérée



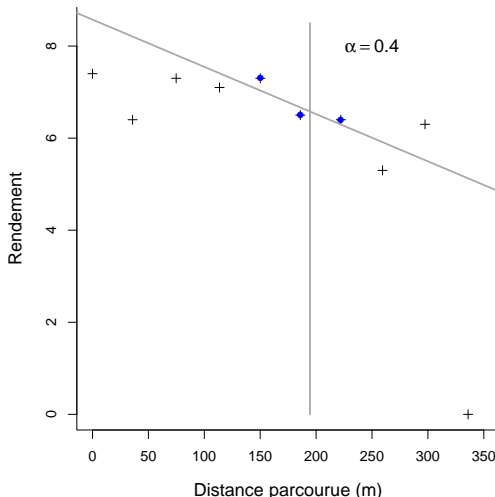


Régression locale pondérée



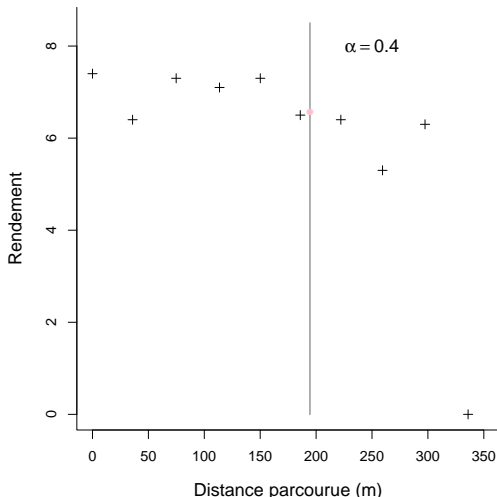


Régression locale pondérée



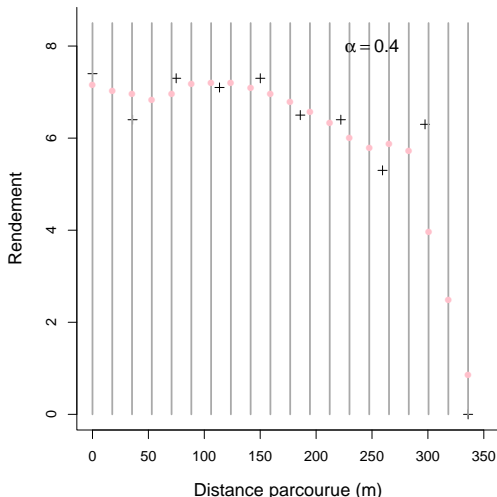


Régression locale pondérée



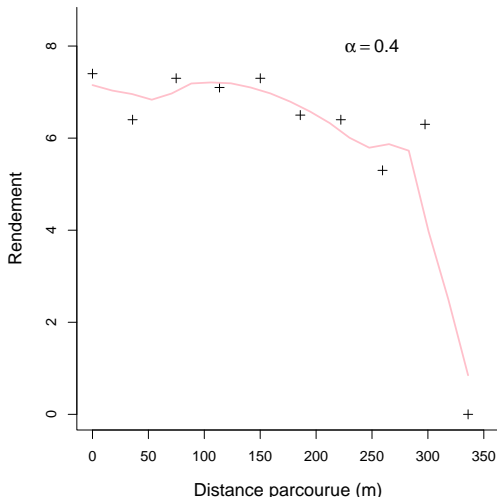


Régression locale pondérée



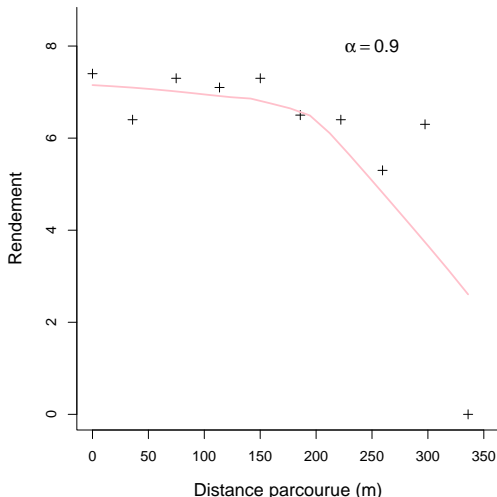


Régression locale pondérée



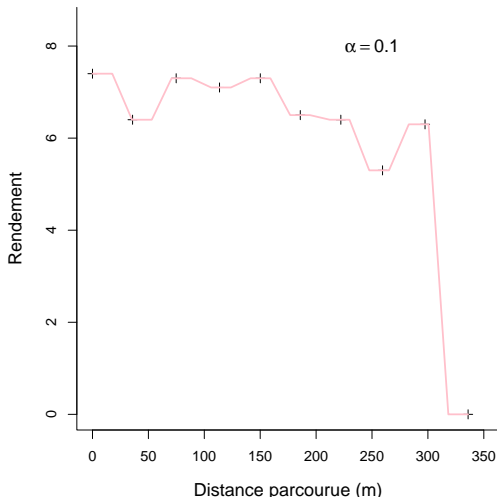


Régression locale pondérée



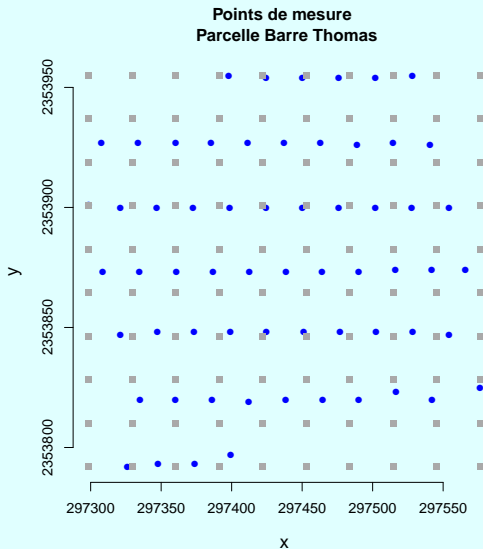


Régression locale pondérée



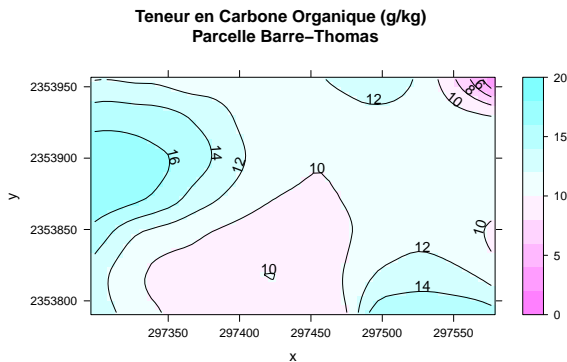


Régression locale pondérée

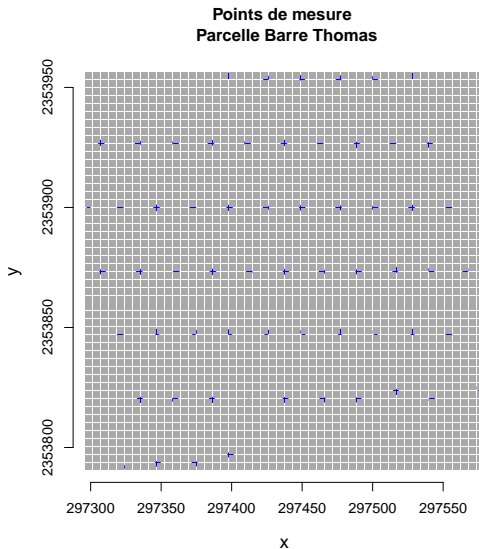




Régression locale pondérée

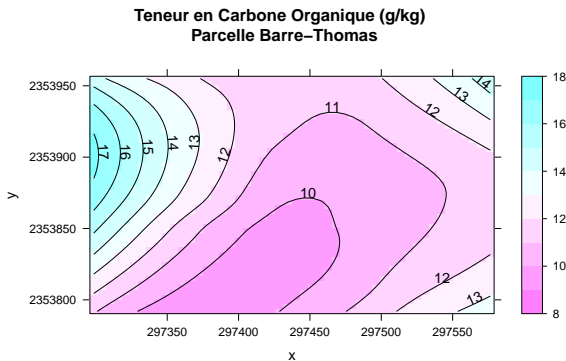


Régression locale pondérée





Régression locale pondérée





Régression locale pondérée - Optimisation

Choix de α

- Maximisation de la qualité d'ajustement

$$\text{SCER}(\alpha) = \sum_{i=1}^n [Z(s_i) - \hat{Z}(s_i)]^2$$



Régression locale pondérée - Optimisation

