Géostatistique spatio-temporelle

Denis Allard

Biostatistique et Processus Spatiaux (BioSP), MIA, INRA
Avignon, France

Atelier Statistique: Introduction aux méthodes spatiales
23–24 juin 2016, IHP, Paris

Introduction

- On collecte des données dans l'espace (S) et dans le temps (T)
- Il existe des dépendances à la fois dans le temps et dans l'espace

$\hookrightarrow$ A la fois une généralisation des statistiques spatiales et des séries temporelles

Introduction

Quelques configurations possibles

1. Séries denses dans le temps, rares spatialement (capteurs, stations météo, ...)
   $\hookrightarrow$ séries temporelles multivariées.
   Mais quel modèle spatial? comment interpoler en un nouveau site?
2. Échantillonnages denses spatialement, peu nombreux (images satellites, campagnes halieutique)
   $\hookrightarrow$ champs spatiaux multivariés.
   Mais quelle évolution temporelle? quelle prédiction temporelle?
3. Echantillonnages denses dans le temps et dans l'espace
   On va chercher à modéliser l'interaction espace $\times$ temps
4. Cas plus complexes: pas d'alignement temporel, faibles fréquences spatiales et temporelles, trajectoires, ...

Introduction

Quelques configurations possibles

1. Séries denses dans le temps, rares spatialement (capteurs, stations météo, ...)
   $\hookrightarrow$ séries temporelles multivariées.
   Mais quel modèle spatial? comment interpoler en un nouveau site?
2. Échantillonnages denses spatialement, peu nombreux (images satellites, campagnes halieutique)
   $\hookrightarrow$ champs spatiaux multivariés.
   Mais quelle évolution temporelle? quelle prédiction temporelle?
3. Echantillonnages denses dans le temps et dans l'espace
   On va chercher à modéliser l'interaction espace $\times$ temps
4. Cas plus complexes: pas d'alignement temporel, faibles fréquences spatiales et temporelles, trajectoires, ...

INRA SCIENCE & IMPACT département mia Biostatistique BIO/∫Π & Processus Spatiaux

Introduction

Quelques configurations possibles

1. Séries denses dans le temps, rares spatialement (capteurs, stations météo, ...)
   $\hookrightarrow$ séries temporelles multivariées.
   Mais quel modèle spatial? comment interpoler en un nouveau site?
2. Échantillonnages denses spatialement, peu nombreux (images satellites, campagnes halieutique)
   $\hookrightarrow$ champs spatiaux multivariés.
   Mais quelle évolution temporelle? quelle prédiction temporelle?
3. Echantillonnages denses dans le temps et dans l'espace
   On va chercher à modéliser l'interaction espace $\times$ temps
4. Cas plus complexes: pas d'alignement temporel, faibles fréquences spatiales et temporelles, trajectoires, ...

Introduction

Quelques configurations possibles

1. Séries denses dans le temps, rares spatialement (capteurs, stations météo, ...)
   $\hookrightarrow$ séries temporelles multivariées.
   Mais quel modèle spatial? comment interpoler en un nouveau site?
2. Échantillonnages denses spatialement, peu nombreux (images satellites, campagnes halieutique)
   $\hookrightarrow$ champs spatiaux multivariés.
   Mais quelle évolution temporelle? quelle prédiction temporelle?
3. Echantillonnages denses dans le temps et dans l'espace
   On va chercher à modéliser l'interaction espace $\times$ temps
4. Cas plus complexes: pas d'alignement temporel, faibles fréquences spatiales et temporelles, trajectoires, ...

Introduction

Quelques configurations possibles

1. Séries denses dans le temps, rares spatialement (capteurs, stations météo, ...)
   $\hookrightarrow$ séries temporelles multivariées.
   Mais quel modèle spatial? comment interpoler en un nouveau site?
2. Échantillonnages denses spatialement, peu nombreux (images satellites, campagnes halieutique)
   $\hookrightarrow$ champs spatiaux multivariés.
   Mais quelle évolution temporelle? quelle prédiction temporelle?
3. Echantillonnages denses dans le temps et dans l'espace
   On va chercher à modéliser l'interaction espace $\times$ temps
4. Cas plus complexes: pas d'alignement temporel, faibles fréquences spatiales et temporelles, trajectoires, ...

Modèles spatio-temporels

1. On dispose d'un modèle physique descriptif
   - Filtre de Kalman (KF)
   - Ensemble KF

   Dans un contexte spatialisé
2. On construit un modèle statistique

$$Z(\mathbf{s},t) = \mu(\mathbf{s},t) + Y(\mathbf{s},t) + \epsilon(\mathbf{s},t), \quad (\mathbf{s},t) \in D \times T \subset \mathbb{R}^d \times \mathbb{R}$$

   avec
   - $$g[\mu(\mathbf{s},t)] = \sum_{k=1}^{p} \beta_k f_k(X_k(\mathbf{s})) + \sum_{l=1}^{q} \alpha_l f_l(X_l(t))$$
   - $Y(\mathbf{s},t)$ est un champ aléatoire d'ordre 2

$$\text{stationnaire} \Leftrightarrow \mathrm{Cov}\{Y(\mathbf{s}+\mathbf{h},t+u), Y(\mathbf{s},t)\} = C(\mathbf{h},u)$$

   - $\epsilon(\mathbf{s},t)$ bruit aléatoire (pépite)
3. Quelles fonctions de covariance $C(\mathbf{h},u)$?

Modèles spatio-temporels

1. On dispose d'un modèle physique descriptif
   - Filtre de Kalman (KF)
   - Ensemble KF

   Dans un contexte spatialisé
2. On construit un modèle statistique

$$Z(\mathbf{s},t) = \mu(\mathbf{s},t) + Y(\mathbf{s},t) + \epsilon(\mathbf{s},t), \quad (\mathbf{s},t) \in D \times T \subset \mathbb{R}^d \times \mathbb{R}$$

   avec
   - $$g[\mu(\mathbf{s},t)] = \sum_{k=1}^{p} \beta_k f_k(X_k(\mathbf{s})) + \sum_{l=1}^{q} \alpha_l f_l(X_l(t))$$
   - $Y(\mathbf{s},t)$ est un champ aléatoire d'ordre 2

     $$\text{stationnaire} \Leftrightarrow \mathrm{Cov}\{Y(\mathbf{s}+\mathbf{h}, t+u), Y(\mathbf{s},t)\} = C(\mathbf{h},u)$$
   - $\epsilon(\mathbf{s},t)$ bruit aléatoire (pépite)
3. Quelles fonctions de covariance $C(\mathbf{h},u)$?

Modèles spatio-temporels

1. On dispose d'un modèle physique descriptif
   - Filtre de Kalman (KF)
   - Ensemble KF

   Dans un contexte spatialisé
2. On construit un modèle statistique

$$Z(\mathbf{s},t) = \mu(\mathbf{s},t) + Y(\mathbf{s},t) + \epsilon(\mathbf{s},t), \quad (\mathbf{s},t) \in D \times T \subset \mathbb{R}^d \times \mathbb{R}$$

   avec
   - $$g[\mu(\mathbf{s},t)] = \sum_{k=1}^{p} \beta_k f_k(X_k(\mathbf{s})) + \sum_{l=1}^{q} \alpha_l f_l(X_l(t))$$
   - $Y(\mathbf{s},t)$ est un champ aléatoire d'ordre 2

$$\text{stationnaire} \Leftrightarrow \mathrm{Cov}\{Y(\mathbf{s}+\mathbf{h},t+u), Y(\mathbf{s},t)\} = C(\mathbf{h},u)$$

   - $\epsilon(\mathbf{s},t)$ bruit aléatoire (pépite)
3. Quelles fonctions de covariance $C(\mathbf{h},u)$?

