



Le semi-variogramme : exemple d'utilisation pour l'échantillonnage spatialisé et l'estimation

B. Tisseyre

UMR ITAP, Montpellier SupAgro/IRSTEA, France

Heliospir Novembre 2013

Introduction

- Développement adoption des systèmes de positionnement par satellites,
- Développement de systèmes de mesure embarqués dans les machines agricoles,

Généralisation de l'information à très haute résolution spatiale en agriculture

Heliospir Novembre 2013

Introduction



- Développement adoption des systèmes de positionnement par satellites,
- Développement de systèmes de mesure embarqués dans les machines agricoles,

Exemple : Cartes de rendement



Sources : ATV

Sources :
Montpellier
SupAgro/Pellenc
S.A.

Heliospir Novembre 2013

Introduction

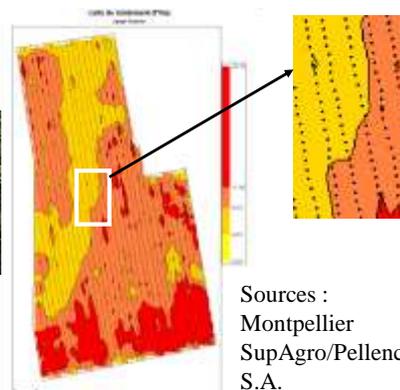


- Développement adoption des systèmes de positionnement par satellites,
- Développement de systèmes de mesure embarqués dans les machines agricoles,

Exemple : Cartes de rendement



Sources : ATV



Sources :
Montpellier
SupAgro/Pellenc
S.A.

Heliospir Novembre 2013

Une nouvelle vision de la Variance



D : domaine d'étude (parcelle)

VAR_T

1000 sites de mesure /ha

Heliospir Novembre 2013

Une nouvelle vision de la Variance



D : domaine d'étude (parcelle)

VAR_T

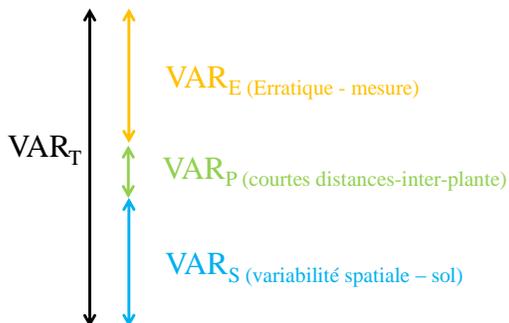
1000 sites de mesure /ha

Heliospir Novembre 2013

Une nouvelle vision de la Variance



D : domaine d'étude (parcelle)



Une définition qui dépend de l'utilisation/point de vue
Des phénomènes en apparence contradictoires

Une nouvelle vision de la Variance

- Variable régionalisée

D : domaine d'étude (parcelle),
Notion de fonction aléatoire sur le Domaine D,

Un semé de mesures effectué sur un domaine D est UNE réalisation de la fonction aléatoire

- en chaque site « s_i », $z(s_i)$ est une réalisation d'une variable aléatoire $Z(s_i)$,
- chaque variable aléatoire $Z(s_i)$, a son espérance « $E(Z(s_i))$ » et sa variance $VAR(Z(s_i))$,

$$Z(s) = \{Z(s_1), Z(s_2), \dots, Z(s_i)\}, \{s_1, s_2, \dots, s_i\} \in D$$

Où $Z(s)$ est la fonction aléatoire du domaine D

Une nouvelle vision de la Variance

- Variable régionalisée en géostatistique



D : domaine d'étude (parcelle),
Notion de fonction aléatoire sur le Domaine D,

Un semi de mesures effectué sur un domaine D est UNE réalisation de la fonction aléatoire

- en chaque site « s_i », $z(s_i)$ est une réalisation d'une variable aléatoire $Z(s_i)$,
- chaque variable aléatoire $Z(s_i)$, a son espérance « $E(Z(s_i))$ » et sa variance $VAR(Z(s_i))$,

$$Z(s) = \{Z(s_1), Z(s_2), \dots, Z(s_i)\}, \{s_1, s_2, \dots, s_i\} \in D$$

Où $Z(s)$ est la fonction aléatoire du domaine D

- Les Variables aléatoire ne sont pas indépendantes
(semi)-variance : mesure dissemblance entre deux variables aléatoires,

$$\gamma(s_i, s_j) = \frac{1}{2} \text{var}(Z(s_i) - Z(s_j))$$

Heliospir Novembre 2013

Une nouvelle vision de la Variance

- Variable régionalisée en géostatistique



Une seule réalisation de la fonction aléatoire

L'inférence de la fonction aléatoire nécessite de poser des hypothèses.