Choix de α

- Maximisation de la qualité d'ajustement

$$\text{SCER}(\alpha) = \sum_{i=1}^n [Z(s_i) - \hat{Z}(s_i)]^2$$

- Compromis qualité d'ajustement/complexité du modèle

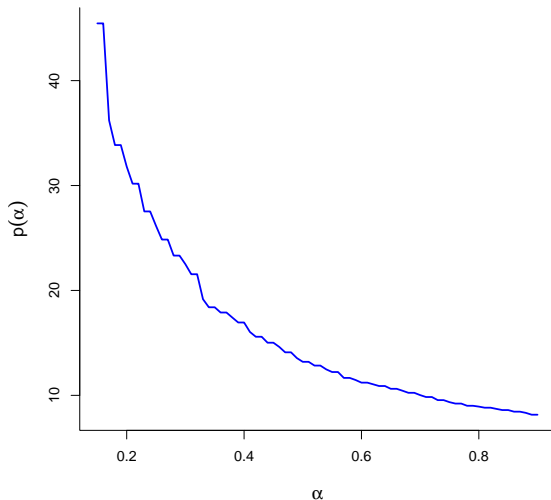
$$\text{AIC}(\alpha) = n \log \left[\frac{\text{SCER}(\alpha)}{n} \right] + 2p(\alpha)$$

où $p(\alpha)$ est le nombre équivalent de paramètres



Régression locale pondérée - Optimisation

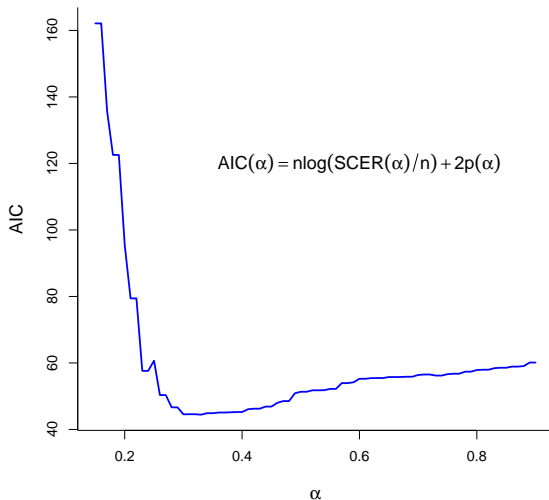
Choix de α . Nombre équivalent de paramètres





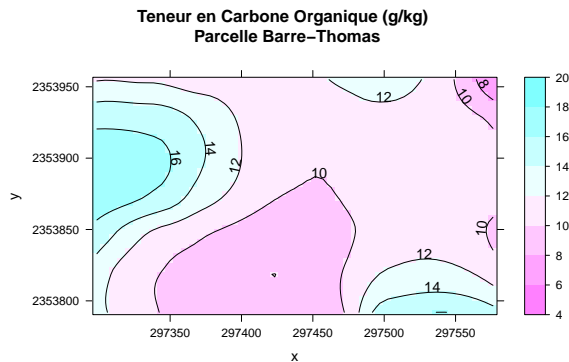
Régression locale pondérée - Optimisation

Choix de α





Régression locale pondérée - Optimisation



Plan du cours

- 1 Introduction
 - Problématiques
 - Modélisation spatiale
- 2 Description statistique de la variabilité spatiale
 - Données géostatistiques
 - Lissage de données spatiales
- 3 **Modélisation géostatistique**
 - Modèles de dépendance spatiale
 - Analyse variographique
 - Prédiction
- 4 Modélisation multivariée
- 5 Perspectives