Rappels

- Stationnarité

$$\text{Cov}\{Y(\mathbf{s}+\mathbf{h}, t+u), Y(\mathbf{s}, t)\} = C(\mathbf{h}, u)$$

- Séparabilité

$$C(\mathbf{h}, u) = C_S(\mathbf{h}).C_T(u) \Leftarrow Y(\mathbf{s}, t) = Y_S(\mathbf{s}).Y_T(t), \quad Z_S(\mathbf{s}) \perp Z_T(t)$$

Matrice de covariance

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_T \otimes \boldsymbol{\Sigma}_S$$

Propriété

$$Z(\mathbf{s}, t) \perp Z(\mathbf{s}', t') \mid Z(\mathbf{s}, t')$$

- Symétrie complète (Gneiting et al., 2006)

$$C(\mathbf{h}, u) = C(\mathbf{h}, -u) = C(-\mathbf{h}, u) = C(-\mathbf{h}, -u)$$

Séparabilité $\Longrightarrow$ symétrie complète

Rappels

- Stationnarité

$$\mathrm{Cov}\{Y(\mathbf{s}+\mathbf{h}, t+u), Y(\mathbf{s}, t)\} = C(\mathbf{h}, u)$$

- Séparabilité

$$C(\mathbf{h}, u) = C_S(\mathbf{h}).C_T(u) \Leftarrow Y(\mathbf{s}, t) = Y_S(\mathbf{s}).Y_T(t), \quad Z_S(\mathbf{s}) \perp Z_T(t)$$

Matrice de covariance

$$\mathbf{\Sigma} = \mathbf{\Sigma}_T \otimes \mathbf{\Sigma}_S$$

Propriété

$$Z(\mathbf{s}, t) \perp Z(\mathbf{s}', t') \mid Z(\mathbf{s}, t')$$

- Symétrie complète (Gneiting et al., 2006)

$$C(\mathbf{h}, u) = C(\mathbf{h}, -u) = C(-\mathbf{h}, u) = C(-\mathbf{h}, -u)$$

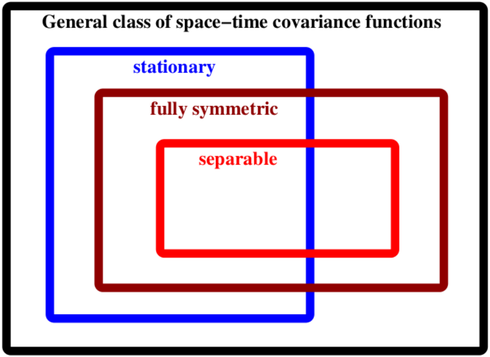
Séparabilité $\Longrightarrow$ symétrie complète

Rappels

- Stationnarité
$$\mathrm{Cov}\{Y(\mathbf{s}+\mathbf{h}, t+u), Y(\mathbf{s}, t)\} = C(\mathbf{h}, u)$$
- Séparabilité
$$C(\mathbf{h}, u) = C_S(\mathbf{h}).C_T(u) \Leftarrow Y(\mathbf{s}, t) = Y_S(\mathbf{s}).Y_T(t), \quad Z_S(\mathbf{s}) \perp Z_T(t)$$
Matrice de covariance
$$\mathbf{\Sigma} = \mathbf{\Sigma}_T \otimes \mathbf{\Sigma}_S$$
Propriété
$$Z(\mathbf{s}, t) \perp Z(\mathbf{s}', t') \mid Z(\mathbf{s}, t')$$
- Symétrie complète (Gneiting et al., 2006)
$$C(\mathbf{h}, u) = C(\mathbf{h}, -u) = C(-\mathbf{h}, u) = C(-\mathbf{h}, -u)$$
Séparabilité $\Longrightarrow$ symétrie complète

INRA SCIENCE & IMPACT — département mia — Biostatistique BIO/Π & Processus Spatiaux

Hello



D'après Gneiting et al. (2006)

Objectif: aller au delà de la séparabilité

Pour aller au-delà: quelques cas simples

- Modèle somme-produit (De Iaco et al., 2013)

$$C(\mathbf{h},u) = a_0 C_S^0(\mathbf{h}).C_T^0(u) + a_1 C_S^1(\mathbf{h}) + a_2 C_T^2(u)$$

  avec $a_0, a_1, a_2 > 0$.
- "Frozen model"

$$C(\mathbf{h},u) = C_S(\mathbf{h} - \mathbf{v}u)$$

  où $\mathbf{v}$ est une vélocité. En EDP, on parle de fronts d'onde (travelling waves)
- Généralisation (Cox and Isham, 1988)

$$C(\mathbf{h},u) = \mathrm{E}[C_S(\mathbf{h} - \mathbf{V}u)]$$

  où $\mathbf{V}$ est un vecteur vitesse aléatoire. Exemple: $\mathbf{V} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{D}^2)$, alors (Schlather, 2010)

$$C(\mathbf{h},u) = \frac{1}{\det\{\mathbf{1} + u^2\mathbf{D}\}^{1/2}} C_S\left\{(\mathbf{h} - u\boldsymbol{\mu})^t(\mathbf{1} + u^2\mathbf{D})^{-1}(\mathbf{h} - u\boldsymbol{\mu})\right\}$$

Modèle avec vélocité dynamique (Ailliot et al., 2011)

déplacement

Soit $\phi(\mathbf{s}, t, t')$ le déplacement du point $\mathbf{s} \in \mathbb{R}^d$ durant les temps t et $t' \in T = [t_0 - \Delta; t_0 + \Delta]$. On pose

- $\phi(\mathbf{s}, t, t) = \mathbf{s}, \quad \forall (\mathbf{s}, t) \in \mathbb{R}^d \times T$
- $\phi(\mathbf{s}, t, t') = \phi(\mathbf{s}, u, t') \circ \phi(\mathbf{s}, t, u), \quad \forall \mathbf{s} \in \mathbb{R}^d$ et $\forall t \leq u \leq s \in T$.
- L'inverse de $\phi(\mathbf{s}, t, t')$ existe

Modèle de champ alétoire

Posons

$$Z(\mathbf{s}, t) = Y(\phi^{-1}(\mathbf{s}, t_0, t), t),$$

où $Y(\cdot, \cdot)$ est un champ aléatoire sur $\mathbb{R}^d \times \mathbb{R}$, avec une fonction de covariance $C(\mathbf{h}, u)$. Alors,

$$\text{Cov}\left\{Z(\mathbf{s}, t), Z(\mathbf{s} + \mathbf{h}, t')\right\} = C\left(\phi^{-1}(\mathbf{s} + \mathbf{h}, t_0, t') - \phi^{-1}(\mathbf{s}, t_0, t), t' - t\right)$$

Modèle avec vélocité dynamique

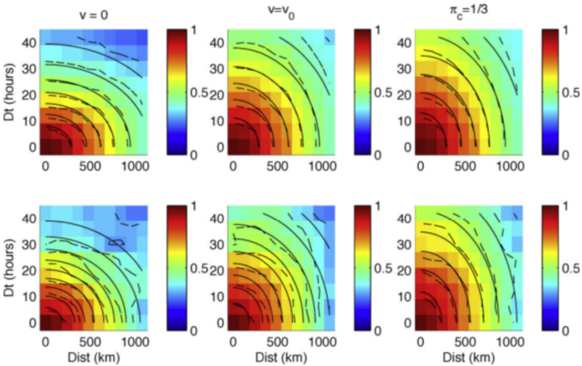


Figure 10. Empirical correlation function of the Lagrangian field computed using ERA-Interim data (top panels) and satellite data (bottom panels). Left panels: no motion ($\mathbf{V}(\mathbf{p}, t) = 0$), middle panels: constant velocity ($\mathbf{V}(\mathbf{p}, t) = \mathbf{V}_0$), right panels: dynamic velocity with dispersion relation ($\pi_c = 1/3$). The dashed lines represent the levels of the empirical covariance function and the full lines the levels of the fitted parametric model. This figure is available in color online at wileyonlinelibrary.com/journal/environmetrics

Ailliot et al. (2011)

Quelques rappels théoriques

Définie positivité

Une fonction $C : \mathbb{R}^d \times \mathbb{R} \mapsto \mathbb{R}$ est une fonction de covariance ssi
$\forall n, \forall \mathbf{s}_1, \ldots, \mathbf{s}_n \in \mathbb{R}^d$, et $\forall a_1, \ldots, a_n \in \mathbb{R}$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_i C(\mathbf{s}_i, t_i; \mathbf{s}_j, t_j) a_j \geq 0$$

C est une fonction (semi-) définie positive.