HP intrinsèque :

$$1) \quad \forall s, s+h, E[Z(s+h) - Z(s)] = 0$$

$$2) \quad \forall s, s+h, \text{var}[Z(s+h) - Z(s)] = 2 \cdot \gamma(h)$$

La variance de $Z(s+h) - Z(s)$ ne dépend donc que de h .
si $\gamma(h)$ est borné la covariance existe.

Heliospir Novembre 2013

Une nouvelle vision de la Variance

- Variable régionalisée en géostatistique

L'inférence du (semi)-variogramme

Le (semi)-variogramme : ensemble des semi-variances moyennes obtenues sur le domaine D pour des sites séparés d'une distance h.

Il est estimé par :

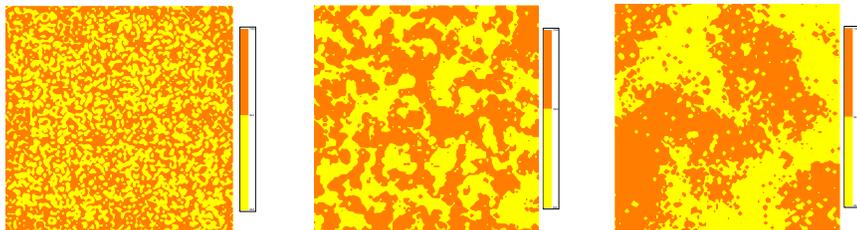
$$\hat{\gamma}(h) = \frac{1}{2 \times |N(h)|} \sum_{N(h)} (z(s_i) - z(s_j))^2$$

$$N(h) = \{ i, j / |s_i - s_j| = h \}$$

$|N(h)|$: nombre de paires distinctes de l'ensemble N(h)

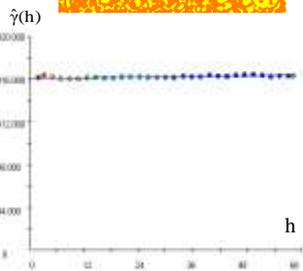
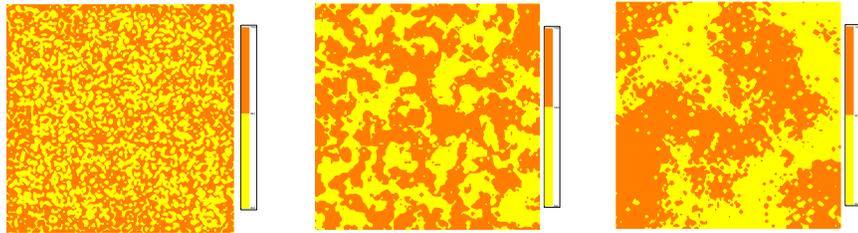
Une nouvelle vision de la Variance

- Exemple de semi-variogramme et clé de lecture



Une nouvelle vision de la Variance

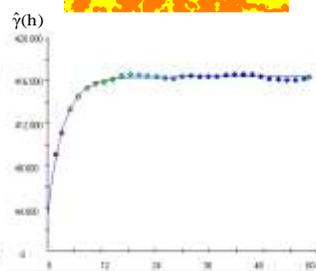
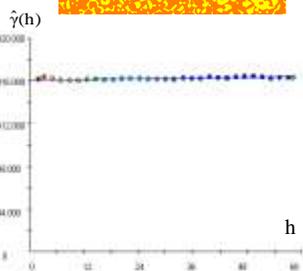
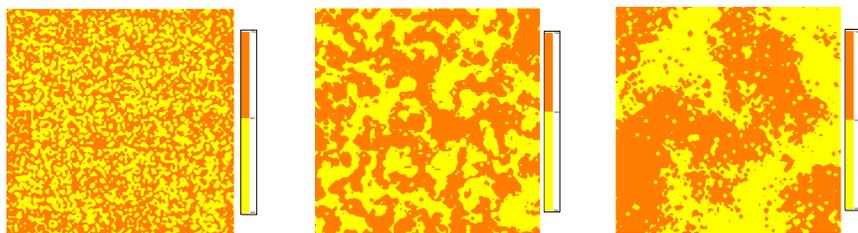
- Exemple de semi-variogramme et clé de lecture