Impact de la dépendance

Modèle de dépendance spatiale 1D

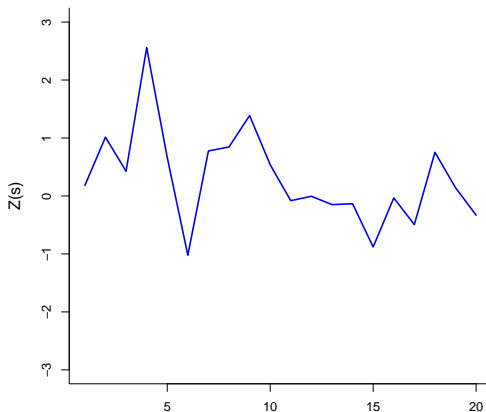
$$\mathbb{E}[Z(s)] = \mu, \quad \text{Cov}[Z(s_1), Z(s_2)] = \sigma^2 \rho^{|s_2 - s_1|}, \quad 0 < \rho < 1$$

Impact de la dépendance

Modèle de dépendance spatiale 1D

$$\mathbb{E}[Z(s)] = \mu, \quad \text{Cov}[Z(s_1), Z(s_2)] = \sigma^2 \rho^{|s_2 - s_1|}, \quad 0 < \rho < 1$$

Simulation de Z . $\rho = 0$

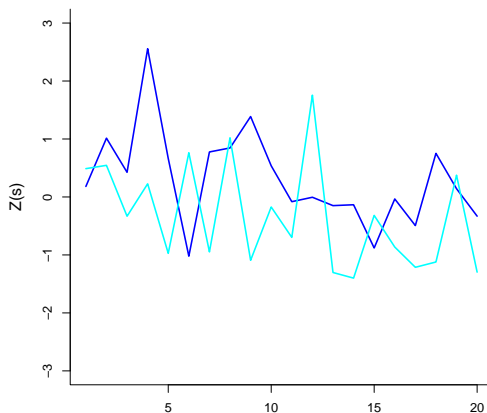


Impact de la dépendance

Modèle de dépendance spatiale 1D

$$\mathbb{E}[Z(s)] = \mu, \quad \text{Cov}[Z(s_1), Z(s_2)] = \sigma^2 \rho^{|s_2 - s_1|}, \quad 0 < \rho < 1$$

Simulation de Z . $\rho = 0$

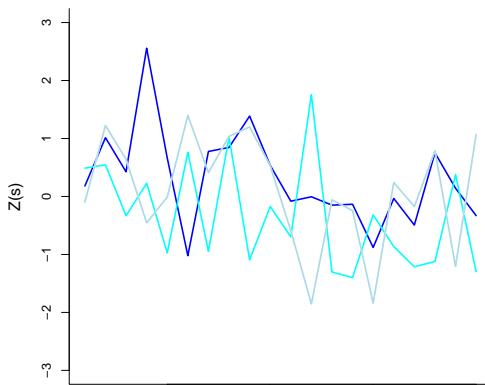


Impact de la dépendance

Modèle de dépendance spatiale 1D

$$\mathbb{E}[Z(s)] = \mu, \quad \text{Cov}[Z(s_1), Z(s_2)] = \sigma^2 \rho^{|s_2 - s_1|}, \quad 0 < \rho < 1$$

Simulation de Z. $\rho = 0$

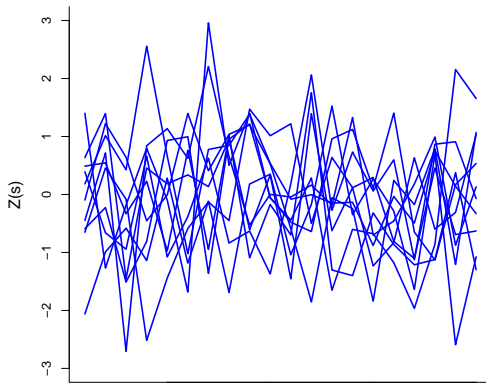


Impact de la dépendance

Modèle de dépendance spatiale 1D

$$\mathbb{E}[Z(s)] = \mu, \quad \text{Cov}[Z(s_1), Z(s_2)] = \sigma^2 \rho^{|s_2 - s_1|}, \quad 0 < \rho < 1$$

Simulation de Z. $\rho = 0$



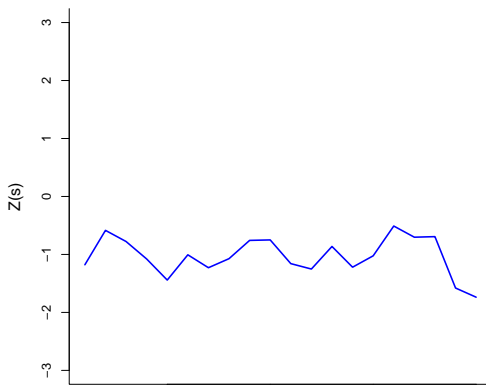


Impact de la dépendance

Modèle de dépendance spatiale 1D

$$\mathbb{E}[Z(s)] = \mu, \quad \text{Cov}[Z(s_1), Z(s_2)] = \sigma^2 \rho^{|s_2 - s_1|}, \quad 0 < \rho < 1$$

Simulation de Z . $\rho = 0.9$



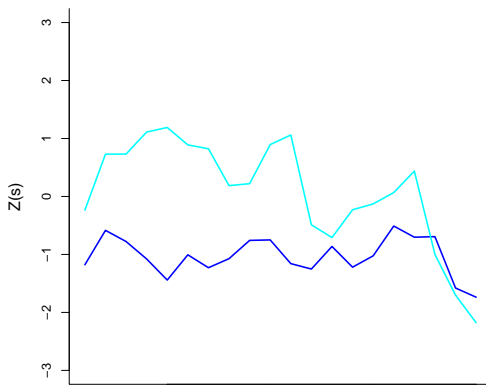


Impact de la dépendance

Modèle de dépendance spatiale 1D

$$\mathbb{E}[Z(s)] = \mu, \quad \text{Cov}[Z(s_1), Z(s_2)] = \sigma^2 \rho^{|s_2 - s_1|}, \quad 0 < \rho < 1$$

Simulation de Z . $\rho = 0.9$



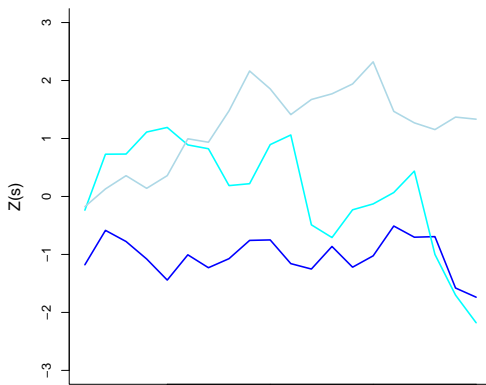


Impact de la dépendance

Modèle de dépendance spatiale 1D

$$\mathbb{E}[Z(s)] = \mu, \quad \text{Cov}[Z(s_1), Z(s_2)] = \sigma^2 \rho^{|s_2 - s_1|}, \quad 0 < \rho < 1$$