Stabilité

Les fonctions définies positives sont stables pour le produit, les sommes positives et par la continuité $\Rightarrow$ stables pour les mélanges positifs.

Théorème de Bochner

C est une fonction def. pos. sur $\mathbb{R}^d \times \mathbb{R}$ ssi

$$C(\mathbf{h}, u) = \int\int e^{i\omega^t \mathbf{h} + i\tau u} f(\omega; u) \mathrm{d}\omega \mathrm{d}u, \quad (\mathbf{h}; u) \in \mathbb{R}^d \times \mathbb{R}$$

Quelques rappels théoriques

Définie positivité

Une fonction $C : \mathbb{R}^d \times \mathbb{R} \mapsto \mathbb{R}$ est une fonction de covariance ssi $\forall n, \forall \mathbf{s}_1, \ldots, \mathbf{s}_n \in \mathbb{R}^d$, et $\forall a_1, \ldots, a_n \in \mathbb{R}$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_i C(\mathbf{s}_i, t_i; \mathbf{s}_j, t_j) a_j \geq 0$$

C est une fonction (semi-) définie positive.

Stabilité

Les fonctions définies positives sont stables pour le produit, les sommes positives et par la continuité $\Rightarrow$ stables pour les mélanges positifs.

Théorème de Bochner

C est une fonction def. pos. sur $\mathbb{R}^d \times \mathbb{R}$ ssi

$$C(\mathbf{h}, u) = \int \int e^{i\omega^t \mathbf{h} + i\tau u} f(\omega; u) \mathrm{d}\omega \mathrm{d}u, \quad (\mathbf{h}; u) \in \mathbb{R}^d \times \mathbb{R}$$

Quelques rappels théoriques

Définie positivité

Une fonction $C : \mathbb{R}^d \times \mathbb{R} \mapsto \mathbb{R}$ est une fonction de covariance ssi $\forall n, \forall \mathbf{s}_1, \ldots, \mathbf{s}_n \in \mathbb{R}^d$, et $\forall a_1, \ldots, a_n \in \mathbb{R}$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_i C(\mathbf{s}_i, t_i; \mathbf{s}_j, t_j) a_j \geq 0$$

C est une fonction (semi-) définie positive.

Stabilité

Les fonctions définies positives sont stables pour le produit, les sommes positives et par la continuité $\Rightarrow$ stables pour les mélanges positifs.

Théorème de Bochner

C est une fonction def. pos. sur $\mathbb{R}^d \times \mathbb{R}$ ssi

$$C(\mathbf{h}, u) = \int \int e^{i\boldsymbol{\omega}^t \mathbf{h} + i\tau u} f(\boldsymbol{\omega}; u) \mathrm{d}\boldsymbol{\omega} \mathrm{d}u, \quad (\mathbf{h}; u) \in \mathbb{R}^d \times \mathbb{R}$$

Critère de Cressie and Huang (1999)

Théorème de Cressie-Huang (1999)

Une fonction continue, bornée, symétrique et intégrable $C(\mathbf{h}, u)$ définie sur $\mathbb{R}^d \times \mathbb{R}$ est une fonction de covariance spatio-temporelle ssi

$$C_{\omega}(u) = \int e^{-i\omega^t \mathbf{h}} C(\mathbf{h}, u) \, \mathrm{d}\mathbf{h}, \quad u \in \mathbb{R},$$

est une fonction de covariance (temporelle) pour presque tout $\omega \in \mathbb{R}^d$.

Cette construction nécessite de pouvoir évaluer les transformées de Fourier.

Covariances isotropiques et fonctions complètement monotones

Fonction complètement monotones

Une fonction $\varphi : \mathbb{R}^+ \mapsto \mathbb{R}$ est complètement monotone si, $\forall n$ $\varphi^{(n)}$ existe, et $(-1)^n \varphi^{(n)}(t) \geq 0$. Alors,

$$\varphi(t) = \int_0^\infty \exp(-rt)\,\mathrm{d}F(r) \Leftrightarrow \varphi = \mathcal{L}(F)(t)$$

Théorème Schoenberg, 1938

$C(\mathbf{h}) = \varphi(||\mathbf{h}||^2)$ est def. pos. dans $\mathbb{R}^d, \forall d > 0$ $\Leftrightarrow \varphi(||\mathbf{h}||)$ est complètement monotone.

Mélanges de Gaussiennes $\subset$ fonctions de covariance.

Fonction	Param.
$\varphi(t) = \exp(-ct^\gamma)$	$c > 0, 0 < \gamma \leq 1$
$\varphi(t) = (1 + ct^\gamma)^{-\nu}$	$c > 0, 0 < \gamma \leq 1, \nu > 0$
$\varphi(t) = [2^{\nu-1}\Gamma(\nu)]^{-1}(ct^{1/2})^\nu K_\nu(ct^{1/2})$	$c > 0, \nu > 0$
$\ldots$	

Covariances isotropiques et fonctions complètement monotones

Fonction complètement monotones

Une fonction $\varphi : \mathbb{R}^+ \mapsto \mathbb{R}$ est complètement monotone si, $\forall n\ \varphi^{(n)}$ existe, et $(-1)^n \varphi^{(n)}(t) \geq 0$. Alors,

$$\varphi(t) = \int_0^\infty \exp(-rt)\, \mathrm{d}F(r) \Leftrightarrow \varphi = \mathcal{L}(F)(t)$$

Théorème Schoenberg, 1938

$C(\mathbf{h}) = \varphi(||\mathbf{h}||^2)$ est def. pos. dans $\mathbb{R}^d, \forall d > 0 \quad \Leftrightarrow \varphi(||\mathbf{h}||)$ est complètement monotone.

Mélanges de Gaussiennes $\subset$ fonctions de covariance.

Fonction	Param.
$\varphi(t) = \exp(-ct^\gamma)$	$c > 0, 0 < \gamma \leq 1$
$\varphi(t) = (1 + ct^\gamma)^{-\nu}$	$c > 0, 0 < \gamma \leq 1, \nu > 0$
$\varphi(t) = [2^{\nu-1}\Gamma(\nu)]^{-1}(ct^{1/2})^\nu K_\nu(ct^{1/2})$	$c > 0, \nu > 0$
$\ldots$	

INRA SCIENCE & IMPACT — département mia — Biostatistique BIO/Π & Processus Spatiaux

Covariances isotropiques et fonctions complètement monotones

Fonction complètement monotones

Une fonction $\varphi : \mathbb{R}^+ \mapsto \mathbb{R}$ est complètement monotone si, $\forall n\ \varphi^{(n)}$ existe, et $(-1)^n \varphi^{(n)}(t) \geq 0$. Alors,

$$\varphi(t) = \int_0^\infty \exp(-rt)\, \mathrm{d}F(r) \Leftrightarrow \varphi = \mathcal{L}(F)(t)$$

Théorème Schoenberg, 1938

$C(\mathbf{h}) = \varphi(||\mathbf{h}||^2)$ est def. pos. dans $\mathbb{R}^d, \forall d > 0 \quad \Leftrightarrow \varphi(||\mathbf{h}||)$ est complètement monotone.

