MODELOS ESPAÇO-TEMPORAIS NÃO GAUSSIANOS

Thais C O da Fonseca - IM - UFRJ

Em colaboração com Prof Mark F J Steel - Warwick University

Abril, 2011

TYPICAL PROBLEM

- Given: observations $Z(s_i, t_j)$ at a finite number locations $s_i,\ i = 1, \ldots, I$ and time points $t_j,\ j = 1, \ldots, J$.

- Desired: predictive distribution for the unknown value $Z(s_0, t_0)$ at the space-time coordinate (s_0, t_0).

- Focus: continuous space and continuous time which allow for prediction and interpolation at any location and any time.

$$Z(s,t),\ (s,t) \in D \times T,\ \text{where } D \subseteq \Re^d,\ T \subseteq \Re$$

TYPICAL PROBLEM

- Given: observations $Z(s_i, t_j)$ at a finite number locations $s_i, \; i = 1, \ldots, I$ and time points $t_j, \; j = 1, \ldots, J$.
- Desired: predictive distribution for the unknown value $Z(s_0, t_0)$ at the space-time coordinate (s_0, t_0).
- Focus: continuous space and continuous time which allow for prediction and interpolation at any location and any time.

$$Z(s,t), \; (s,t) \in D \times T, \text{ where } D \subseteq \Re^d, \; T \subseteq \Re$$

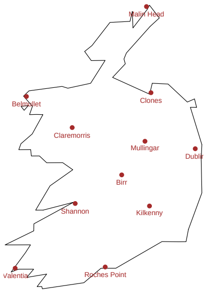
TYPICAL PROBLEM

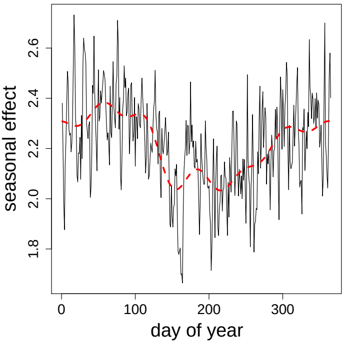
- Given: observations $Z(s_i, t_j)$ at a finite number locations $s_i,\ i = 1, \ldots, I$ and time points $t_j,\ j = 1, \ldots, J$.
- Desired: predictive distribution for the unknown value $Z(s_0, t_0)$ at the space-time coordinate (s_0, t_0).
- Focus: continuous space and continuous time which allow for prediction and interpolation at any location and any time.

$$Z(s,t),\ (s,t) \in D \times T, \text{ where } D \subseteq \Re^d,\ T \subseteq \Re$$

EXAMPLE: IRISH WIND DATA [HASLETT AND RAFTERY, 1989]

Daily average wind speed in m/s at 11 meteorological stations in Ireland during the period 1961-1970.





Plot 1: Location of the 11 stations in Ireland;

Plot 2: Mean wind over all stations and years for each day of the year and fitted mean.

REAL-VALUED STOCHASTIC PROCESSES

The uncertainty of the unobserved parts of the process can be expressed probabilistically by a random function in space and time:

$$\{Z(s,t); s \in D \subset \Re^d, t \in T \subseteq \Re_+\}$$

MEAN FUNCTION:

$$m(s,t) = E(Z(s,t)) = \int z(s,t)dF(z),$$

COVARIANCE FUNCTION:

$$\mathrm{Cov}(Z(s_1,t_1), Z(s_2,t_2)) = \int [z(s_1,t_1) - m(s_1,t_1)][z(s_2,t_2) - m(s_2,t_2)]dF(z_1,z_2),$$

where $(s,t), (s_1,t_1), (s_2,t_2) \in D \times T$.

REAL-VALUED STOCHASTIC PROCESSES

The uncertainty of the unobserved parts of the process can be expressed probabilistically by a random function in space and time:

$$\{Z(s,t); s \in D \subset \Re^d, t \in T \subseteq \Re_+\}$$

MEAN FUNCTION:

$$m(s,t) = E(Z(s,t)) = \int z(s,t)dF(z),$$

COVARIANCE FUNCTION:

$$\text{Cov}(Z(s_1,t_1),Z(s_2,t_2)) = \int [z(s_1,t_1)-m(s_1,t_1)][z(s_2,t_2)-m(s_2,t_2)]dF(z_1,z_2),$$

where $(s,t),(s_1,t_1),(s_2,t_2) \in D \times T$.

REAL-VALUED STOCHASTIC PROCESSES

The uncertainty of the unobserved parts of the process can be expressed probabilistically by a random function in space and time:

$$\{Z(s,t); s \in D \subset \Re^d, t \in T \subseteq \Re_+\}$$

MEAN FUNCTION:

$$m(s,t) = E(Z(s,t)) = \int z(s,t)dF(z),$$

COVARIANCE FUNCTION:

$$\mathrm{Cov}(Z(s_1,t_1), Z(s_2,t_2)) = \int [z(s_1,t_1) - m(s_1,t_1)][z(s_2,t_2) - m(s_2,t_2)]dF(z_1,z_2),$$

where $(s,t),(s_1,t_1),(s_2,t_2) \in D \times T$.

VALID FUNCTIONS

- We need to specify a valid covariance structure for the process.

$$C(s_1, s_2; t_1, t_2) = \mathrm{Cov}(Z(s_1, t_1), Z(s_2, t_2))$$

- Positive definiteness: C has to imply that $\sum_{i=1}^{n} a_i Z(s_i, t_i)$ has positive variance for any $(s_1, t_1), \ldots, (s_n, t_n)$, any real a_1, ..., a_n, and any positive integer n.