Simulation de Z . $\rho = 0.9$

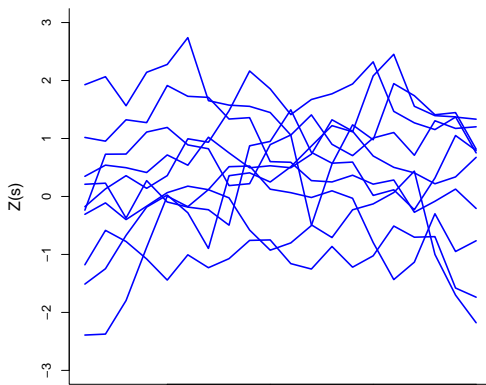


Impact de la dépendance

Modèle de dépendance spatiale 1D

$$\mathbb{E}[Z(s)] = \mu, \quad \text{Cov}[Z(s_1), Z(s_2)] = \sigma^2 \rho^{|s_2 - s_1|}, \quad 0 < \rho < 1$$

Simulation de Z . $\rho = 0.9$





Impact de la dépendance

Modèle de dépendance spatiale 1D

$$\mathbb{E}[Z(s)] = \mu, \quad \text{Cov}[Z(s_1), Z(s_2)] = \sigma^2 \rho^{|s_2 - s_1|}, \quad 0 < \rho < 1$$

Modèle des variations locales

- Stationnarité stricte

$$\begin{aligned} \mathcal{L}[\varepsilon(s_1), \varepsilon(s_2), \dots, \varepsilon(s_n)] = \\ \mathcal{L}[\varepsilon(s_1 + h), \varepsilon(s_2 + h), \dots, \varepsilon(s_n + h)] \end{aligned}$$

- Stationnarité à l'ordre 2

$$\forall(s, s'), \text{Cov}[\varepsilon(s), \varepsilon(s')] = C(s - s')$$

où C est la fonction d'autocovariance de ε .

Auto-covariance d'un processus spatial

Propriétés

- Symétrie : $C(h) = C(-h)$
- Majoration : $|C(h)| \leq C(0)$
- Portée : a tel que $C(h) = 0, \forall h, ||h|| \geq a$

Analyse de C :

- Direction de h : structuration spatiale de la variabilité
- Norme de h : régularité des variations spatiales

Isotropie des variations locales

$$\forall (s, s'), \text{Cov} [\varepsilon(s), \varepsilon(s')] = C(||s - s'||)$$

Variogramme

Accroissements d'ordre h d'origine s

$$\Delta_h(s) = \varepsilon(s + h) - \varepsilon(s)$$

Variogramme $\gamma(h)$

$$\begin{aligned} \text{Var}[\Delta_h(s)] &= 2[C(0) - C(h)], \\ &= 2\gamma(h) \end{aligned}$$

Propriétés

- Symétrie : $\gamma(h) = \gamma(-h)$
- Comportement en 0 : $\gamma(h) \geq 0$ et $\gamma(0) = 0$
- Majoration : $|\gamma(h)| \leq 2C(0)$
- Portée d'un site : a tel que $\gamma(h) = C(0), \forall h, \|h\| \geq a$



Variogramme et régularité d'un processus spatial

- Irrégularité : indépendance locale $[\gamma(h) = C(0)]$
- Régularité : dépendance locale maximale $[\gamma(h) = 0]$

Régularité :

- Non-covariance entre deux sites : γ constant
- Non-continuité de ε : effet de pépite

$$\lim_{h \rightarrow 0} E \left\{ [\varepsilon(s+h) - \varepsilon(s)]^2 \right\} \neq 0,$$

$$\lim_{h \rightarrow 0} \{\gamma(h)\} \neq 0,$$

γ n'est pas continue en 0

- Non-dérivabilité de ε

$$\lim_{h \rightarrow 0} E \left\{ \left[\frac{\varepsilon(s+h) - \varepsilon(s)}{h} - \varepsilon'(s) \right]^2 \right\} \neq 0,$$

γ n'est pas dérivable au second ordre en 0 et $\gamma'(0) \neq 0$



Variogramme empirique

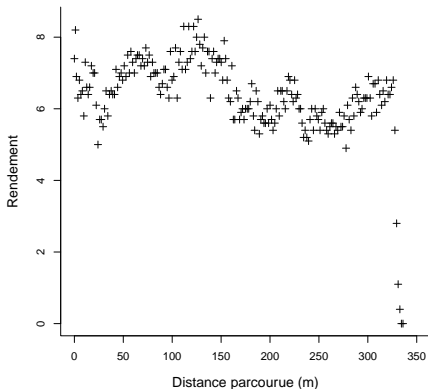
Estimation $\hat{\gamma}$ du variogramme

$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$

Variogramme empirique

Estimation $\hat{\gamma}$ du variogramme

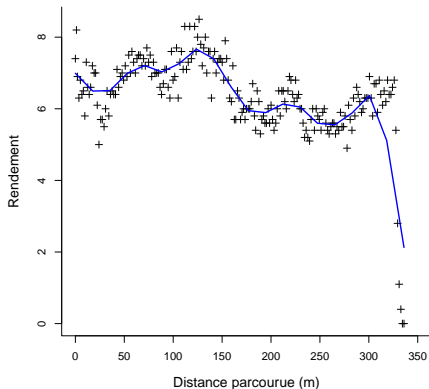
$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$



Variogramme empirique

Estimation $\hat{\gamma}$ du variogramme

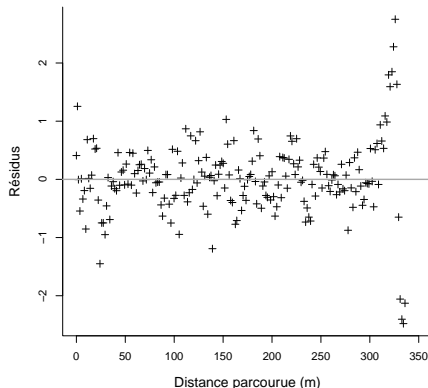
$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$