Heliospir Novembre 2013

Une nouvelle vision de la Variance

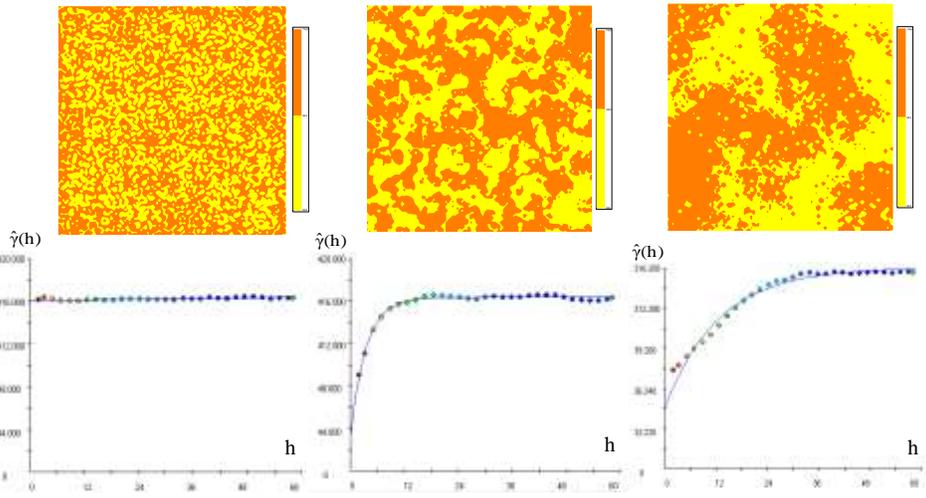
- Exemple de semi-variogramme et clé de lecture



Heliospir Novembre 2013

Une nouvelle vision de la Variance

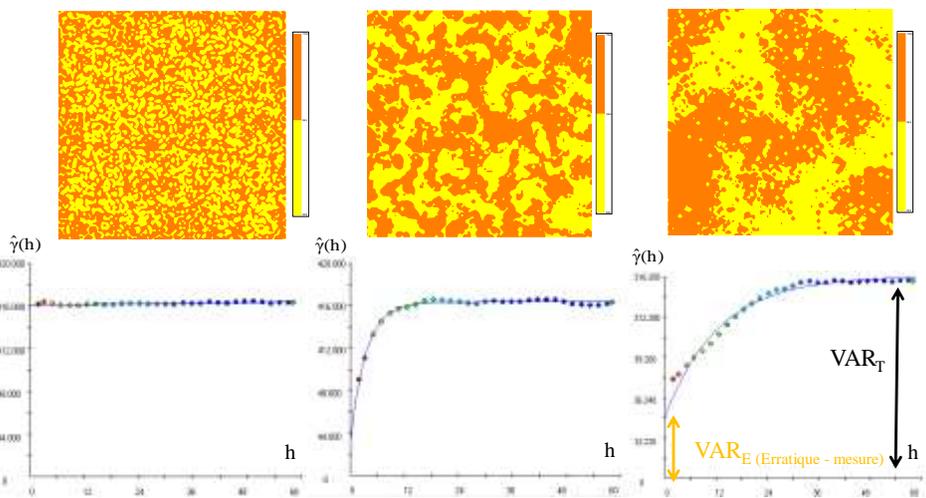
- Exemple de semi-variogramme et clé de lecture



Heliospir Novembre 2013

Une nouvelle vision de la Variance

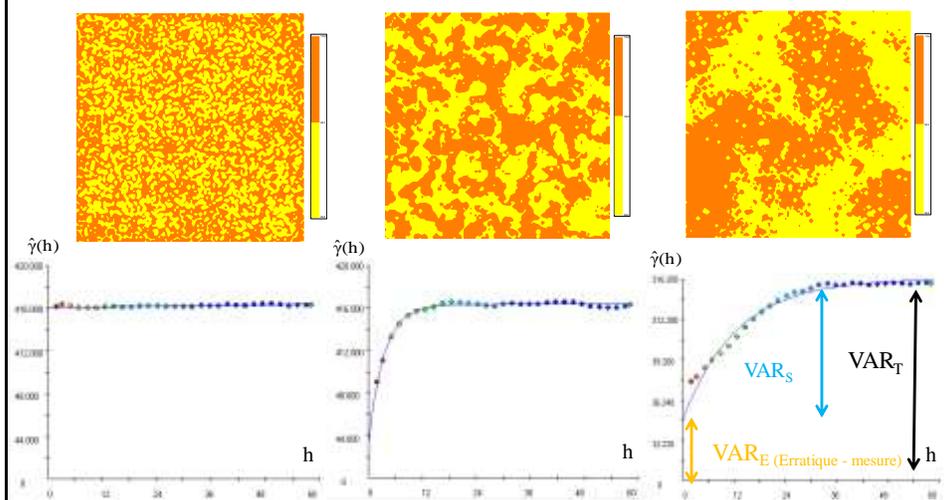
- Exemple de semi-variogramme et clé de lecture



