Géostatistique spatio-temporelle (2)

Denis Allard

Biostatistique et Processus Spatiaux (BioSP), MathNum, INRAE
Avignon, France

Option Géostatistique et Probabilité Appliquées, Mines-ParisTech
17-18 Janvier 2023

Plan

- Jour 1
  1. Introduction, premières notions et premier TP
  2. Modèles de covariance spatio-temporelle, avec exercice
  3. Présentation: modélisation de la précipitation à fine échelle

- Jour 2
  1. Estimation et TP estimation (ou J1, selon avancement)
  2. Prédiction (et TP prédiction)
  3. Méthode de simulation
  4. Présentation: extension(s) au cadre multivarié

Quiz

Quiz

European Cultural Centre
TIME
SPACE
EXISTENCE
FUTURE PRACTICE, PRACTICE FUTURE
BIENNALE ARCHITETTURA 2016
PALAZZO ROSSINI

Estimation: plusieurs approches possibles

Paramètres $\theta = (\mu, \sigma^2, a_S, a_T, ...)$

- Ajustement variogramme: variogramme empirique + Moindres Carrés Pondérés (`gstat, Rgeostats`)

$$\hat{\theta} = \arg\min_{\theta} \sum_{k,j} w_{kj} \left(\gamma(h_k, u_j; \theta) - \hat{\gamma}(h_k, u_j)\right)^2$$

- Maximum de vraisemblance :

$$\hat{\theta}_\ell = \arg\max_{\theta} \ell(\theta; \mathbf{z})$$

avec $\ell(\theta; \mathbf{z}) = g_n(\mathbf{z}; \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta)$. Rarement utilisable en contexte ST.

- Maximum de Vraisemblance composite, e.g. vraisemblance de paires = Pairwise Likelihood (PL) (`Rgeostats, GeoModels`)

Estimation: plusieurs approches possibles

Paramètres $\theta = (\mu, \sigma^2, a_S, a_T, ...)$

- Ajustement variogramme: variogramme empirique + Moindres Carrés Pondérés (`gstat, Rgeostats`)

$$\hat{\theta} = \arg\min_{\theta} \sum_{k,j} w_{kj} \left(\gamma(h_k, u_j; \theta) - \hat{\gamma}(h_k, u_j)\right)^2$$

- Maximum de vraisemblance :

$$\hat{\theta}_\ell = \arg\max_{\theta} \ell(\theta; \mathbf{z})$$

avec $\ell(\theta; \mathbf{z}) = g_n(\mathbf{z}; \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta)$. Rarement utilisable en contexte ST.

- Maximum de Vraisemblance composite, e.g. vraisemblance de paires = Pairwise Likelihood (PL) (`Rgeostats, GeoModels`)

Estimation: plusieurs approches possibles

Paramètres $\theta = (\mu, \sigma^2, a_S, a_T, ...)$

- Ajustement variogramme: variogramme empirique + Moindres Carrés Pondérés (`gstat, Rgeostats`)

$$\hat{\theta} = \arg\min_{\theta} \sum_{k,j} w_{kj} \left(\gamma(h_k, u_j; \theta) - \hat{\gamma}(h_k, u_j)\right)^2$$

- Maximum de vraisemblance :

$$\hat{\theta}_\ell = \arg\max_{\theta} \ell(\theta; \mathbf{z})$$

avec $\ell(\theta; \mathbf{z}) = g_n(\mathbf{z}; \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta)$. Rarement utilisable en contexte ST.

- Maximum de Vraisemblance composite, e.g. vraisemblance de paires = Pairwise Likelihood (PL) (`Rgeostats, GeoModels`)

Rappels sur le Maximum de Vraisemblance

Soit $(Y_1, \ldots, Y_n)$ un échantillon i.i.d. d'une variable aléatoire continue ayant pour densité f_θ, $\theta \in \Theta$
La vraisemblance de $(Y_1, \ldots, Y_n)$ est:

$$l(Y_1, \ldots, Y_n; \theta) = \prod_{i=1}^{n} f_\theta(Y_i)$$

Il est souvent plus commode de travailler avec la log-vraisemblance:

$$L(Y_1, \ldots, Y_n; \theta) = \log l(Y_1, \ldots, Y_n; \theta) = \sum_{i=1}^{n} \log(f_\theta(Y_i))$$

L'estimateur du maximum de vraisemblance (EMV) de θ est

$$\widehat{\theta} = \arg\max_{\theta \in \Theta} l(Y_1, \ldots, Y_n; \theta) = \arg\max_{\theta \in \Theta} L(Y_1, \ldots, Y_n; \theta)$$

Exercice

Soit $(Y_1, \ldots, Y_n)$ un échantillon i.i.d. d'une variable aléatoire $\mathcal{N}(\mu, \sigma^2)$

1. Ecrire la log-vraisemblance
2. Calculer l'EMV de μ
3. compute l'EMV de σ^2

Propriétés asymptotiques de l'EMV

- Consistance (i.e. $\widehat{\theta}_n \xrightarrow{\mathcal{P}} \theta$) lorsque $n \to \infty$
- Normalité asymptotique
- Efficacité (atteint la limite inférieure de Cramer-Rao)

EMV pour un champ Gaussien

Soit $Y(\mathbf{s}, t)$ une champ Gaussien spatio-temporel et $\mathbf{Y} = (Y(\mathbf{s}_1, t_1), \ldots, Y(\mathbf{s}_n, t_n))$ caractérisé par une fonction de covariance $C_\theta(\mathbf{h}, u)$.
La vraisemblance de $\mathbf{Y}$ est alors:

$$l(\theta; \mathbf{Y}) = (2\pi)^{-n/2} |C_\theta|^{-1/2} \exp\left(-\frac{1}{2}\|\mathbf{Y} - \mu\mathbf{1}\|^2_{C_\theta^{-1}}\right)$$

où $\|A\|^2_B = A^t B A$

L'EMV de μ et θ sont

$$(\widehat{\mu}_{ML}, \widehat{\theta}_{ML}) = \arg \max_{\mu \in \mathbb{R}, \theta \in \Theta} l(\theta; \mathbf{Y}) = \arg \max_{\mu \in \mathbb{R}, \theta \in \Theta} L(\theta; \mathbf{Y})$$

où

$$L(\theta; \mathbf{Y}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log|C_\theta| - \frac{1}{2}\|\mathbf{Y} - \mu\mathbf{1}\|^2_{C_\theta^{-1}}$$