Variogramme empirique

Estimation $\hat{\gamma}$ du variogramme

$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$

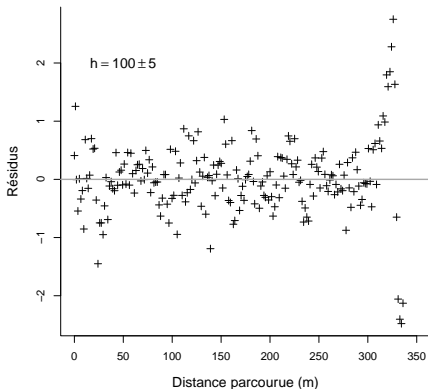




Variogramme empirique

Estimation $\hat{\gamma}$ du variogramme

$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$

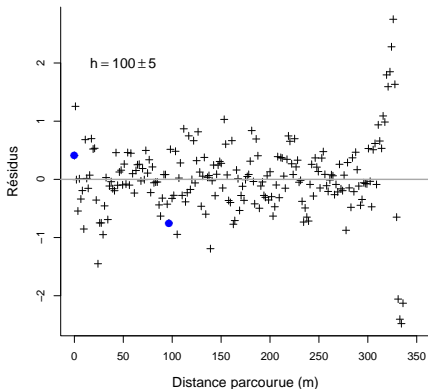




Variogramme empirique

Estimation $\hat{\gamma}$ du variogramme

$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$

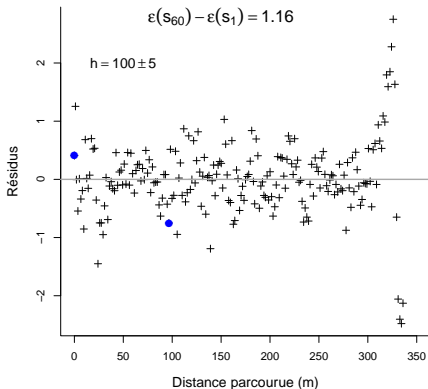




Variogramme empirique

Estimation $\hat{\gamma}$ du variogramme

$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$

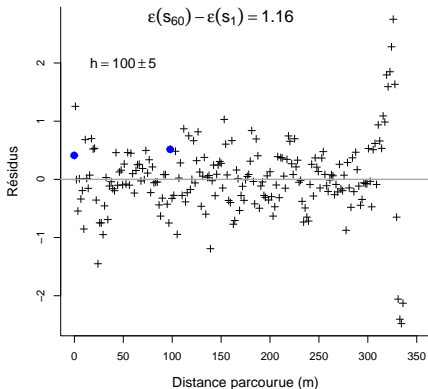




Variogramme empirique

Estimation $\hat{\gamma}$ du variogramme

$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$

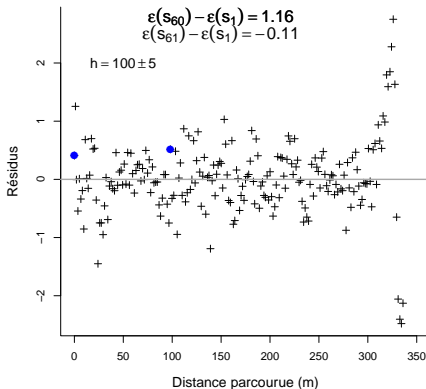




Variogramme empirique

Estimation $\hat{\gamma}$ du variogramme

$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$

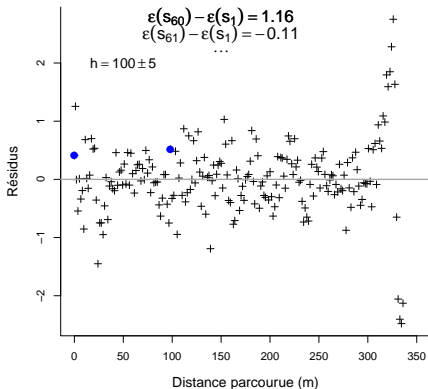




Variogramme empirique

Estimation $\hat{\gamma}$ du variogramme

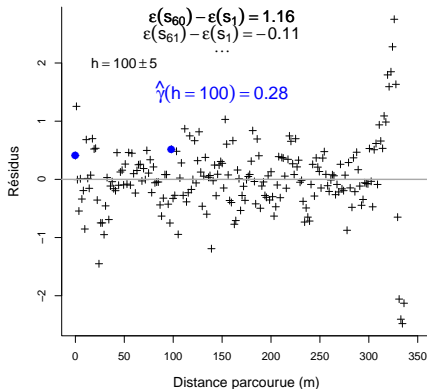
$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$



Variogramme empirique

Estimation $\hat{\gamma}$ du variogramme

$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$

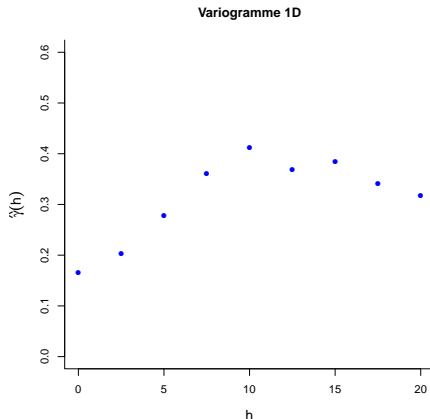




Variogramme empirique

Estimation $\hat{\gamma}$ du variogramme

$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$

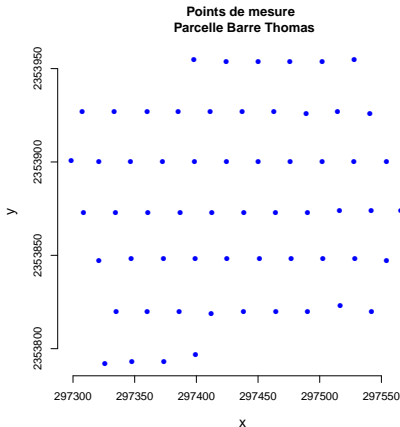




Variogramme empirique

Estimation $\hat{\gamma}$ du variogramme

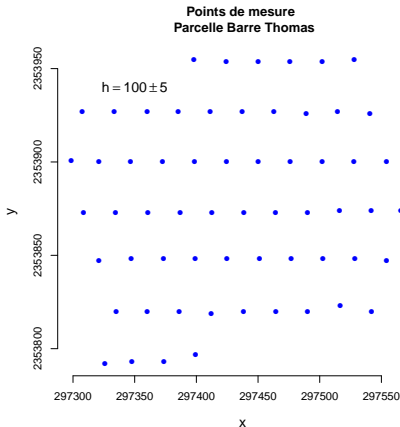
$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$