Heliospir Novembre 2013

Une nouvelle vision de la Variance

- Exemple de semi-variogramme et clé de lecture



Heliospir Novembre 2013

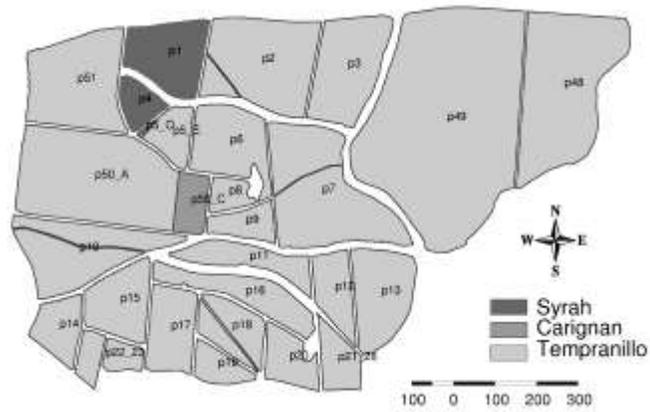
Exemples d'applications :

- Cas 1 : Echantillonnage,
- Cas 2 : Estimation globale,

Heliospir Novembre 2013

Cas 1 : Echantillonnage

Vignoble de 150 ha (32 différentes parcelles).



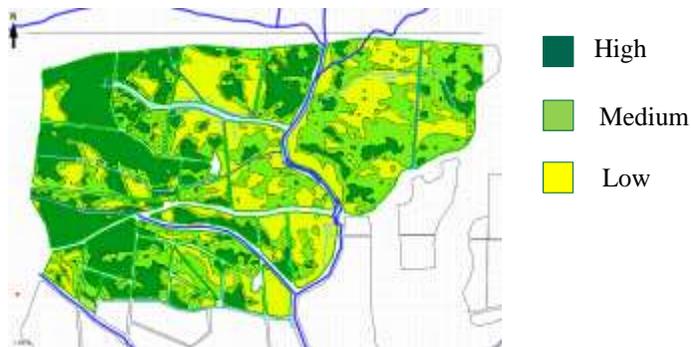
Heliospir Novembre 2013

Cas 1 : Echantillonnage

Image multi-spectrale (NDVI) de 30 cm de résolution.

- Classification (iso-efficatif)

- Identification de grande zones de fonctionnement à valider par des données terrain



NDVI map derived from multispectral image at veraison

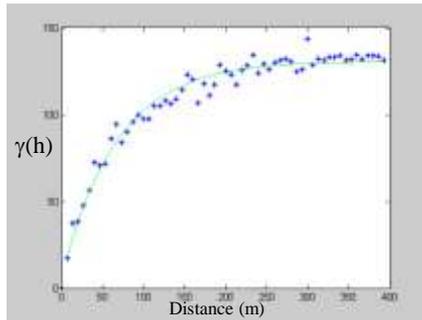
Heliospir Novembre 2013

Santesteban et al., 2013

Cas 1 : Echantillonnage

Définition d'une grille d'échantillonnage

- Contraintes de mesure (temps, matériel, etc.) -> nombre de sites de mesure
- Hypothèse sur la stabilité temporelle des motifs spatiaux observés,
- Une grille d'échantillonnage définie en fonction du semi-variogramme du NDVI,



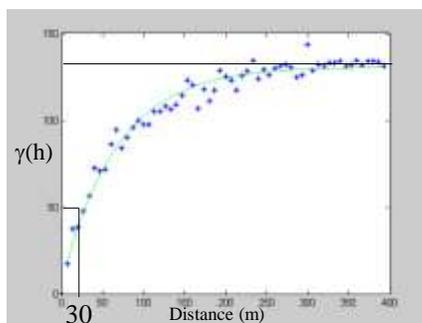
Semi-variance of the NDVI

Heliospir Novembre 2013

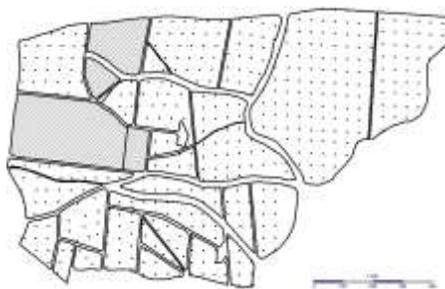
Cas 1 : Echantillonnage

Définition d'une grille d'échantillonnage

- Contraintes de mesure (temps, matériel, etc.) -> nombre de sites de mesure
- Hypothèse sur la stabilité temporelle des motifs spatiaux observés,
- Une grille d'échantillonnage définie en fonction du semi-variogramme du NDVI,
- La distance entre site -> **75 %** de la variance du NDVI,