INRAE — département MathNum — Biostatistique BIOSP & Processus Spatiaux

Recherche du MV

- Optimisation non-linéaire sous contrainte (sur Θ)
- Grid search
- Descente de gradient

$$\theta^{(k+1)} = \theta^{(k)} + M^{(k)} g^{(k)}$$

avec
  - $M^{(k)}$ is a $m \times m$ matrix
  - $g^{(k)} = \frac{\partial L(\theta;\mathbf{Z})}{\partial \theta}|_{\theta=\theta^{(k)}}$
- 2 exemples
  - Newton-Raphson: $M^{(k)} = (B^{(k)})^{-1}$, $B_{ij}^{(k)} = -\frac{\partial^2 L(\theta;\mathbf{Z})}{\partial \theta_i \partial \theta_j}|_{\theta=\theta^{(k)}}$
  - Fisher scoring: $M^{(k)} = (B^{(k)})^{-1}$, $B_{ij}^{(k)} = E\left[-\frac{\partial^2 L(\theta;\mathbf{Z})}{\partial \theta_i \partial \theta_j}|_{\theta=\theta^{(k)}}\right]$

Problèmes computationels

Lorsque n est grand, l'évaluation de L devient très lourde :

$$L(\theta; \mathbf{Y}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log|C_\theta| - \frac{1}{2}\|\mathbf{Y} - \mu\mathbf{1}\|^2_{C_\theta^{-1}}$$

- stockage en $O(n^2)$
- calculs en $O(n^3)$

Avantages et inconvénients de l'EMV

Avantages

- Automatique; pas de paramétrisation potentiellement sensible (lags, distances, etc.)
- Bonnes propriétés statistiques des estimateurs sous hypothèse Gaussienne

Inconvénients

- Repose sur une hypothèse Gaussienne
- Optimisation délicate si les paramètres sont nombreux; dépend des valeurs initiales
- Calculs assez lourds

Vraisemblance de paires (PL)

PL: Définition

$$\mathrm{PL}(\theta; \mathbf{z}) = \prod_{i,j} f_{\theta;i,j}(z_i, z_j)$$

où $f_{\theta;i,j}(z_i, z_j)$ est l'une des densités liées à

$$(Y(s_i,t_i), Y(s_j,t_j)); \qquad Y(s_i,t_i) \mid Y(s_j,t_j); \qquad Y(s_i,t_i) - Y(s_j,t_j)$$

- $$\hat{\theta}_{\mathrm{PL}} = \arg\max_{\theta} \mathrm{PL}(\theta; \mathbf{z})$$
- Coût de calcul réduit en n'utilisant que les paires "proches" en espace et en temps (et réduit aussi parfois le biais)
- Bonnes propriétés (consistance, non biais), mais $\mathrm{Var}(\hat{\theta}_{\mathrm{PL}}) \leq \mathrm{Var}(\hat{\theta}_{\ell})$
- Attention ! Intervalles de confiance, tests d'hypothèse

INRAE — département MathNum — Biostatistique BIOSP & Processus Spatiaux

Comparaison PL / FL sur simulations

Simulations qui miment les conditions des données climatiques en France

- Trois variables, 13 stations, 30 jours = 1170 valeurs, simulés par Cholesky
- 100 répétitions
- FL, PL

Pour chaque paramètre p, on calcule

$$\text{bs}_p = (\bar{\hat{\theta}}_p - \theta_p); \quad \text{sd}_p^2 = \sum_{j=1}^{100} (\hat{\theta}_{j,p} - \bar{\hat{\theta}}_p)^2/100; \quad \text{rmse}_p = \sqrt{\text{bs}_p^2 + \text{sd}_p^2}$$

Puis l'Efficacité Relative de FL comparé à PL:

$$\text{rre}_p = \text{rmse}_p^{FL}/\text{rmse}_p^{PL},$$

- rre > 0.65, sur la plupart des paramètres

Comparaison PL / FL sur simulations

Trois champs spatio-temporels avec même covariance temporelle et différentes covariances spatiales

	$\sigma_1 = 1$		$\sigma_2 = 1$		$\sigma_3 = 1$		$1/r_1 = 250$		$1/r_2 = 200$		$1/r_3 = 350$	
	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	218.5	911.6	61.8	473.9	516.6	4539.3
sd^2 (10^{-3})	1.73	0.84	1.40	0.59	1.24	0.63	2300.0	856.1	1127.1	682.4	9629	6788.9
rmse	0.042	0.030	0.038	0.025	0.036	0.028	50.2	42.0	34.5	34.0	100.7	106.4
rre	0.72		0.66		0.79		0.84		0.99		1.06	
	$\nu_1 = 0.7$		$\nu_2 = 0.8$		$\nu_3 = 9.4$		$\alpha = 0.9$		$a = 0.5$		$b = 0.8$	
	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	

Autres approches possibles

- Approximation aux plus proches voisins (Vecchia); pas encore disponible dans le cadre ST
- Approximations de rang faible (assez mauvais en estimation)
- Tapering (idem)
- . . .
- SPDE

Sélection de modèle

Encore plus crucial que dans le cas spatial, p.ex. : non-séparabilité (positive, nulle, négative); vélocité; déplacement; ...

Plusieurs approches

- comparaison visuelle
- Critère AIC, CLIC (Composite Likelihood Information Criterion)
- validation croisée (cf plus tard)

Messages principaux

- Il existe un grand nombre de façons de passer d'un modèle S $\rightarrow$ ST; plusieurs peuvent être acceptables
- Estimer correctement la dépendance ST est cruciale pour la qualité de la prédiction
- De plus en plus de choix dans les différents package `R` – Malheureusement, pas encore `RGeostats` :-((
  A ce jour, `GeoModels` est le package le plus complet
- L'ajustement visuel des fonctions de covariance (via leur variogrammes souvent) est une tâche difficile $\Rightarrow$ méthodes automatiques d'estimation des paramètres
- Ajustement semi-automatique par optimisation d'un critère (MCP, PL, ...): vérifier la qualité de l'optimisation; initialiser l'optimisateur à des valeurs différentes;
- Toujours vérifier que le modèle estimé est cohérent avec les données

TP - séquence 2

- Ajuster les covariances (ou les variogrammes) spatio-temporels
- Représenter des variogrammes théoriques spatio-temporels
- Utilisation de `GeoModels`
  - `GeoVariogram`
  - `GeoFit`
  - `GeoCV`
  - `GeoKrig`