Variogramme empirique

Estimation $\hat{\gamma}$ du variogramme

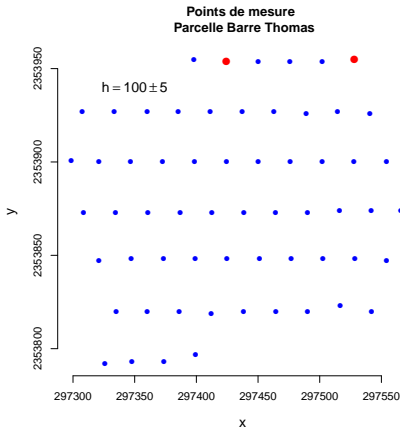
$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$



Variogramme empirique

Estimation $\hat{\gamma}$ du variogramme

$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$

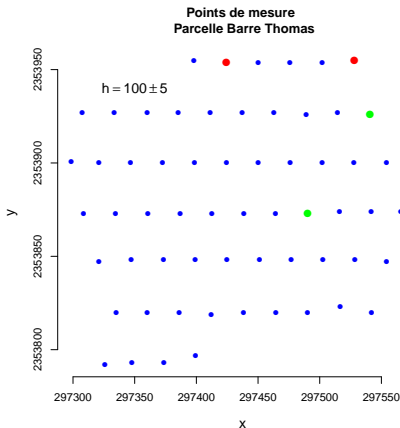




Variogramme empirique

Estimation $\hat{\gamma}$ du variogramme

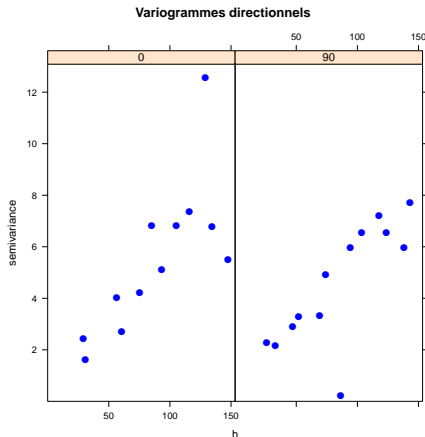
$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$



Variogramme empirique

Estimation $\hat{\gamma}$ du variogramme

$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$

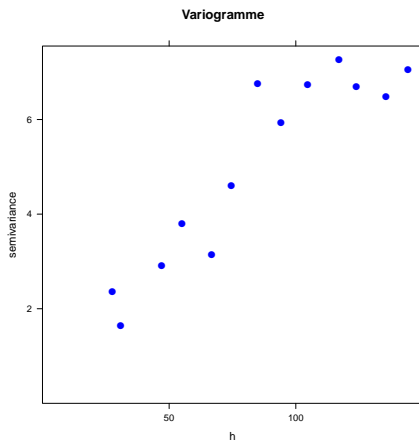




Variogramme empirique

Estimation $\hat{\gamma}$ du variogramme

$$\hat{\gamma}(h) = \widehat{\text{Var}}[\Delta_h(s)]$$





Modèles de variogramme

Modèles non-stationnaires

- Variogramme linéaire : $\gamma(h) = \gamma_0 + \beta h$, $\gamma_0 > 0$, $\beta > 0$
- Variogramme puissance : $\gamma(h) = \gamma_0 + \beta h^a$, $a > 0$



Modèles de variogramme

Modèles non-stationnaires

- Variogramme linéaire : $\gamma(h) = \gamma_0 + \beta h$, $\gamma_0 > 0$, $\beta > 0$
- Variogramme puissance : $\gamma(h) = \gamma_0 + \beta h^a$, $a > 0$

Modèles stationnaires

- Variogramme gaussien : $\gamma(h) = \gamma_0 + \sigma^2 \left[1 - e^{-\left(\frac{h}{a}\right)^2} \right]$

où $\sigma^2 > 0$, $\gamma_0 > 0$, $a > 0$.

Modèles de variogramme

Modèles non-stationnaires

- Variogramme linéaire : $\gamma(h) = \gamma_0 + \beta h, \gamma_0 > 0, \beta > 0$
- Variogramme puissance : $\gamma(h) = \gamma_0 + \beta h^a, a > 0$

Modèles stationnaires

- Variogramme gaussien : $\gamma(h) = \gamma_0 + \sigma^2 \left[1 - e^{-\left(\frac{h}{a}\right)^2} \right]$
- Variogramme exponentiel : $\gamma(h) = \gamma_0 + \sigma^2 \left[1 - e^{-\frac{h}{a}} \right]$

où $\sigma^2 > 0, \gamma_0 > 0, a > 0$.

Modèles de variogramme

Modèles non-stationnaires

- Variogramme linéaire : $\gamma(h) = \gamma_0 + \beta h$, $\gamma_0 > 0$, $\beta > 0$
- Variogramme puissance : $\gamma(h) = \gamma_0 + \beta h^a$, $a > 0$

Modèles stationnaires

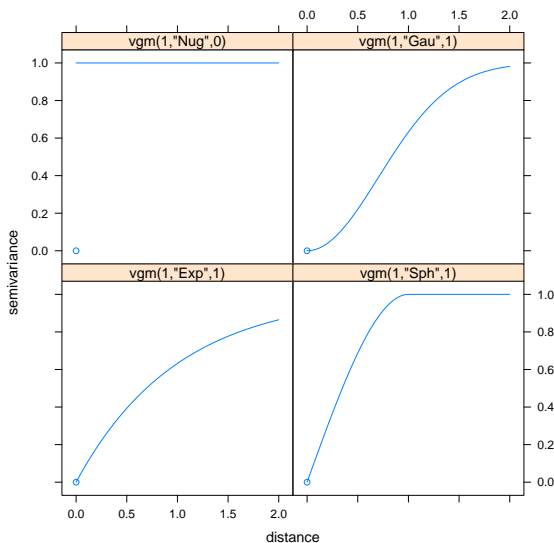
- Variogramme gaussien : $\gamma(h) = \gamma_0 + \sigma^2 \left[1 - e^{-\left(\frac{h}{a}\right)^2} \right]$
- Variogramme exponentiel : $\gamma(h) = \gamma_0 + \sigma^2 \left[1 - e^{-\frac{h}{a}} \right]$
- Variogramme sphérique :

$$\gamma(h) = \gamma_0 + \sigma^2 \begin{cases} \frac{3}{2} \frac{\|h\|}{a} - \frac{1}{2} \frac{\|h\|^3}{a^3} & \text{si } \|h\| < a \\ 1 & \text{si } \|h\| \geq a \end{cases},$$

où $\sigma^2 > 0$, $\gamma_0 > 0$, $a > 0$.