Semi-variance of the NDVI

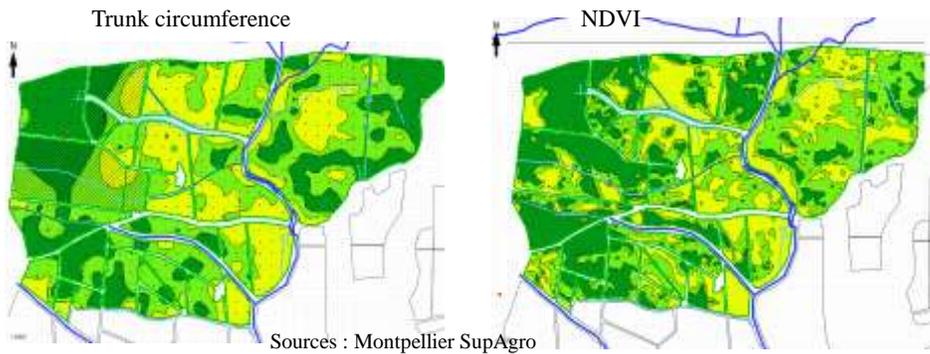


Sampling grid (30 x 30 m), 256 sites

Heliospir Novembre 2013

Cas 1 : Echantillonnage

Définition d'une grille d'échantillonnage

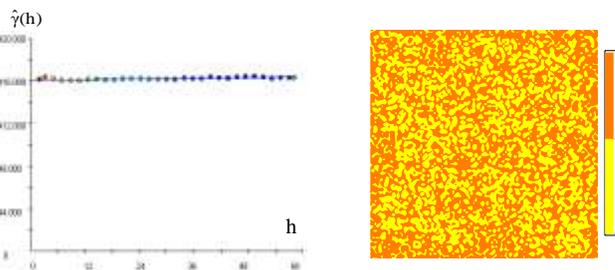


Heliospir Novembre 2013

Cas 2 : Estimation globale

Estimation de la moyenne du rendement avant vendange sur une parcelle

Approche pratique courante

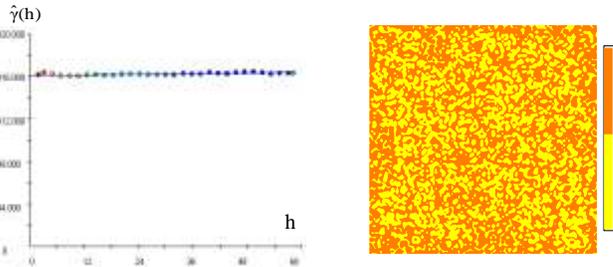


$Z(s) = \{Z(s_1), Z(s_2), \dots, Z(s_i)\}, \{s_1, s_2, \dots, s_i\} \in D$
VA identiquement distribuées et indépendantes

Heliospir Novembre 2013

Cas 2 : Estimation globale

Estimation de la moyenne du rendement avant vendange sur une parcelle
Approche pratique courante

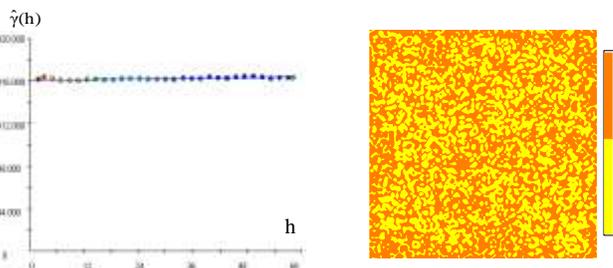


$Z(s) = \{Z(s_1), Z(s_2), \dots, Z(s_i)\}, \{s_1, s_2, \dots, s_i\} \in D$
VA identiquement distribuées et indépendantes

n échantillons $z(s_1), z(s_2), z(s_3), \dots, z(s_n)$
Estimation de l'espérance (\bar{x}) et de la variance (s^2)

Cas 2 : Estimation globale

Estimation de la moyenne du rendement avant vendange sur une parcelle
Approche pratique courante