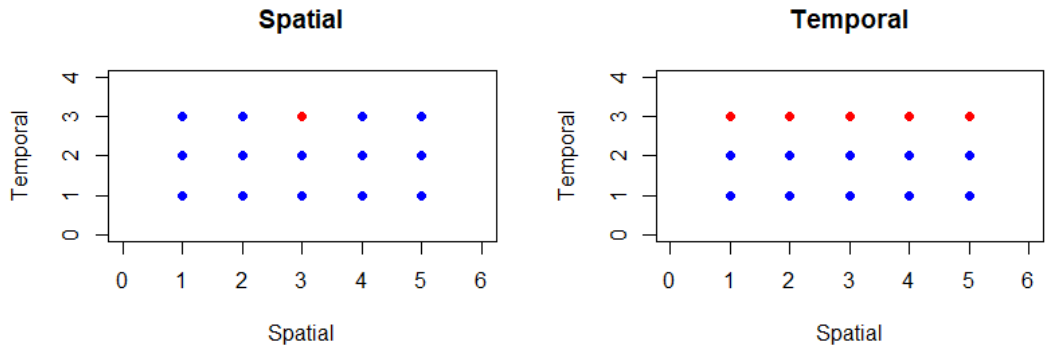
Rappel sur le krigeage

On suppose qu'on connaît la fonction de covariance $C(\mathbf{h}, u)$

Krigeage = Best Linear Unbiased Predictor

$$Z^*(s_0, t_0) = \sum_{(s_\alpha, t_k)} \lambda_{\alpha,k} Z(s_\alpha, t_k)$$

avec $(s_0, t_0) \neq (s_\alpha, t_k),\ \forall s_\alpha, \forall t_k$



Rappel sur le krigeage

On suppose qu'on connaît la fonction de covariance $C(\mathbf{h}, u)$

Krigeage = Best Linear Unbiased Predictor

$$Z^*(s_0, t_0) = \sum_{(s_\alpha, t_k)} \lambda_{\alpha,k} Z(s_\alpha, t_k)$$

On note: $\mathbf{C} = \text{Var}(\mathbf{Z})$ et $\mathbf{C}_0 = \text{Cov}(Z(s_0, t_0), \mathbf{Z})$

Non biais et $\text{Var}(Z^*(s_0, t_0) - Z(s_0, t_0))$ minimum:

- Krigeage simple (espérance nulle)

$$Z^*(s_0, t_0) = \mathbf{C}_0^\top \mathbf{C}^{-1} \mathbf{Z}$$

- Krigeage avec dérives $f_1(s, t), \ldots, f_p(s, t)$

$$Z^*(s_0, t_0) = \left(\mathbf{C}_0 + \mathbf{F}(\mathbf{F}^\top \mathbf{C}^{-1} \mathbf{F})^{-1}(\mathbf{f}_0 - \mathbf{F}^\top \mathbf{C}^{-1} \mathbf{C}_0)\right)^\top \mathbf{C}^{-1} \mathbf{Z}$$

INRAE — département MathNum — Biostatistique BIOSP & Processus Spatiaux

Complexité krigeage

Krigeage simple

$$Z^*(s_0, t_0) = \mathbf{c}_0^\top \mathbf{C}^{-1} \mathbf{Z}$$

Lorsque $n_S \times n_T$ est grand (ça arrive vite !), inverser $\mathbf{C}$ devient impossible

- Voisinages glissants: taille? règles de définition? Voisinages petits pour que cela soit utile $\Rightarrow$ discontinuités
- "tapering" = on remplace $C(\mathbf{h}, u)$ par $C(\mathbf{h}, u) \times C_0(\mathbf{h}, u)$ où $C_0(\mathbf{h}, u)$ est une fonction de covariance à support compact puis utilisation de méthodes d'inversion utilisant les matrices creuses
- Modèles "low rank" = on remplace $\mathbf{C}$ par une matrice de rang faible, m, fixé par l'utilisateur.
- SPDE

Validation et scores de prédiction

Validation

- Règle générale: estimer des données retirées du jeu analysé puis comparer à la vraie valeur
- Validation externe: diviser le jeu de données en deux parties (pas forcément 50/50). Selon quel schéma ? Supprimer des dates / des séries ???
- Validation croisée: supprimer une donnée / une série

Scores

- Erreur quadratique

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|Z_i^* - Z_i|; \quad \text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Z_i^* - Z_i)^2; \quad \text{NMSE} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{Z_i^* - Z_i}{\sigma_i^*}\right)^2$$

Comparaison PL/FL sur simulations

Même modèles que ceux utilisés pour la comparaison en termes d'estimation des paramètres

	PL		FL	
	RMSE	MAE	MSE	MAE
Sep	0.490	0.389	0.484	0.383
Non-sep	0.485	0.383	0.484	0.383

TP - séquence 3

- Faire le krigeage pour l'une des variables climatiques
- Faire une validation croisée
- Utilisation de `GeoModels`
  - `GeoVariogram`
  - `GeoFit`
  - `GeoCV`
  - `GeoKrig`

Simuler un champ ST de la classe de Gneiting

Classe de Gneiting

$$C(\mathbf{h}, u) = \frac{1}{(\gamma(u)+1)^{d/2}} \varphi\left(\frac{|\mathbf{h}|^2}{\gamma(u)+1}\right)$$

où φ est une fonction complètement monotone et γ est un variogramme non borné. On pose: $\varphi(\mathbf{0}) = 1$ et $\psi(\mathbf{0}) = 1$

- Cholesky si $n < 10\,000$
- Méthodes spectrales ? Bandes tournantes ?