Modèles de variogramme





Paramètres d'un modèle de variogramme

- γ_0 est la pépité du processus,
- σ^2 est le paramètre de seuil :

$$\lim_{h \rightarrow +\infty} [\gamma(h) - \gamma_0] = \sigma^2.$$

- a est le paramètre de portée



Ajustement d'un modèle de variogramme

Méthode des moindres carrés

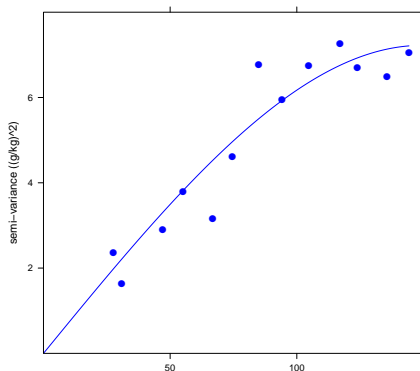
$$\hat{\theta} = \arg \min_{\theta} \sum_{i=1}^n [\hat{\gamma}(h_i) - \gamma(h_i ; \theta)]^2$$

Ajustement d'un modèle de variogramme

Méthode des moindres carrés

$$\hat{\theta} = \arg \min_{\theta} \sum_{i=1}^n [\hat{\gamma}(h_i) - \gamma(h_i; \theta)]^2$$

Ajustement d'un variogramme sphérique
Nugget = 0, Psill = 7.22, Range = 149.24



Krigeage

Modèle spatial : $Z(s) = \mu(s) + \varepsilon(s)$, $\text{Var}(\varepsilon) = C$

- Krigeage ordinaire : $\mu(s) = \mu$
- Krigeage universel : $\mu(s)$ paramétrique linéaire



Krigeage

Modèle spatial : $Z(s) = \mu(s) + \varepsilon(s)$, $\text{Var}(\varepsilon) = C$

- Krigeage ordinaire : $\mu(s) = \mu$
- Krigeage universel : $\mu(s)$ paramétrique linéaire

Interpolation : $Z(s_0)$ prédite par $\hat{Z}(s_0)$

- $\hat{Z}(s_0)$ combinaison linéaire des $Z(s_i)$,
- $\mathbb{E} [Z(s_0) - \hat{Z}(s_0)] = 0$,
- $\text{Var} [Z(s_0) - \hat{Z}(s_0)]$ minimale.



Krigeage

Modèle spatial : $Z(s) = \mu(s) + \varepsilon(s)$, $\text{Var}(\varepsilon) = C$

- Krigeage ordinaire : $\mu(s) = \mu$
- Krigeage universel : $\mu(s)$ paramétrique linéaire

Interpolation : $Z(s_0)$ prédite par $\hat{Z}(s_0)$

- $\hat{Z}(s_0)$ combinaison linéaire des $Z(s_i)$,
- $\mathbb{E} [Z(s_0) - \hat{Z}(s_0)] = 0$,
- $\text{Var} [Z(s_0) - \hat{Z}(s_0)]$ minimale.

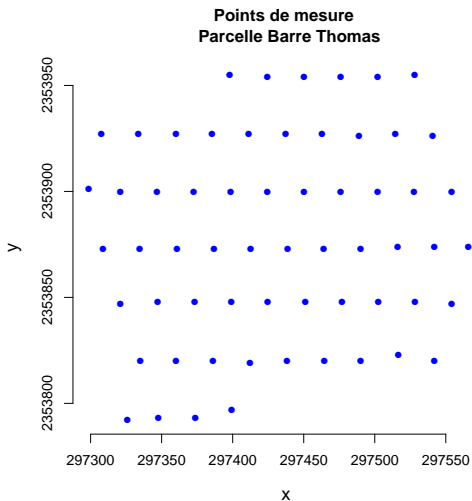
Prédicteur optimal :

$$\hat{Z}(s_0) = \hat{\mu}(s_0) + c(S, s_0)' C^{-1} [Z_n - \hat{\mu}(s)],$$

$$\hat{Z}(s_i) = Z(s_i) \quad [\text{Universalité}]$$

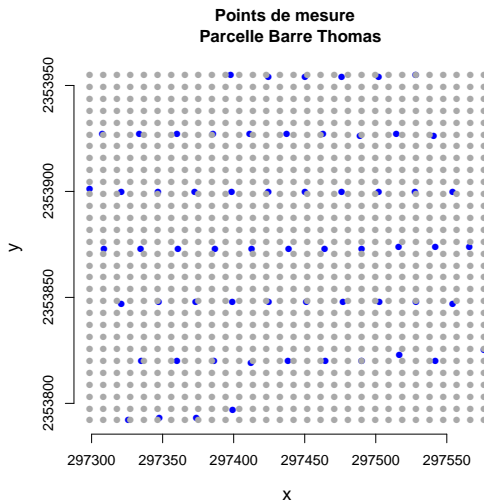


Krigeage



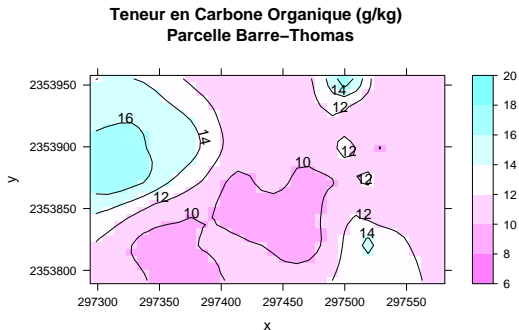


Krigeage





Krigeage



Plan du cours

- 1 Introduction
 - Problématiques
 - Modélisation spatiale
- 2 Description statistique de la variabilité spatiale
 - Données géostatistiques
 - Lissage de données spatiales
- 3 Modélisation géostatistique
 - Modèles de dépendance spatiale
 - Analyse variographique
 - Prédiction
- 4 **Modélisation multivariée**
- 5 Perspectives