Mélanges de Gaussiennes $\subset$ fonctions de covariance.

Fonction	Param.
$\varphi(t) = \exp(-ct^\gamma)$	$c > 0, 0 < \gamma \leq 1$
$\varphi(t) = (1 + ct^\gamma)^{-\nu}$	$c > 0, 0 < \gamma \leq 1, \nu > 0$
$\varphi(t) = [2^{\nu-1}\Gamma(\nu)]^{-1}(ct^{1/2})^\nu K_\nu(ct^{1/2})$	$c > 0, \nu > 0$
...	...

Construction de Gneiting

Théorème Gneiting (2002)

Soit $\varphi(t)$ une fonction complètement monotone et soit $\psi(t) = c + \gamma(t)$, avec $c > 0$ et $\gamma(t)$ un variogramme temporel. Alors

$$C(\mathbf{h}, u) = \frac{\sigma^2}{\psi(u^2)^{1/2}} \varphi\left(\frac{||\mathbf{h}||^2}{\psi(u^2)}\right), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}$$

est une fonction de covariance.

Preuve

1. Montrer que

$$C(\mathbf{h}, u) = \exp(-au^2) \frac{\sigma^2}{\psi(u^2)^{1/2}} \varphi\left(\frac{||\mathbf{h}||^2}{\psi(u^2)}\right)$$

   est définie positive.
2. Faire tendre $a \to 0$

Construction de Gneiting

Théorème Gneiting (2002)

Soit $\varphi(t)$ une fonction complètement monotone et soit $\psi(t) = c + \gamma(t)$, avec $c > 0$ et $\gamma(t)$ un variogramme temporel. Alors

$$C(\mathbf{h}, u) = \frac{\sigma^2}{\psi(u^2)^{1/2}} \varphi\left(\frac{||\mathbf{h}||^2}{\psi(u^2)}\right), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}$$

est une fonction de covariance.

Preuve

1. Montrer que

$$C(\mathbf{h}, u) = \exp(-au^2) \frac{\sigma^2}{\psi(u^2)^{1/2}} \varphi\left(\frac{||\mathbf{h}||^2}{\psi(u^2)}\right)$$

   est définie positive.
2. Faire tendre $a \to 0$

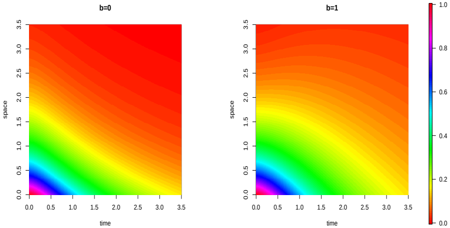
Exemple

Gneiting-Matérn

$$\phi(\mathbf{h}) = \mathcal{M}(\mathbf{h}; r, \nu); \qquad \psi(t) = (\alpha t^a + 1)^b, \quad t \geq 0,$$

avec $\alpha, r, \nu > 0, 0 < a \leq 1, 0 \leq b \leq 1$.

$$C(\mathbf{h}, u) = \frac{\sigma^2}{(\alpha|u|^{2a} + 1)^{b/2}} \mathcal{M}\left(\frac{\mathbf{h}}{(\alpha|u|^{2a} + 1)^b}; r, \nu,\right), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R},$$



Encore plus général

Schlather (2010)

Soit $G : \mathbb{R}^{2d} \mapsto \mathbb{R}^{p\times p}$ une matrice de variogramme croisé sur $\mathbb{R}^d$;
Soit $M \in \mathbb{R}^{p\times p}$ une matrice définie positive, t.q. $M + G(\mathbf{s}, \mathbf{s}')$ est s.d.p $\forall \mathbf{s}, \mathbf{s}' \in \mathbb{R}^d$;
Soit $H : \mathbb{R}^d \mapsto \mathbb{R}^p$:

$$C(\mathbf{s}, \mathbf{s}') = \frac{\varphi\{[(H(\mathbf{s}) - H(\mathbf{s}'))^t (M + G(\mathbf{s}, \mathbf{s}'))^{-1} (H(\mathbf{s}) - H(\mathbf{s}'))]^{1/2}\}}{\sqrt{\det(M + G(\mathbf{s}, \mathbf{s}'))}}$$

- Si $p = d$, $H(\mathbf{s}) = \mathbf{s}$, $M = m.\mathbf{I}_d$ et $G((\mathbf{s}, t), (\mathbf{s}', t')) = \gamma(t' - t)\mathbf{I}_d$

$$\Rightarrow C(\mathbf{h}, \tau) = \frac{\sigma^2}{[m + \gamma(\tau)]^{d/2}} \varphi\left(\frac{||h||^2}{m + \gamma(\tau)}\right), h \in \mathbb{R}^d, \tau \in \mathbb{R}$$

  construction de Gneiting (2002).
- Si $H(\mathbf{s}) = \mathbf{s}$, $M = 0$ et $G(\mathbf{s}, \mathbf{s}') = G^*(\mathbf{s}) + G^*(\mathbf{s}')$
  $\Rightarrow$ construction de Stein (2005).
- Si $M = \mathbf{I}_{d+1}$, $H(\mathbf{s}, t) = \mathbf{s} + \mu t$ et $G((\mathbf{s}, t); (\mathbf{s}', t')) = (t - t')^2 \mathbf{D}$
  $\Rightarrow$ construction de Cox and Isham (1988).

Encore plus général

Schlather (2010)

Soit $G : \mathbb{R}^{2d} \mapsto \mathbb{R}^{p\times p}$ une matrice de variogramme croisé sur $\mathbb{R}^d$;
Soit $M \in \mathbb{R}^{p\times p}$ une matrice définie positive, t.q. $M + G(\mathbf{s}, \mathbf{s}')$ est s.d.p $\forall \mathbf{s}, \mathbf{s}' \in \mathbb{R}^d$;
Soit $H : \mathbb{R}^d \mapsto \mathbb{R}^p$:

$$C(\mathbf{s}, \mathbf{s}') = \frac{\varphi\{[(H(\mathbf{s}) - H(\mathbf{s}'))^t (M + G(\mathbf{s}, \mathbf{s}'))^{-1} (H(\mathbf{s}) - H(\mathbf{s}'))]^{1/2}\}}{\sqrt{\det(M + G(\mathbf{s}, \mathbf{s}'))}}$$

- Si $p = d$, $H(\mathbf{s}) = \mathbf{s}$, $M = m.\mathbf{I}_d$ et $G((\mathbf{s}, t), (\mathbf{s}', t')) = \gamma(t' - t)\mathbf{I}_d$

$$\Rightarrow C(\mathbf{h}, \tau) = \frac{\sigma^2}{[m + \gamma(\tau)]^{d/2}} \varphi\left(\frac{||h||^2}{m + \gamma(\tau)}\right), h \in \mathbb{R}^d, \tau \in \mathbb{R}$$

construction de Gneiting (2002).

- Si $H(\mathbf{s}) = \mathbf{s}$, $M = 0$ et $G(\mathbf{s}, \mathbf{s}') = G^*(\mathbf{s}) + G^*(\mathbf{s}')$
$\Rightarrow$ construction de Stein (2005).
- Si $M = \mathbf{I}_{d+1}$, $H(\mathbf{s}, t) = \mathbf{s} + \mu t$ et $G((\mathbf{s}, t); (\mathbf{s}', t')) = (t - t')^2\mathbf{D}$
$\Rightarrow$ construction de Cox and Isham (1988).