Allard, D. Emery X., Lacaux, C., Lantuéjoul, C. (2020) Simulating space-time random fields with nonseparable Gneiting-type covariance functions. *Statistics and Computing*, 30(5): 1479-1495

Classe de Gneiting

Classe de Gneiting

$$C(\mathbf{h}, \mathbf{u}) = \frac{1}{(\gamma(\mathbf{u}) + 1)^{d/2}} \varphi\left(\frac{|\mathbf{h}|^2}{\gamma(\mathbf{u}) + 1}\right)$$

Composante spatiale

Sous hypothèse d'isotropie

$$C_S(\mathbf{h}) = \varphi(||\mathbf{h}||^2) = \int_{\mathbb{R}^+} e^{-r||\mathbf{h}||^2} \mu(dr) := \int_{\mathbb{R}^+} C_G(\mathbf{h}; r)\mu(dr)$$

où $C_G(\mathbf{h}; r)$ est une fonction de covariance Gaussienne de paramètre d'échelle $1/r$

Composante temporelle

$$C_T(u) = (\gamma(u) + 1)^{-d/2}$$

- Pour que $C_T(u) \to 0$ il faut et il suffit que $\gamma(u) \to \infty$

Pour résumer

- Classe de Gneiting

$$C(\mathbf{h}, u) = \frac{1}{(\gamma(u)+1)^{d/2}} \varphi\left(\frac{|\mathbf{h}|^2}{\gamma(u)+1}\right)$$

- Mélange de Gaussienne pour la composante **spatiale**
- Variograme non borné pour la composante **temporelle**

- **Pas d'expression analytique connue pour la Transformée de Fourier ST**

Simulations des champs spatiaux ou spatio-temporels

Méthodes pour simuler en des points pré-determinés

- Cholesky: limité à N petit ou modéré
- Simulation séquentielle: universelle, mais lente et approximative
- Circular embedding matrices: grille régulière, effets de bords et limité à $N = 2^K$

Méthodes pour simuler en tout point de $\mathbb{R}^d \times \mathbb{R}$

- **Bandes Spectrales**

Approche spectrale

Méthode spectrale (basique)

Simuler un champ Gaussien $Z(\mathbf{x})$ ayant pour covariance $C(\mathbf{h})$, et mesure spectrale associée F

1. Choisir p grand (p.ex. 5000)
2. Tirer $\mathbf{\Omega}_j \sim F$ et $\Phi_j \sim \mathcal{U}_{|}(0, 2\pi)$, avec $j = 1, \ldots, p$
3. En tout point $\mathbf{x}$, calculer

$$\tilde{Z}(\mathbf{x}) = \sqrt{\frac{2}{p}} \sum_{j=1}^{p} \cos\left(\langle \mathbf{\Omega}_j, \mathbf{x} \rangle + \Phi_j\right),$$

Convergence to a Gaussian Random Field is achieved as $p \to \infty$ thanks to the Central Limit Theorem

Approche spectrale générique pour une covariance ST

Idée générale

Pour $(\mathbf{x}, t) \in \mathbb{R}^d \times \mathbb{R}$,

$$\tilde{Z}(\mathbf{x}, t) = \sum_{j=1}^{p} \sqrt{\frac{-2\ln(U_j)}{p}} \cos\left(\langle \mathbf{\Omega}_j^S, \mathbf{x} \rangle + \Omega_j^T t + \Phi_j\right)$$

- $(\mathbf{\Omega}_j^S, \Omega_j^T) \sim F$, est la mesure spectrale jointe
- $U_j \sim \mathcal{U}(0,1)$ sert à obtenir une marginale Gaussienne, et donc à accélérer la convergence du TCL
- Même si $F(\omega, \tau)$ n'est pas connue explicitement, il est possible de simuler le vecteur de fréquence spatio-temporelle $(\mathbf{\Omega}^S, \Omega^T)$

Approche spectrale pour la classe de Gneiting

La fonction φ étant une fonction c.m., d'après le théorème de Schoenberg:

$$\varphi(x) = \int_0^\infty e^{-rx^2} \mu(dr)$$

Covariance de la classe de Gneiting

$$C(\mathbf{h}, u) = \int_0^\infty C(\mathbf{h}, u \mid r)\mu(dr)$$

avec

$$C(\mathbf{h}, u \mid r) = \frac{1}{(\gamma(\mathbf{u}) + 1)^{d/2}} \exp\left(-\frac{r|\mathbf{h}|^2}{\gamma(\mathbf{u}) + 1}\right)$$

Mesure spectrale associée

$$F(\boldsymbol{\omega}, \tau) = \int_0^\infty F(\boldsymbol{\omega}, \tau \mid r)\mu(dr)$$

avec

$$F(\boldsymbol{\omega}, \tau \mid r) = F_S(\boldsymbol{\omega} \mid r) F_T(\tau \mid \boldsymbol{\omega}, r)$$

Approche spectrale générique pour une covariance ST

$F_S(\omega \mid r)$

En posant $u = 0$, on a

$$C_S(\mathbf{h} \mid r) = \exp\left(-r|\mathbf{h}|^2\right) ; \text{ donc } \quad f_S(\boldsymbol{\omega} \mid r) = \frac{1}{(4\pi r)^{d/2}} \exp\left(-\frac{|\boldsymbol{\omega}|^2}{4r}\right)$$

$F_T(\tau \mid \omega, r)$

Soit $\Omega_T(\boldsymbol{\omega}, r) \sim F_T(\tau \mid \boldsymbol{\omega}, r)$. Alors on peut montrer que sa fonction caractéristique est

$$\mathbb{E}\left[e^{iu\,\Omega_T(\boldsymbol{\omega},r)}\right] = \exp\left(-|\boldsymbol{\omega}|^2\,\gamma(u)/4r\right)$$

$\Rightarrow$ Algorithme de simulation pour générer $\Omega_T(\boldsymbol{\omega}, r)$:

- **Algorithme générique** utilisant un processus de points inhomogène (non détaillé ici)
- **Algorithme spécifique** pour chaque famille paramétrique de $\gamma(u)$, cf ci-dessous pour un exemple

Algorithme général

Algorithm 1 Simulation spectrale d'un champs aléatoire Gaussien avec une covariance ST de la classe de Gneiting

Require: d and p
Require: μ, γ
1: **for** $j = 1$ to p **do**
2: Simuler $R_j \sim \mu$;
3: Simuler $(\mathbf{\Omega}_j^S \mid R_j) \sim \sqrt{2R}\mathcal{N}_d(\mathbf{0}, \mathbf{I}_d)$;
4: Simuler $(\Omega_j^T \mid \mathbf{\Omega}_j^S, R_j) \sim F_T$ (avec un algorithme générique ou spécifique);
5: Simuler $\Phi_j \sim \mathcal{U}(0, 2\pi)$;
6: Simuler $U_j \sim \mathcal{U}(0, 1)$;
7: **end for**
8: Pour tout $(\mathbf{x}, t) \in \mathbb{R}^d \times \mathbb{R}$ calculer et retourner

$$\tilde{Z}(\mathbf{x}, t) = \sum_{j=1}^{p} \sqrt{-2\ln(U_j)/p}\, \cos\left(\langle \mathbf{\Omega}_j^S, \mathbf{x}\rangle + \Omega_j^T t + \Phi_j\right)$$

.