Modèle géostatistique multivarié

Données spatiales

<i>Site</i>	LOCALISATION DU SITE			CARACTÉRISTIQUES			
<i>S</i>	<i>X</i>	<i>Y</i>	...	<i>Z</i> ₁	<i>Z</i> ₂	...	<i>Z</i> _{<i>k</i>}
<i>s</i> ₁	<i>x</i> ₁	<i>y</i> ₁	...	<i>Z</i> ₁ (<i>s</i> ₁)	<i>Z</i> ₂ (<i>s</i> ₁)	...	<i>Z</i> _{<i>k</i>} (<i>s</i> ₁)
<i>s</i> ₂	<i>x</i> ₂	<i>y</i> ₂	...	<i>Z</i> ₁ (<i>s</i> ₂)	?	...	<i>Z</i> _{<i>k</i>} (<i>s</i> ₂)
⋮	⋮	⋮		⋮	⋮	...	⋮
<i>s</i> _{<i>n</i>}	<i>x</i> ₁	<i>y</i> ₁	...	<i>Z</i> ₁ (<i>s</i> _{<i>n</i>})	<i>Z</i> ₂ (<i>s</i> _{<i>n</i>})	...	?

Modèle de dépendance multivariée : cross-variogrammes

$$\gamma_{jk}(h) = \frac{1}{2} \text{Cov}(\Delta_h^{(j)}(s), \Delta_h^{(k)}(s)) \quad \text{ou}$$

$$\gamma_{jk}(h) = \frac{1}{2} \text{Var}[Z_k(s+h) - Z_j(s)]$$



Co-krigeage

Interpolation : $Z_j(s_0)$ prédite par $\hat{Z}_j(s_0)$

- $\hat{Z}_j(s_0)$ combinaison linéaire des $Z_k(s_i)$, $k = 1, 2, \dots$,
- $\mathbb{E} [Z_j(s_0) - \hat{Z}_j(s_0)] = 0$,
- $\text{Var} [Z(s_0) - \hat{Z}(s_0)]$ minimale.

Co-krigeage

Interpolation : $Z_j(s_0)$ prédite par $\hat{Z}_j(s_0)$

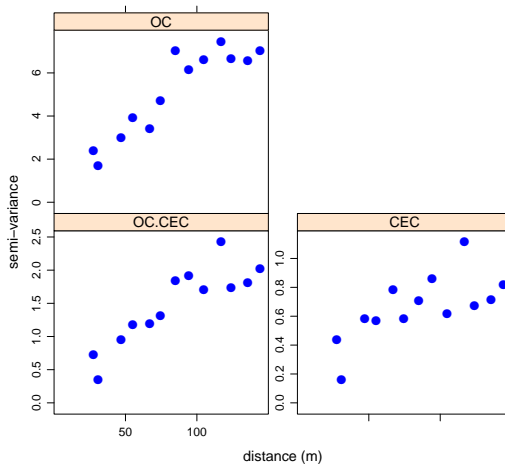
- $\hat{Z}_j(s_0)$ combinaison linéaire des $Z_k(s_i)$, $k = 1, 2, \dots$,
- $\mathbb{E} [Z_j(s_0) - \hat{Z}_j(s_0)] = 0$,
- $\text{Var} [Z(s_0) - \hat{Z}(s_0)]$ minimale.

Choix de la variable associée : corrélation (Carbone Org, CEC)
= 0.82



Co-krigeage

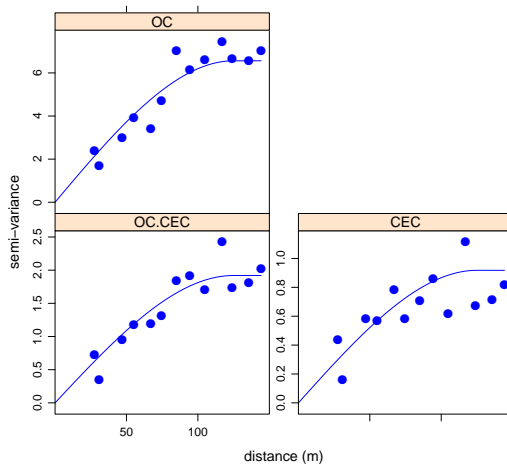
Ajustement d'un covariogramme sphérique





Co-krigeage

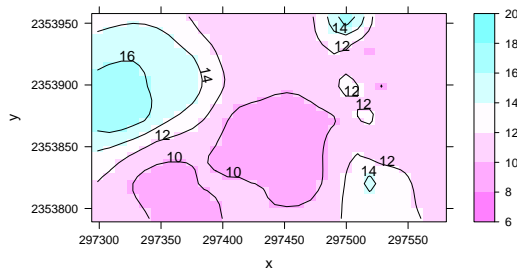
Ajustement d'un covariogramme sphérique



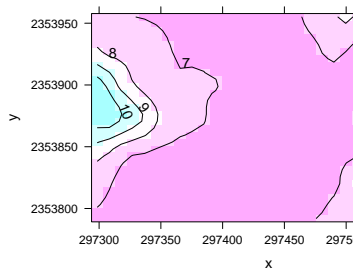


Co-krigeage

Teneur en Carbone Organique (g/kg)
Parcelle Barre-Thomas



CEC (cmol/kg)
Parcelle Barre-Thomas



Plan du cours

- 1 Introduction
 - Problématiques
 - Modélisation spatiale
- 2 Description statistique de la variabilité spatiale
 - Données géostatistiques
 - Lissage de données spatiales
- 3 Modélisation géostatistique
 - Modèles de dépendance spatiale
 - Analyse variographique
 - Prédiction
- 4 Modélisation multivariée
- 5 Perspectives



Perspectives

- Ouverture à des domaines d'application non-traditionnels
- Modélisation multivariée pour cartographies cohérentes
- Optimisation des ressources expérimentales
 - Choix des variables corrélées
 - Choix des sites expérimentaux