Encore plus général

Schlather (2010)

Soit $G : \mathbb{R}^{2d} \mapsto \mathbb{R}^{p \times p}$ une matrice de variogramme croisé sur $\mathbb{R}^d$;
Soit $M \in \mathbb{R}^{p \times p}$ une matrice définie positive, t.q. $M + G(\mathbf{s}, \mathbf{s}')$ est s.d.p $\forall \mathbf{s}, \mathbf{s}' \in \mathbb{R}^d$;
Soit $H : \mathbb{R}^d \mapsto \mathbb{R}^p$:

$$C(\mathbf{s}, \mathbf{s}') = \frac{\varphi\{[(H(\mathbf{s}) - H(\mathbf{s}'))^t (M + G(\mathbf{s}, \mathbf{s}'))^{-1} (H(\mathbf{s}) - H(\mathbf{s}'))]^{1/2}\}}{\sqrt{\det(M + G(\mathbf{s}, \mathbf{s}'))}}$$

- Si $p = d$, $H(\mathbf{s}) = \mathbf{s}$, $M = m.\mathbf{I}_d$ et $G((\mathbf{s}, t), (\mathbf{s}', t')) = \gamma(t' - t)\mathbf{I}_d$

  $$\Rightarrow C(\mathbf{h}, \tau) = \frac{\sigma^2}{[m + \gamma(\tau)]^{d/2}} \varphi\left(\frac{||h||^2}{m + \gamma(\tau)}\right), h \in \mathbb{R}^d, \tau \in \mathbb{R}$$

  construction de Gneiting (2002).
- Si $H(\mathbf{s}) = \mathbf{s}$, $M = 0$ et $G(\mathbf{s}, \mathbf{s}') = G^*(\mathbf{s}) + G^*(\mathbf{s}')$
  $\Rightarrow$ construction de Stein (2005).
- Si $M = \mathbf{I}_{d+1}$, $H(\mathbf{s}, t) = \mathbf{s} + \mu t$ et $G((\mathbf{s}, t); (\mathbf{s}', t')) = (t - t')^2 \mathbf{D}$
  $\Rightarrow$ construction de Cox and Isham (1988).

Encore plus général

Schlather (2010)

Soit $G : \mathbb{R}^{2d} \mapsto \mathbb{R}^{p \times p}$ une matrice de variogramme croisé sur $\mathbb{R}^d$;
Soit $M \in \mathbb{R}^{p \times p}$ une matrice définie positive, t.q. $M + G(\mathbf{s}, \mathbf{s}')$ est s.d.p $\forall \mathbf{s}, \mathbf{s}' \in \mathbb{R}^d$;
Soit $H : \mathbb{R}^d \mapsto \mathbb{R}^p$:

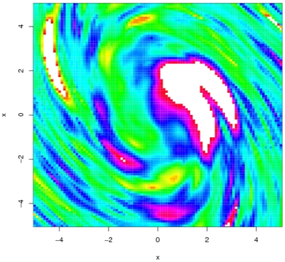
$$C(\mathbf{s}, \mathbf{s}') = \frac{\varphi\{[(H(\mathbf{s}) - H(\mathbf{s}'))^t (M + G(\mathbf{s}, \mathbf{s}'))^{-1} (H(\mathbf{s}) - H(\mathbf{s}'))]^{1/2}\}}{\sqrt{\det(M + G(\mathbf{s}, \mathbf{s}'))}}$$

- Si $p = d$, $H(\mathbf{s}) = \mathbf{s}$, $M = m.\mathbf{I}_d$ et $G((\mathbf{s}, t), (\mathbf{s}', t')) = \gamma(t' - t)\mathbf{I}_d$

  $$\Rightarrow C(\mathbf{h}, \tau) = \frac{\sigma^2}{[m + \gamma(\tau)]^{d/2}} \varphi\left(\frac{||h||^2}{m + \gamma(\tau)}\right), h \in \mathbb{R}^d, \tau \in \mathbb{R}$$

  construction de Gneiting (2002).
- Si $H(\mathbf{s}) = \mathbf{s}$, $M = 0$ et $G(\mathbf{s}, \mathbf{s}') = G^*(\mathbf{s}) + G^*(\mathbf{s}')$
  $\Rightarrow$ construction de Stein (2005).
- Si $M = \mathbf{I}_{d+1}$, $H(\mathbf{s}, t) = \mathbf{s} + \mu t$ et $G((\mathbf{s}, t); (\mathbf{s}', t')) = (t - t')^2 \mathbf{D}$
  $\Rightarrow$ construction de Cox and Isham (1988).

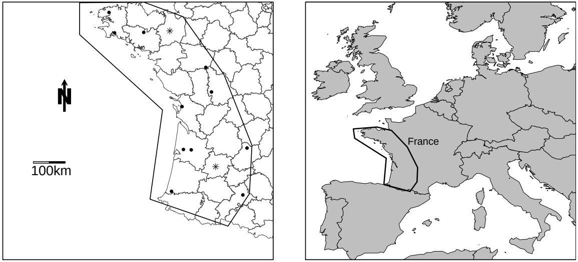
Illustration "Cyclone"



©Martin Schlather. Simulé grâce au package `RandomFields`.

Modèle multivarié spatio-temporel

Contexte



Emplacement de 13 stations Météo-France. Etoiles: stations de validation

Retour sur la séparabilité

Séparabilité complète

$$\mathbf{C}(\mathbf{k}) = \mathbf{A}\, \rho_S(\mathbf{h})\, \rho_T(u)$$

Alors,

$$\boldsymbol{\Sigma} = \mathbf{A} \otimes \mathbf{C}_S \otimes \mathbf{C}_T,$$

Séparabilité spatio-temporelle

$$\rho_{ST}(\mathbf{k}) = \rho_S(\mathbf{h})\, \rho_T(u)$$

$$\Longleftrightarrow$$

$$\mathrm{Cov}\left\{Z(\mathbf{s},t), Z(\mathbf{s}',t')\right\} \propto \mathrm{Cov}\left\{Z(\mathbf{s},t), Z(\mathbf{s}',t)\right\} \mathrm{Cov}\left\{Z(\mathbf{s}',t), Z(\mathbf{s}',t')\right\},$$

- Les interactions complexes ne sont pas modélisées
- En contexte multivarié, cela implique que chaque variable est caractérisée par la même fonction de corrélation spatio-temporelle.

Modèle multivarié spatio-temporel

Idée générale

Combinerla non séparabilité spatio-temporelle du modèle de Gneiting avec la non séparabilité du modèle multivarié de Matérn flexible.

Résultat principal Bourotte et al. (2016)

Le modèle multivarié de Gneiting-Matérn noté $\mathbf{C}^{\mathcal{M}} = \left[C_{ij}^{\mathcal{M}}(\cdot,\cdot)\right]_{i,j=1}^{p}$, avec

$$C_{ij}^{\mathcal{M}}(\mathbf{h},u) = \frac{\sigma_i\sigma_j}{\psi(u^2)^{d/2}}\rho_{ij}\mathcal{M}\left(\frac{\mathbf{h}}{\psi(u^2)^{1/2}}; r_{ij},\nu_{ij}\right), \quad (\mathbf{h},u) \in \mathbb{R}^d \times \mathbb{R}, \qquad (1)$$

est une matrice de fonctions de covariances valides si, pour tout $i, j = 1, \ldots, p$,

$$\begin{aligned} r_{ij} &= (r_i^2/2 + r_j^2/2)^{1/2}; \qquad \nu_{ij} = (\nu_i + \nu_j)/2, \\ \rho_{ij} &= \beta_{ij}\frac{\Gamma(\nu_{ij})}{\Gamma(\nu_i)^{1/2}\Gamma(\nu_j)^{1/2}}\frac{r_{ij}^{\nu_{ij}}}{(r_i^{\nu_i}r_j^{\nu_j})^{1/2}}, \end{aligned}$$

avec $r_i, \nu_i > 0$ et $0 < \lambda \leq 2$, et où $\boldsymbol{\beta} = \left[\beta_{ij}\right]_{i,j=1}^{p}$ est une matrice de corrélation, et $\psi(t), t \geq 0$, est une fonction positive qui peut s'écrire $c + \gamma(t)$, où $c > 0$ et $\gamma(t)$ est un variogramme.