Illustration

- Modèle utilisé pour l'analyse du "Irish wind dataset" (Gneiting, 2002)
- $$\gamma(u) = (1 + a|u|^{\alpha})^{\beta} - 1$$

  avec $a > 0,\ 0 < \alpha \leq 2,\ 0 < \beta \leq 1,$
- Donc $C_T(u) = (1 + a|u|^{\alpha})^{-\beta}$ appartient à la classe dite de Cauchy
- Dans ce cas spécifique,

$$\mathbb{E}\left[e^{iu\,\Omega_T(\boldsymbol{\omega},r)}\right] = \exp\left(-\lambda[(1 + a|u|^{\alpha})^{\beta} - 1]\right), \quad \lambda = |\boldsymbol{\omega}|^2/4r$$

- $\Omega_T(\boldsymbol{\omega}, r)$ est un mélange de lois stables (Allard et al., 2020)

Algorithm 2 Simuler $\Omega_T(\boldsymbol{\omega} \mid r)$ avec une covariance temporelle de Cauchy

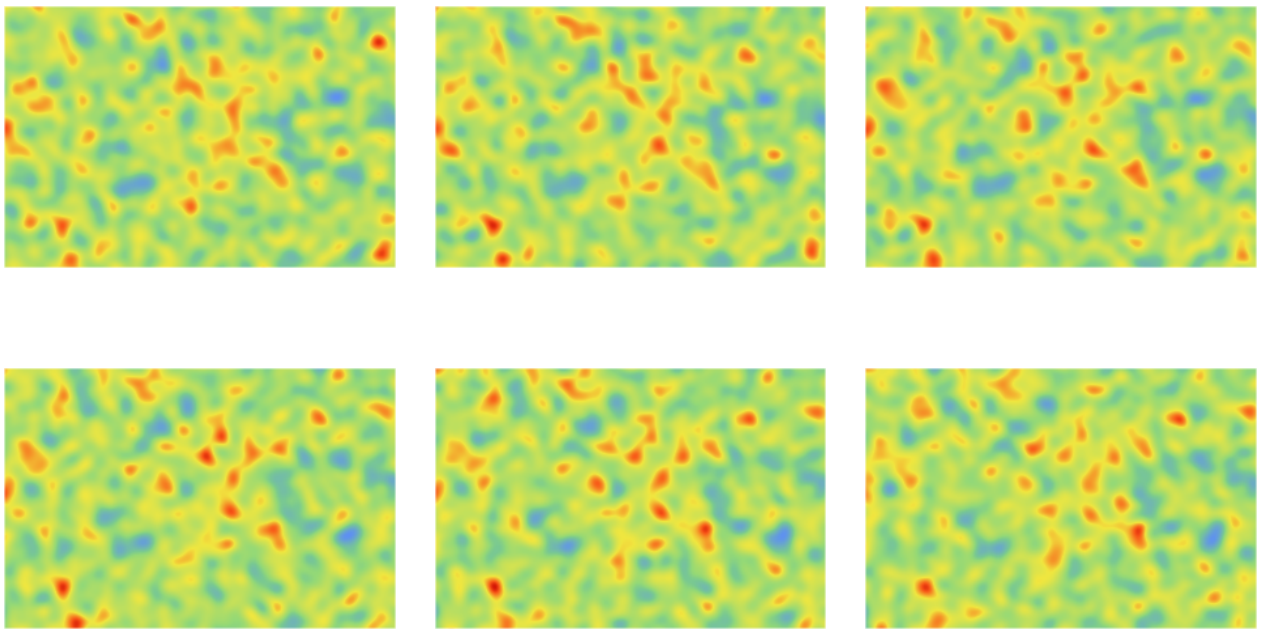
Require: $a > 0,\ 0 < \alpha \leq 2,\ 0 < \beta \leq 1,\ \lambda = \frac{|\boldsymbol{\omega}|^2}{4r}$

1: Simuler $S \sim \mathcal{S}^{+}(\beta, \lambda^{1/\beta})$;
2: Simuler $T \sim \mathcal{S}(\alpha)$;
3: Retourner $\Omega^T(\boldsymbol{\omega}, r) = T\,(Sa\lambda^{1/\beta})^{1/\alpha}$.