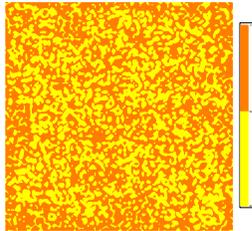
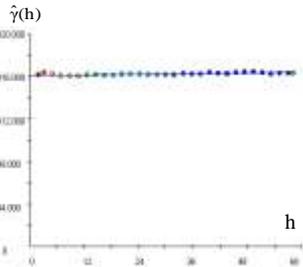
$Z(s) = \{Z(s_1), Z(s_2), \dots, Z(s_i)\}, \{s_1, s_2, \dots, s_i\} \in D$
VA identiquement distribuées et indépendantes

n échantillons $z(s_1), z(s_2), z(s_3), \dots, z(s_n)$
Estimation de l'espérance (\bar{x}) et de la variance (s^2)

$$\text{Intervalle de confiance } I_c = \left[\bar{x} - t_\alpha \frac{s}{\sqrt{n}} ; \bar{x} + t_\alpha \frac{s}{\sqrt{n}} \right]$$

Cas 2 : Estimation globale

Estimation de la moyenne du rendement avant vendange sur une parcelle
Approche pratique courante



Question du positionnement spatial des échantillons d'observations ?

$Z(s) = \{Z(s_1), Z(s_2), \dots, Z(s_i)\}, \{s_1, s_2, \dots, s_i\} \in D$
VA identiquement distribuées et indépendantes

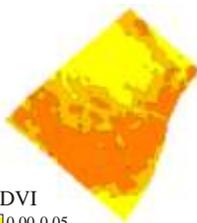
n échantillons $z(s_1), z(s_2), z(s_3), \dots, z(s_n)$

Estimation de l'espérance (\bar{x}) et de la variance (s^2)

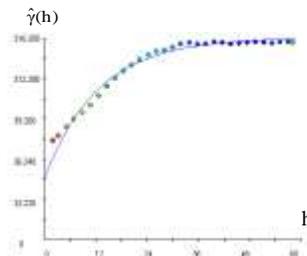
$$\text{Intervalle de confiance } I_c = \left[\bar{x} - t_\alpha \frac{s}{\sqrt{n}} ; \bar{x} + t_\alpha \frac{s}{\sqrt{n}} \right]$$

Cas 2 : Estimation globale

Estimation de la moyenne du rendement avant vendange sur une parcelle
Utilisation d'une donnée auxiliaire à haute résolution spatiale



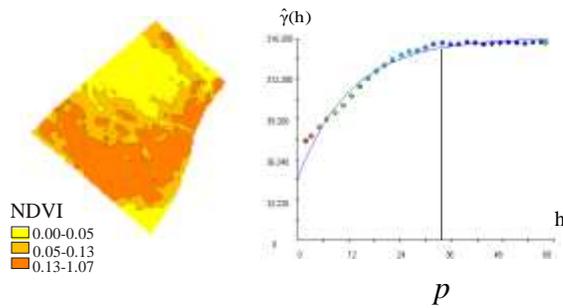
NDVI
 0.00-0.05
 0.05-0.13
 0.13-1.07



$Z(s) = \{Z(s_1), Z(s_2), \dots, Z(s_i)\}, \{s_1, s_2, \dots, s_i\} \in D$
VA identiquement distribuées **et indépendantes**

Cas 2 : Estimation globale

Estimation de la moyenne du rendement avant vendange sur une parcelle
Utilisation d'une donnée auxiliaire à haute résolution spatiale



NDVI
 0.00-0.05
 0.05-0.13
 0.13-1.07

$Z(s) = \{Z(s_1), Z(s_2), \dots, Z(s_i)\}, \{s_1, s_2, \dots, s_i\} \in D$
 VA identiquement distribuées **et indépendantes**

n échantillons $z(s_1), z(s_2), z(s_3), \dots, z(s_n)$ tq $|s_i - s_j| > p$

Conclusion

- Une très brève initiation aux méthodes utilisées pour décrire les données spatialisées dans le domaine du vivant,
- D'autres utilisation du semi-variogramme telle que l'estimation locale/interpolation spatiale : krigage,
- Méthodes stochastiques (géostatistiques) formalisées par Matheron (1965),
- Cadre formelle intéressant (phénomène erratique/phénomène structuré) qui nécessite une certaine expertise sur les données (support spatial de la mesure), des hypothèses relatives à la distribution des données .



Merci pour votre attention