Modèle particulier

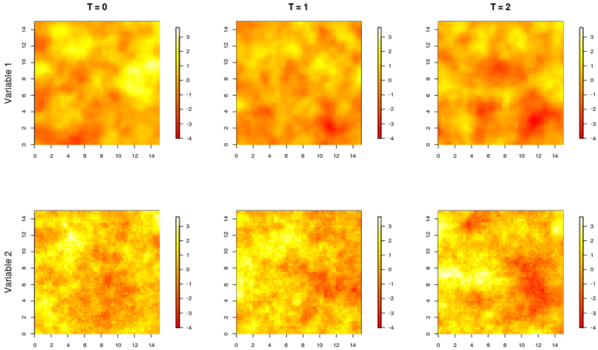
Gneiting-Matérn

$$\psi(x) = (\alpha x^a + 1)^b, \quad x \geq 0,$$

avec $\alpha > 0, 0 < a \leq 1, 0 \leq b \leq 1$.

$$C_{ij}(\mathbf{h}, u) = \frac{\sigma_i \sigma_j}{(\alpha|u|^{2a} + 1)^{b/2}} \rho_{ij} \mathcal{M}\left(\frac{\mathbf{h}}{(\alpha|u|^{2a} + 1)^b}; r_{ij}, \nu_{ij},\right), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R},$$

Simulation



Champ Gaussien Gneiting-Matérn spatio-temporel bivariés à 3 temps consécutifs. En haut: variable régulière ($\nu_1 = 1.5$ and $r_1 = 1$); en bas: variable moins régulière ($\nu_2 = 0.5$ and $r_2 = 0.5$). la corrélation est ρ_{12} is set to 0.5.

Estimation des paramètres

Problème

- le maximum de vraisemblance nécessite $\mathcal{O}\left((np)^3\right)$ opérations (déterminant et inverse de la matrice de covariance $np \times np$)
- Il y a $(p+2)(p+3)/2$ paramètres à estimater, i.e. 15 paramètres lorsque $p = 3$

Solution: utiliser la vraisemblance composite

- produits de vraisemblances plus petites
- aisée à calculer

Cf. Varin et al. (2011) pour une revue

Pairwise Composite Likelihood

- Pairwise marginal Gaussian log-likelihoods computed on pairs of data (Bevilacqua et al., 2014; Bevilacqua and Gaetan, 2015)

$$\ell(i, j, \mathbf{s}_\alpha, \mathbf{s}_\beta, t_\alpha, t_\beta; \theta).$$

- The Weighted Pairwise log-Likelihood (WPL) is thus

$$\mathrm{wpl}(\theta) = \sum_{(i,j,\alpha,\beta)\in\Lambda} \ell(i, j, \mathbf{s}_\alpha, \mathbf{s}_\beta, t_\alpha, t_\beta; \theta) w_{\alpha\beta}, \quad \theta \in \Theta,$$

- The computational cost is $\mathcal{O}\left((np)^2\right)$.
- Can be reduced by only considering pairs such that

$$(\|\mathbf{h}\|, u) \leq (d_S, d_T).$$

Here, $d_S = 500$ km and $d_T = 2$ days

Relative efficiency of estimation

Simulation setting

- Mimicking the conditions of weather data
- Three variables, 13 same locations. Estimation using 11 stations, validation on the others
- 30 days, 13 stations, 3 variables = 1170 values.
- 100 repetitions
- Do FL, WCL and prediction on 2 | 11 stations

For each parameter $i = 1, \ldots, 15$ compute

$$\mathrm{bs}_i = (\bar{\hat{\theta}}_i - \theta_i); \quad \mathrm{sd}_i^2 = \sum_{j=1}^{100} (\hat{\theta}_{j,i} - \bar{\hat{\theta}}_i)^2/100; \quad \mathrm{rmse}_i = \sqrt{\mathrm{bs}_i^2 + \mathrm{sd}_i^2}$$

Then,

$$\mathrm{rre}_i = \mathrm{rmse}_i^{FL}/\mathrm{rmse}_i^{WPL},$$

is the Root Relative Efficiency of WPL compared to FL.

- When rre < 1, FL is more efficient that WPL
- Conversely when rre > 1

Relative efficiency of estimation: results

	σ_1		σ_2		σ_3		β_{12}		β_{13}		β_{23}	
	WPL	FL	WPL	FL	WPL	FL	WPL	FL	WPL	FL	WPL	FL
bs^2 (10^{-5})	0.18	5.13	1.13	2.54	3.48	17.1	1.02	5.77	6.30	4.50	7.96	5.35
sd^2 (10^{-3})	1.73	0.84	1.40	0.59	1.24	0.63	8.47	0.26	7.86	0.29	10.8	0.36
rmse	0.042	0.030	0.038	0.025	0.036	0.028	0.092	0.018	0.089	0.018	0.104	0.020
rre	0.72		0.66		0.79		0.19		0.21		0.19	
	ν_1		ν_2		ν_3		α		a		b	
	WPL	FL	WPL	FL	WPL	FL	WPL	FL	WPL	FL	WPL	FL
bs^2 (10^{-3})	2.63	3.48	1.75	6.30	0.60	2.95	0.132	2.82	0.001	0.018	1.000	27.0
sd^2 (10^{-3})	9.71	4.30	13.1	8.95	3.81	1.82	11.4	4.70	5.08	0.699	52.7	19.4
rmse	0.111	0.088	0.122	0.123	0.066	0.069	0.107	0.087	0.071	0.027	0.232	0.215
rre	0.79		1.01		1.04		0.81		0.38		0.93	
	$1/r_1$		$1/r_2$		$1/r_3$							
	WPL	FL	WPL	FL	WPL	FL						
bs^2	218.5	911.6	61.8	473.9	516.6	4539.3						
sd^2	2300.0	856.1	1127.1	682.4	9629	6788.9						
rmse	50.2	42.0	34.5	34.0	100.7	106.4						
rre	0.84		0.99		1.06							

Validation on prediction

We compare space-time separable (S) and non-separable (NS) models, and models with equal (E) or different (D) smoothness and scale parameters.

1. S-E: $b = 0$; all components have equal smoothness and scale parameters.

$$\mathbf{\Sigma} = \mathbf{A} \otimes \mathbf{C}_S \otimes \mathbf{C}_T$$

2. NS-E: $0 < b \leq 1$; all components have equal smoothness and scale parameters. This model is space-time non-separable, but it has the same space-time covariance for all variables.

$$\mathbf{\Sigma} = \mathbf{A} \otimes \mathbf{C}_{ST}$$

3. S-D: $b = 0$; each component has its own smoothness and scale parameters, but it is space-time separable.

$$\mathbf{\Sigma} = \mathbf{B} \otimes \mathbf{C}_T$$

   where B is non separable spatial multivariate covariance.