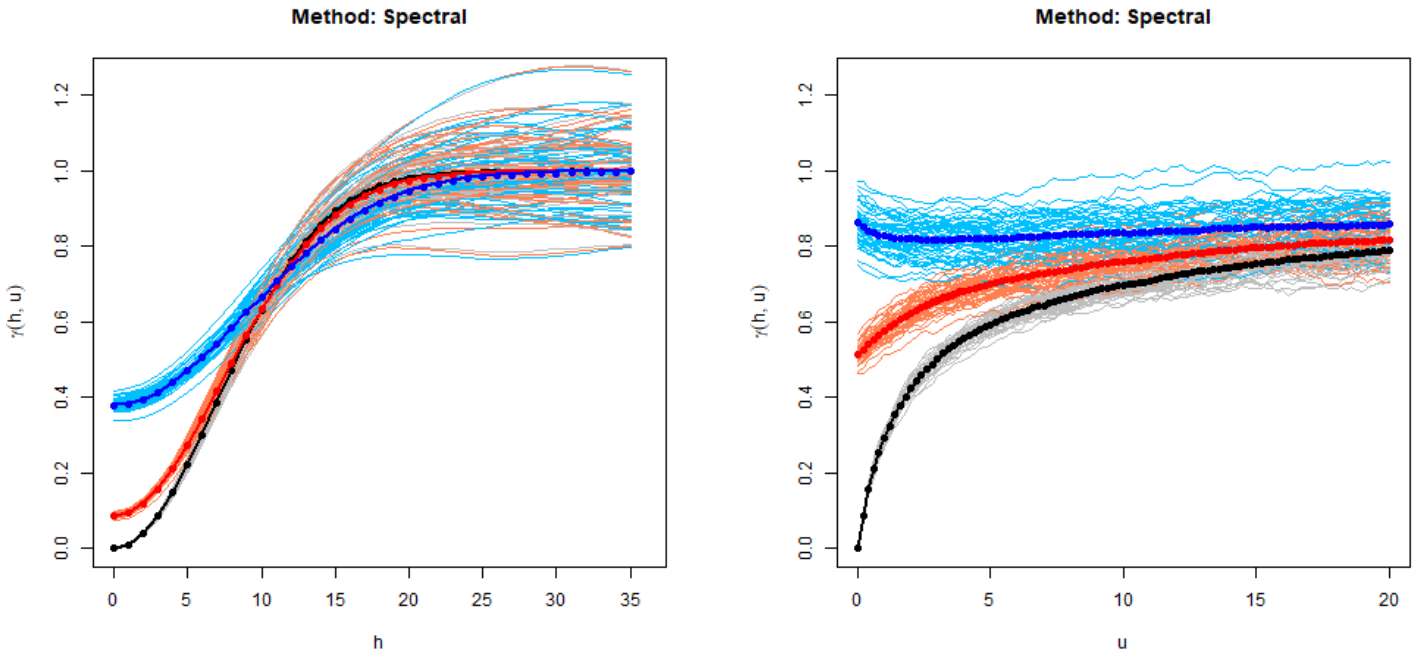
Illustration

Réalisation d'un champ aléatoire spatio-temporel avec une covariance appartenant à la classe de Gneiting

$$C(\mathbf{h}, u) = (1 + |u|)^{-1/2} \exp\left(-0.01\, |\mathbf{h}|^2 (1 + |u|)^{-1/2}\right)$$



Illustration



50 réalisations sur un domaine $100 \times 100 \times 100$ dans $\mathbb{R}^2 \times \mathbb{R}$; discrétisation spatiale 1×1; pas de temps 0.2. Gauche: $u = 0$ (noire), $u = 0.2$ (rouge) and $u = 1.6$ (bleu). Droite: $\mathbf{h} = (0, 0)$ (noir), $\mathbf{h} = (6, 6)$ (rouge) et $\mathbf{h} = (10, 10)$ (bleu).

Approche par substitution: motivation

Simuler $\Omega_T(s\omega, r)$ peut être assez pénible:

- Soit utiliser une algorithme générique (mais c'est un algorithme approché)
- Ou alors un algorithme taillé sur mesure pour chaque famille paramétrique $\gamma(u)$.

Peut-on faire autrement ?

$\Rightarrow$ Simuler une fonction aléatoire $W(t)$ ayant pour variogramme $\gamma(u)$ et insérer $W(t)$ dans l'approche spectrale pour la dimensions spatiale.

Approche par substitution

Posons

$$Z(\mathbf{x}, t) = \sqrt{-2\ln(U)}\cos\left(\sqrt{2R}\,\langle\tilde{\mathbf{\Omega}}, \mathbf{x}\rangle + \frac{|\tilde{\mathbf{\Omega}}|}{\sqrt{2}}W(t) + \Phi\right)$$

où

- $U \sim \mathcal{U}(0,1)$,
- $R \sim \mu$,
- $\tilde{\mathbf{\Omega}} \sim \mathcal{N}_k(\mathbf{0}, \mathbf{I}_k)$,
- $W(t) \sim \mathcal{GP}(\gamma(u))$
- $\Phi \sim \mathcal{U}(0, 2\pi)$,

tous indépendants. Alors,

$$\mathbb{C}\mathrm{ov}\big(Z(\mathbf{x}, t), Z(\mathbf{x}+\mathbf{h}, t+u)\big) = C(\mathbf{h}, u)$$

Preuve basée sur des arguments trigonométriques habituels.

Algorithm

Algorithm 3 Algorithme par substitution

Require: μ et $\gamma(u)$
Require: p

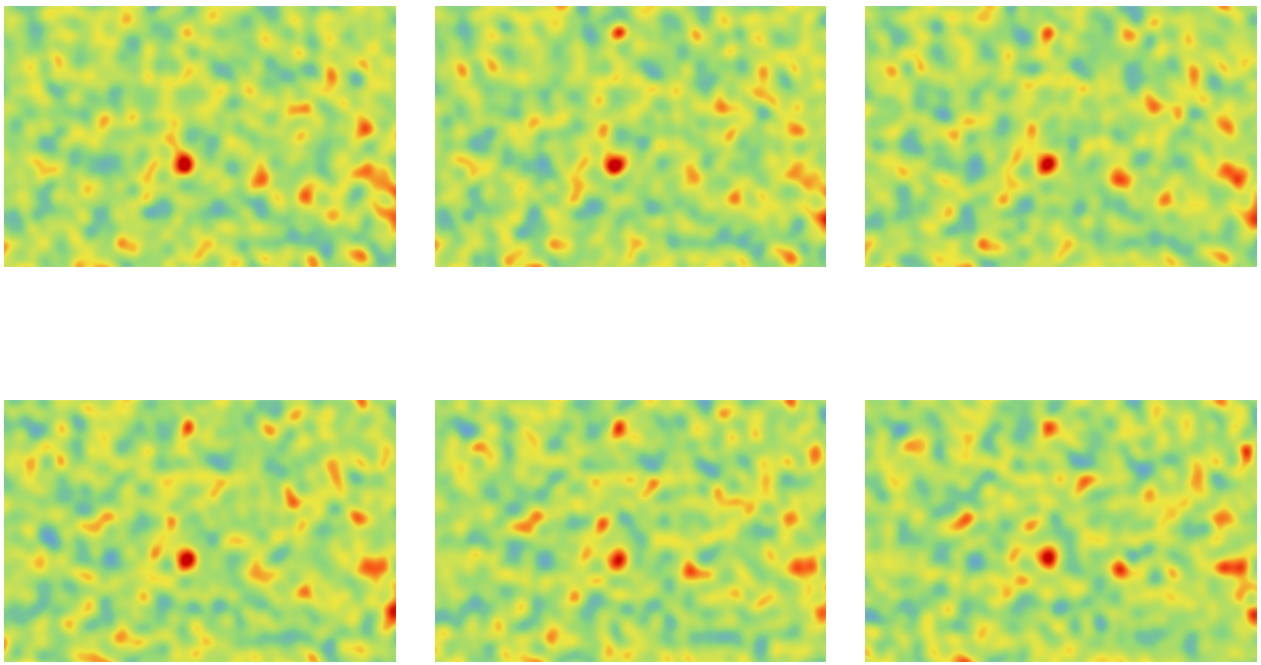
1: **for** $j = 1$ to p **do**
2: Simuler $R_j \sim \mu$;
3: Simuler $\tilde{\mathbf{\Omega}}_j \sim \mathcal{N}_k(\mathbf{0}, \mathbf{I}_k)$;
4: Simuler $\Phi_j \sim \mathcal{U}(0, 2\pi)$;
5: Simuler $U_j \sim \mathcal{U}(0, 1)$;
6: Simuler un processus indépendant avec accroissements Gaussiens, ayant pour variogramme $\gamma(u)$.
7: **end for**
8: Pour tout $(\mathbf{x}, t) \in \mathbb{R}^d \times \mathbb{R}$ calculer et retourner

$$\tilde{Z}(\mathbf{x}, t) = \sum_{j=1}^{p} \sqrt{\frac{-2\ln(U_j)}{p}} \cos\left(\sqrt{2R_j}\langle\tilde{\mathbf{\Omega}}_j, \mathbf{x}\rangle + \frac{|\tilde{\mathbf{\Omega}}_j|}{\sqrt{2}} W_j(t) + \Phi_j\right)$$