4. NS-D: $0 < b \leq 1$; each component has its own smoothness and scale parameters, and it is not space-time separable. It is the new full non-separable model

Validation scores

$$\text{MSE} = \frac{1}{6|T|} \sum_{t \in T} \sum_{j=1}^{6} (\tilde{z}_j^t - z_j^t)^2.$$

$$\text{CRPS} = \frac{1}{6|T|} \sum_{t \in T} \sum_{j=1}^{6} \text{crps}(\Phi_j^t, y_j^t).$$

$$\text{with } \text{crps}(F, z) = \int_{-\infty}^{\infty} [F(u) - H(u - z)]^2 du = \tilde{\sigma}_j^t \{\tilde{z}[2\Phi(\tilde{z}) - 1] + 2\phi(\tilde{z}) - \pi^{-1/2}\}$$

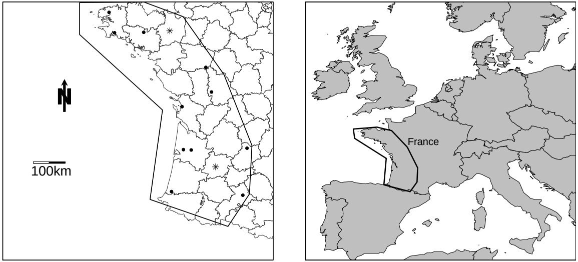
$$\text{LogS} = \frac{1}{6|T|} \sum_{t \in T} \sum_{j=1}^{6} \left\{ \ln \tilde{\sigma}_j + \frac{(\tilde{z}_j^t - z_j^t)^2}{2\tilde{\sigma}_j^2} \right\}.$$

Validation results

	Sim. Data: WPL			Sim. Data: FL		
	RMSE	CRPS	LogS	RMSE	CRPS	LogS
S-E	0.494	0.278	4.070	0.489	0.274	3.792
NS-E	0.488	0.274	3.903	0.488	0.273	3.782
S-D	0.490	0.277	4.091	0.484	0.271	3.733
NS-D	**0.485**	**0.271**	**3.789**	**0.484**	**0.271**	**3.723**

Mean of the prediction scores RMSE, CRPS, LogS, according to models S-E, NS-E, S-D and NS-D. For simulated data, with Weighted pariwise Likelihood (WPL) and Full Likelihood (FL) estimates. For data set, WPL is used.

Western France weather dataset

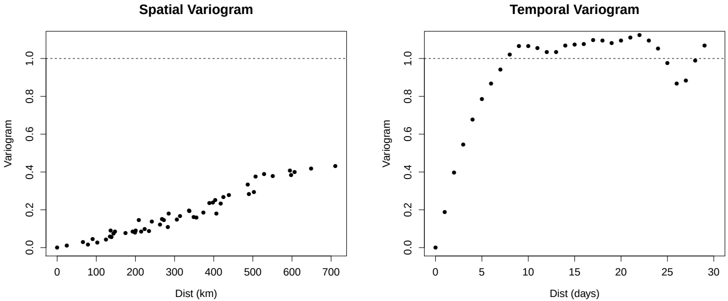


Location of the 13 weather stations over western France. With stars: validation stations

Pre-processing

- Three variables: solar radiation, R, temperature, T, and humidity, H.
- Space-time interaction is expected due to prevailing westerly winds
- Standardization

$$Y_i(\mathbf{s},t) = \mu_i(\mathbf{s}) + \sigma_i(\mathbf{s}) Z_i(\mathbf{s},t) \quad i \in \{\mathrm{R},\mathrm{T},\mathrm{H}\} \quad (\mathbf{s},t) \in \mathbb{R}^2 \times \mathbb{R}.$$



Marginal spatial and temporal variograms for daily January temperature standardized variables.

The full model

Need for a purely temporal process

$$Z_i(\mathbf{s}, t) = X_i(t) + W_i(\mathbf{s}, t) \quad i \in \{\mathrm{R}, \mathrm{T}, \mathrm{H}\}, \quad (\mathbf{s}, t) \in \mathbb{R}^2 \times \mathbb{R}.$$

Hence,

$$\begin{aligned} \gamma_{ii}^{(\mathbf{Z})}(\mathbf{h}, 0) &= \gamma_{ii}^{(\mathbf{X})}(0) + \gamma_{ii}^{(\mathbf{W})}(\mathbf{h}, 0) \underset{\mathbf{h}\to\infty}{\longrightarrow} \left[\sigma_i^{(\mathbf{W})}\right]^2 \\ \gamma_{ii}^{(\mathbf{Z})}(\mathbf{0}, u) &= \gamma_{ii}^{(\mathbf{X})}(u) + \gamma_{ii}^{(\mathbf{W})}(\mathbf{0}, u) \underset{u\to\infty}{\longrightarrow} \left[\sigma_i^{(\mathbf{X})}\right]^2 + \left[\sigma_i^{(\mathbf{W})}\right]^2 = 1. \end{aligned}$$

We further suppose that **W** follows a Gneiting-Matérn covariance function and that the temporal process **X** has covariance function

$$C_{ij}^{(\mathbf{X})}(u) = \sigma_i^{(\mathbf{X})} \sigma_j^{(\mathbf{X})} \beta_{ij}^{(\mathbf{X})} / (\alpha^{(\mathbf{X})} |u|^{2a^{(\mathbf{X})}} + 1),$$

with $i, j \in \{\mathrm{R}, \mathrm{T}, \mathrm{H}\}$ and $u \in \mathbb{R}$.

Model fitting: spatial

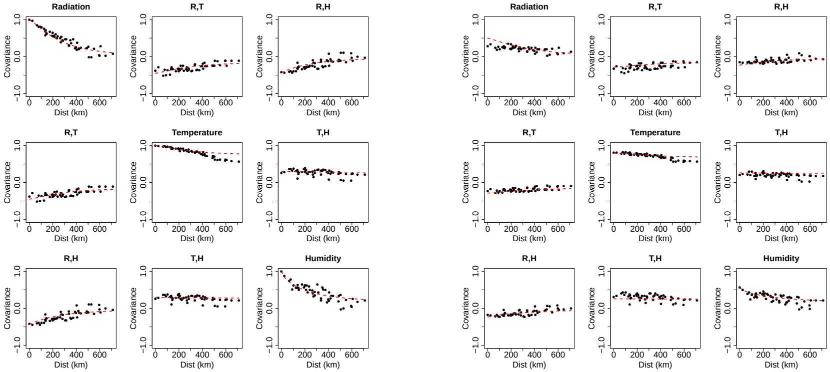


Figure: Spatial direct and cross-covariance functions for R, T and H. Points: empirical values. Solid line: model NS-D with parameters estimated using WPL. Left panel: $u = 0$. Right panel: $u = 1$.

Model fitting: temporal

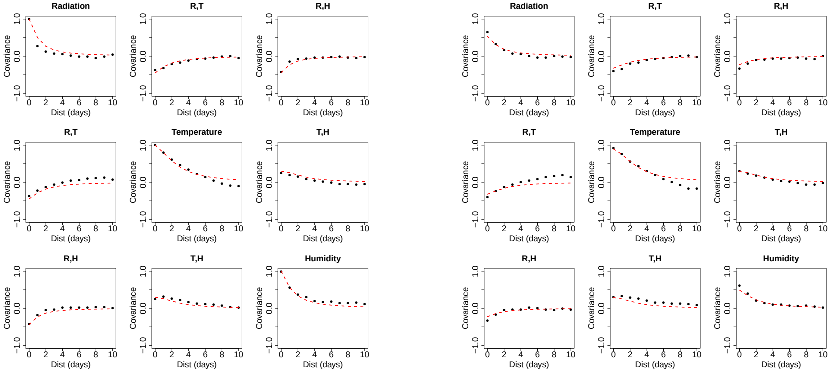


Figure: Temporal direct and cross-covariance functions for R, T and H. Left panel: $\mathbf{h} = 0$. Right panel: $\|\mathbf{h}\| = 199$ km.