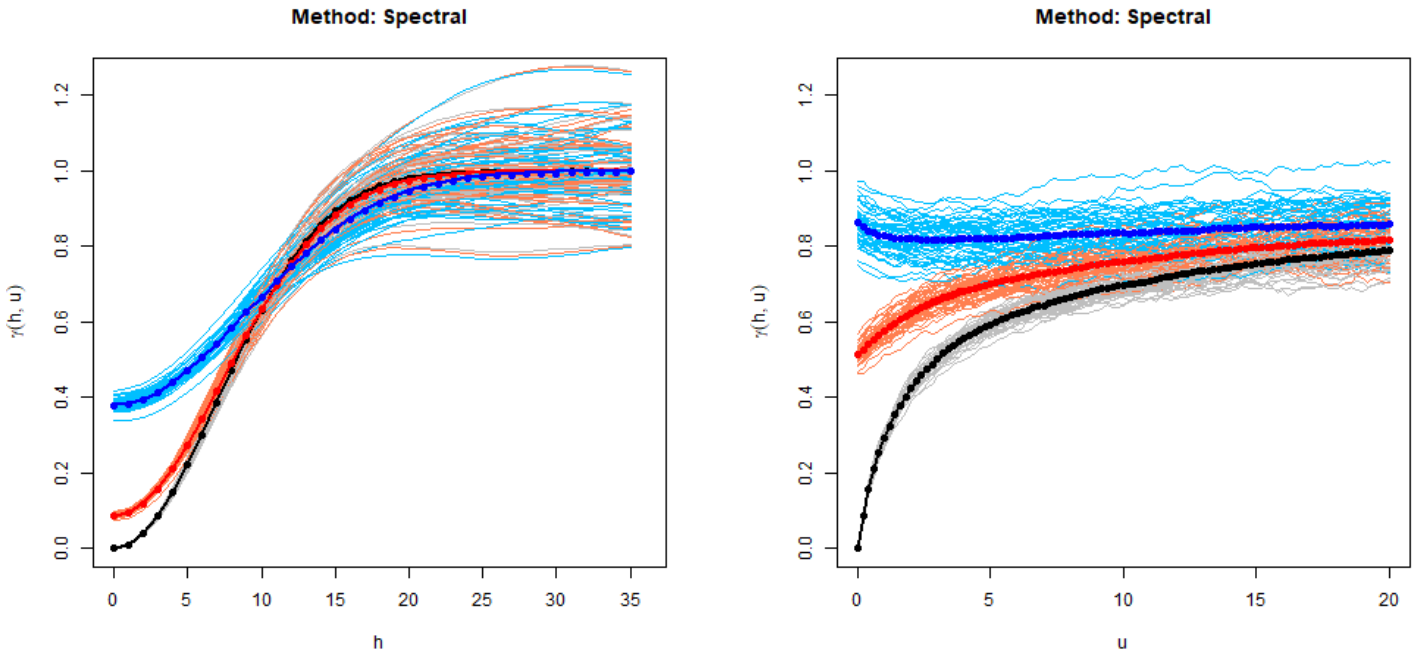
Illustration

Réalisation d'un champ aléatoire spatio-temporel avec une covariance appartenant à la classe de Gneiting

$$C(\mathbf{h}, u) = (1 + |u|)^{-1/2} \exp\left(-0.01 \, |\mathbf{h}|^2 \, (1 + |u|)^{-1/2}\right)$$



Illustration



50 réalisations sur un domaine $100 \times 100 \times 100$ dans $\mathbb{R}^2 \times \mathbb{R}$; discrétisation spatiale 1×1; pas de temps 0.2. Gauche: $u = 0$ (noire), $u = 0.2$ (rouge) and $u = 1.6$ (bleu). Droite: $\mathbf{h} = (0, 0)$ (noir), $\mathbf{h} = (6, 6)$ (rouge) et $\mathbf{h} = (10, 10)$ (bleu).

Spectral vs. Substitution

La simulation dans l'espace est identique

$$F(d\omega, d\tau \mid r) = F_S(d\omega \mid r) F_T(d\tau \mid d\omega, r)$$

avec $R \sim \mu$ et $\mathbf{\Omega}^S(R) \sim \mathcal{N}_d(\mathbf{0}, \mathbf{I}_d)$

La simulation dans le temps est différente

- Approche spectrale : simuler Ω^T, puis projeter

$$\cos\left(\sqrt{2R_j}\langle \mathbf{\Omega}_j^S, \mathbf{x}\rangle + \Omega_j^T t + \Phi_j\right)$$

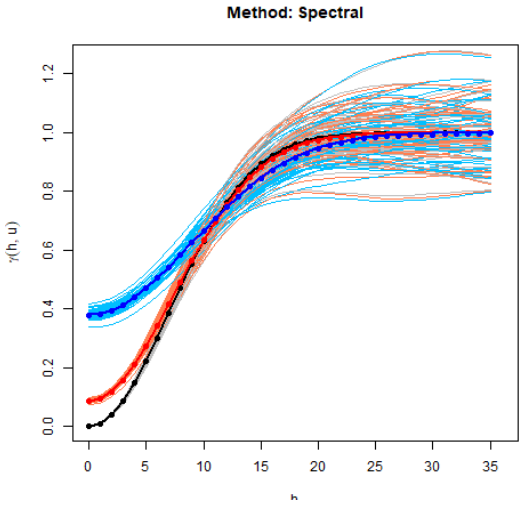
- Approche par substitution : simuler $W(t) \sim \mathcal{GP}(0, \gamma(u))$, puis projeter

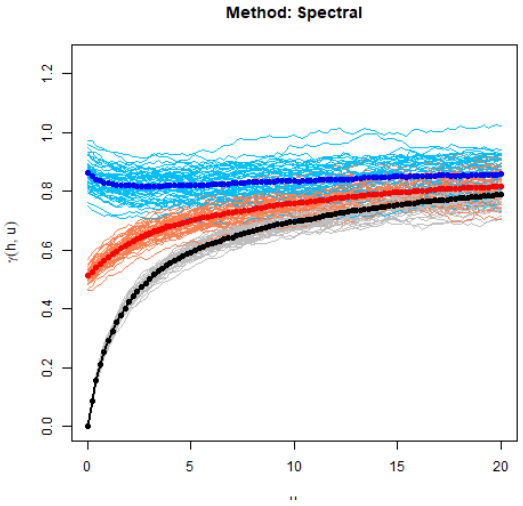
$$\cos\left(\sqrt{2R_j}\langle \tilde{\mathbf{\Omega}}_j, \mathbf{x}\rangle + \frac{|\tilde{\mathbf{\Omega}}_j|}{\sqrt{2}} W_j(t) + \Phi_j\right)$$

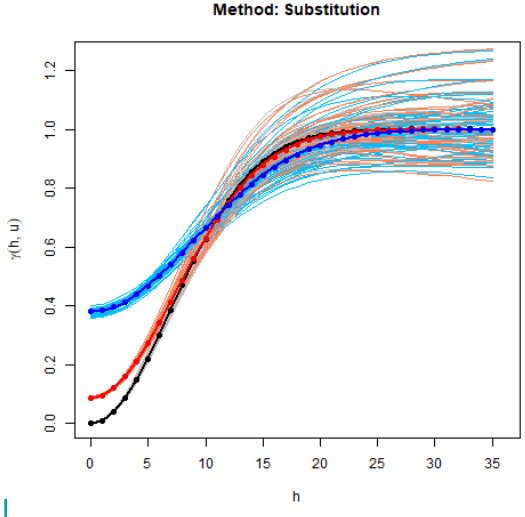
$\surd$ La densité spectral temporelle n'a pas besoin d'être connue

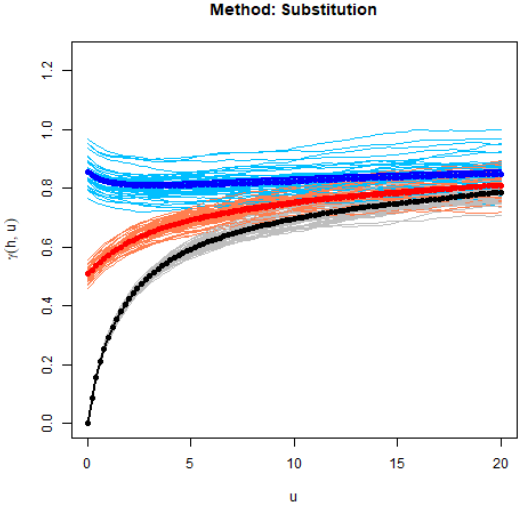
$\surd$ La simulation de $W(t)$ peut se faire par Cholesky (rarement besoin de plus de 10 000 pas de temps).

Illustration









En résumé

- Approches de type spectrales
- Coûts fixes faibles, puis temps proportionnel au nombre de points de simulation
- Nécessite peu de mémoire
- Plus ν est faible, plus p doit être élevé (ici $p = 5000$)
- Extension vers les covariances non-stationnaires en utilisant l'échantillonnage d'importance (*importance sampling*)
- Extensions possibles vers
  - Sphères $\times$ Time
  - Modèles spatio-temporels multivariés
  - Construction de "type Gneiting", plus générales

Références

Allard, D., Emery, X., Lacaux, C., Lantuéjoul, C. (2020).

Simulating space-time random fields with nonseparable Gneiting-type covariance functions.
Statistics and Computing, 30, 1479-?1495. https://doi.org/10.1007/s11222-020-09956-4

Bakka, H., Krainski, E., Bolin, D., Rue, H., Lindgren, F. (2020).

The diffusion-based extension of the Matérn field to space-time.
arXiv:2006.04917.

Cameletti, M., Lindgren, F., Simpson, D., Rue, H. (2013).

Spatio-temporal modeling of particulate matter concentration through the SPDE approach.
AStA Advances in Statistical Analysis, 97(2), 109-131.

Carrizo-Vergara, R., Allard, D., and Desassis, N. (2018).

A general framework for SPDE-based stationary random fields.
arXiv:1806.04999.

Cox, D. R. and Isham, V. (1988).

A simple spatial-temporal model of rainfall.
Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences, 415(1849):317–328.

Cressie, N., Huang, H. C. (1999).

Classes of nonseparable, spatio-temporal stationary covariance functions.
Journal of the American Statistical Association, 94(448), 1330–1339.

INRAE

département MathNum

Biostatistique BIOSP & Processus Spatiaux

Références

De Iaco, S., Myers, D. E., and Posa, D. (2001).

Space–time analysis using a general product–sum model.
Statistics & Probability Letters, 52(1):21–28.

Gneiting, T. (2002).

Nonseparable, stationary covariance functions for space–time data.
Journal of the American Statistical Association, 97(458):590–600.

Huang, H., Sun, Y. (2019)

Visualization and assessment of spatio-temporel covariance properties.
Spatial Statistics, 34, 100272

Paciorek, C. J., Schervish, M. J. (2006).

Spatial modelling using a new class of nonstationary covariance functions.
Environmetrics, 17(5), 483-506.

Rodrigues, A., Diggle, P. J. (2010).

A Class of Convolution-Based Models for Spatio-Temporal Processes with Non-Separable Covariance Structure
Scandinavian Journal of Statistics, 37:553?567,

Schlather, M. (2010).

Some covariance models based on normal scale mixtures.
Bernoulli, 16(3), 780-797.

Schlather, M., Malinowski, A., Menck, P. J., Oesting, M., Strokorb, K., et al. (2015).

Analysis, simulation and prediction of multivariate random fields with package RandomFields.
Journal of Statistical Software, 63(8):1–25.

Schoenberg, I.-J. (1938).

Metric spaces and completely monotone functions.
Annals of Mathematics, 39(4):811–831.

Zastavnyi, V. P. and Porcu, E. (2011).

Characterization theorems for the Gneiting class of space–time covariances.