Validation with model NS-D

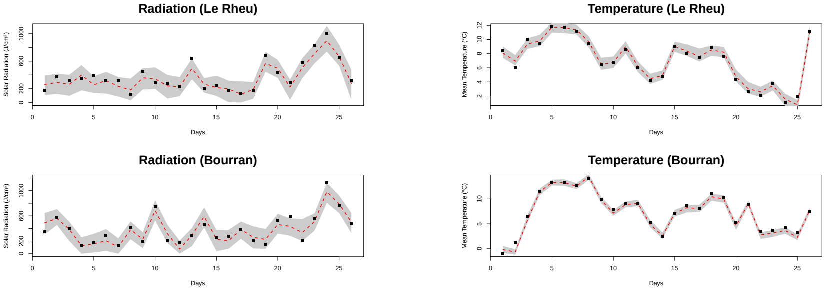


Figure: Prediction of Radiation, Temperature and Humidity at the two validation stations, Le Rheu (Brittany) and Bourran (Aquitaine). Black points: true values. Dotted lines: conditional expectation. Shaded area: 90% envelope interval, computed on 100 simulations.

Validation and interpretation

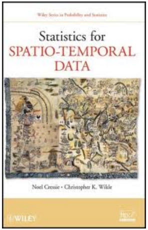
	RMSE	CRPS	LogS
S-E	0.433	0.222	0.440
NS-E	0.446	0.228	0.467
S-D	0.417	**0.215**	0.405
NS-D	**0.417**	0.216	**0.397**

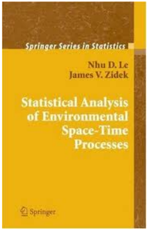
Prediction scores RMSE, CRPS, LogS, according to models S-E, NS-E, S-D and NS-D. Estimation with WPL.

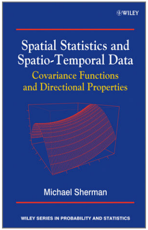
	σ_{RT}	σ_{RH}	σ_{TH}
Process **X**	-0.16	-0.05	0.28
Process **W**	-0.29	-0.39	0.03
Process **Z**	-0.45	-0.44	0.31
Empirical	-0.38	-0.42	0.25

Empirical and estimated covariance coefficients $\sigma_{ij} = \sigma_i \sigma_j \rho_{ij}$, with $i, j \in \{R, T, H\}$ for the model NS-D.

Pour aller plus loin







Packages

Aller lire la Task View "Handling and Analyzing Spatio-Temporal Data" (E. Pebesma)

- `spacetime`: pour typer et gérer des données spatio-temporelles. Créer des types de données différents. Mais aussi `SpatioTemporal`, `stpp`, `adehabitatLT`
- Pour les données géostatististiques, la plupart des packages spatiaux ont des extensions spatio-temporelles: `gstat`, `Randomfields`, mais aussi `spTimer`, `Stem`, ...
- Processus de points: `splancs`, `stpp`, `ptproc`, ...
- Données sur lattice: `surveillance`, `splm`, ...
- Données de trajectoire: `trip`, `diveMove`, `move`, `animalTrack`, ...

RESSTE

RESSTE
Réseau MIA RESSTE

RESeau Statistiques pour données Spatio-TEmporelles

Bienvenue | Lancement: Oct 2014 | Estimation: Mai 2015 | Estimation & Visualisation: Nov 2015 | Atelier d'Avignon, Avril 2016

User login

Username *

Password *

- Create new account
- Request new password

Log in

Bienvenue

Les progrès de l'instrumentation, des systèmes électroniques embarqués et de l'imagerie satellite génèrent de très grandes quantités de données, géolocalisées et répétées dans le temps. Un grand nombre de variables, parfois assez fortement dépendantes, sont disponibles. Les défis posés par ces données spatio-temporelles sont multiples :ils concernent à la fois la visualisation de ces données, la définition de modèles pertinents, les méthodes d'estimation pour ceux-ci et la mise en œuvre de ces méthodes pour des ensembles de données de (très) grande taille.

Le réseau RESSTE (**RES**eau **S**tatistiques pour données **S**patio-**TE**mporelles) propose une animation scientifique autour des modèles, méthodes et algorithmes pour les données spatio-temporelles, en fédérant des statisticiens de différents horizons et des modélisateurs des autres disciplines ayant à traiter de ce type de données, en confrontant les approches et points de vue des différents domaines des statistiques. Il a pour ambition de contribuer à une plus large diffusion des méthodes en statistiques spatio-temporelles, auprès des statisticiens et des modélisateurs. A moyen terme, il s'agit d'initier de nouveaux fronts de recherche en statistiques spatio-temporelles répondant aux enjeux rencontrés par les scientifiques, et de favoriser les collaborations entre équipes de recherche.

RESSTE est financé depuis 2014 par le département MIA de l'INRA. Il compte à ce jour 63 membres dans 19 équipes (voir la liste complète ici) et est animé par un bureau de six personne:

Denis Allard, BioSP, MIA

Liliane Bel, UMR MIA, AgroParisTech

Edith Gabriel, Université d'Avignon

Cécile Hardouin, Université Paris Ouest

Thomas Opitz, BioSP

Eric Parent, UMR MIA, AgroParisTech

La communauté statistiques pour l'environnement du portail math info de l'INRA:

http://informatique-mia.inra.fr/resste/

INRA SCIENCE & IMPACT   département mia   Biostatistique BIO/Π & Processus Spatiaux

Bibliography

Ailliot, P., Baxevani, A., Cuzol, A., Monbet, V., & Raillard, N. (2011). Space?time models for moving fields with an application to significant wave height fields. *Environmetrics*, 22(3), 354-369.

Apanasovich, T. V. and Genton, M. G. (2010). Cross-covariance functions for multivariate random fields based on latent dimensions. *Biometrika*, 97(1):15–30.

Apanasovich, T. V., Genton, M. G., and Sun, Y. (2012). A valid Matérn class of cross-covariance functions for multivariate random fields with any number of components. *Journal of the American Statistical Association*, 107(497):180–193.

Bevilacqua M., Gaetan C., Mateu J., Porcu E. (2014) Estimating Space and Space-Time Covariance Functions for Large Data Sets: A Weighted Composite Likelihood Approach. *Journal of the American Statistical Association*, 107:497, 268-280.

Bevilacqua, M. and Gaetan, C. (2015). Comparing composite likelihood methods based on pairs for spatial Gaussian random fields. *Statistics and Computing*, 25(5):877–892.

Bourotte, M., Allard, D., and Porcu, E. (2016). A flexible class of non-separable cross-covariance functions for multivariate space–time data. *Spatial Statistics*,

Chilès J.P. and Delfiner P., 2nd Edition (2012). *Geostatistics: Modeling Spatial Uncertainty*. John Wiley & Sons, New-York.

Cressie, N., and Huang, H. C. (1999). Classes of nonseparable, spatio-temporal stationary covariance functions. *Journal of the American Statistical Association*, 94(448), 1330-1339.

Cox, D. R., and Isham, V. (1988). A simple spatial-temporal model of rainfall. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 415, pp. 317-328

De Iaco, S., Myers, D., Palma, M., and Posa (2013). Using simultaneous diagonalization to identify a space-time linear coregionalization model. *Mathematical Geosciences*, 45(1):69–86.

Gneiting, T. (2002). Nonseparable, stationary covariance functions for space?time data. *Journal of the American Statistical Association*, 97(458), 590-600.

Gneiting, T., Genton, M., and Guttorp, P. (2006). Geostatistical space-time models, stationarity, separability and full symmetry. In Fintenstädt, B., Held, L., and Isham, V., editors, *Statistical Methods for Spatio-Temporal Systems*, pages 151–175. Chapman & Hall/CRC.

Gneiting, T., Kleiber, W., and Schlather, M. (2010). Matérn cross-covariance functions for multivariate random fields. *Journal of the American Statistical Association*, 105(491):1167–1177.

Schlather, M. (2010). Some covariance models based on normal scale mixtures. *Bernoulli*, 16(3):780–797.

Varin, C., Reid, N., and Firth, D. (2011). An overview of composite likelihood methods. *Statistica Sinica*, 21(1):5–42.