- It is quite difficult to check whether a function is positive definite.

VALID FUNCTIONS

- We need to specify a valid covariance structure for the process.

$$C(s_1, s_2; t_1, t_2) = \mathrm{Cov}(Z(s_1, t_1), Z(s_2, t_2))$$

- Positive definiteness: C has to imply that $\sum_{i=1}^{n} a_i Z(s_i, t_i)$ has positive variance for any $(s_1, t_1), \ldots, (s_n, t_n)$, any real a_1, ..., a_n, and any positive integer n.

- It is quite difficult to check whether a function is positive definite.

VALID FUNCTIONS

- We need to specify a valid covariance structure for the process.

$$C(s_1, s_2; t_1, t_2) = \mathrm{Cov}(Z(s_1, t_1), Z(s_2, t_2))$$

- Positive definiteness: C has to imply that $\sum_{i=1}^{n} a_i Z(s_i, t_i)$ has positive variance for any $(s_1, t_1), \ldots, (s_n, t_n)$, any real a_1, ..., a_n, and any positive integer n.

- It is quite difficult to check whether a function is positive definite.

SIMPLIFYING ASSUMPTIONS

- One way to ensure positive definiteness: separability

$$\mathrm{Cov}(Z(s_1, t_1), Z(s_2, t_2)) = C_1(s_1, s_2)C_2(t_1, t_2),$$

C_1 and C_2 are valid functions in space and time, respectively.

- Other simplifying assumptions:
  - Stationarity: $\mathrm{Cov}(Z(s_1, t_1), Z(s_2, t_2)) = C(s_1 - s_2, t_1 - t_2)$.
  - Isotropy: $\mathrm{Cov}(Z(s_1, t_1), Z(s_2, t_2)) = C(\|s_1 - s_2\|, |t_1 - t_2|)$.
  - Gaussianity: The process has finite dimensional Gaussian distribution.

- Initially, I will consider Gaussian processes with stationary and isotropic covariance functions.

- First I relax the separability assumption.

Simplifying assumptions

- One way to ensure positive definiteness: separability

$$\mathrm{Cov}(Z(s_1,t_1),Z(s_2,t_2)) = C_1(s_1,s_2)C_2(t_1,t_2),$$

C_1 and C_2 are valid functions in space and time, respectively.

- Other simplifying assumptions:
  - Stationarity: $\mathrm{Cov}(Z(s_1,t_1),Z(s_2,t_2)) = C(s_1-s_2,t_1-t_2)$.
  - Isotropy: $\mathrm{Cov}(Z(s_1,t_1),Z(s_2,t_2)) = C(||s_1-s_2||,|t_1-t_2|)$.
  - Gaussianity: The process has finite dimensional Gaussian distribution.

- Initially, I will consider Gaussian processes with stationary and isotropic covariance functions.

- First I relax the separability assumption.

SIMPLIFYING ASSUMPTIONS

- One way to ensure positive definiteness: separability

$$\mathrm{Cov}(Z(s_1, t_1), Z(s_2, t_2)) = C_1(s_1, s_2)C_2(t_1, t_2),$$

C_1 and C_2 are valid functions in space and time, respectively.

- Other simplifying assumptions:
  - Stationarity: $\mathrm{Cov}(Z(s_1, t_1), Z(s_2, t_2)) = C(s_1 - s_2, t_1 - t_2)$.
  - Isotropy: $\mathrm{Cov}(Z(s_1, t_1), Z(s_2, t_2)) = C(||s_1 - s_2||, |t_1 - t_2|)$.
  - Gaussianity: The process has finite dimensional Gaussian distribution.
- Initially, I will consider Gaussian processes with stationary and isotropic covariance functions.
- First I relax the separability assumption.

Definition

Separable covariance function

A stationary isotropic separable covariance function is defined as

$$\text{Cov}(Z(s_1, t_1), Z(s_2, t_2)) = C(s_1 - s_2, t_1 - t_2) = C_1(s_1 - s_2)C_2(t_1 - t_2), \quad (1)$$

where $(s_1, t_1), (s_2, t_2) \in D \times T$.

- It is computationally very convenient: $\Sigma = \Sigma_1 \otimes \Sigma_2$.
- But it is very unrealistic.
- It has severe theoretical limitations.

Definition

Separable covariance function

A stationary isotropic separable covariance function is defined as

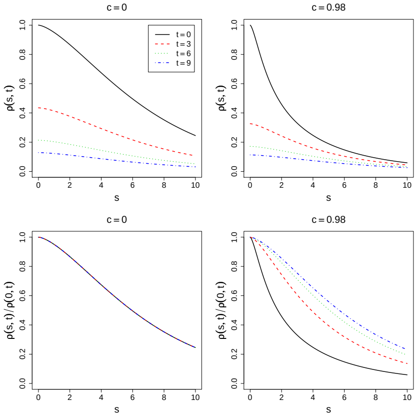
$$\text{Cov}(Z(s_1, t_1), Z(s_2, t_2)) = C(s_1 - s_2, t_1 - t_2) = C_1(s_1 - s_2)C_2(t_1 - t_2), \tag{1}$$

where $(s_1, t_1), (s_2, t_2) \in D \times T$.

- It is computationally very convenient: $\Sigma = \Sigma_1 \otimes \Sigma_2$.
- But it is very unrealistic.
- It has severe theoretical limitations.

THEORETICAL LIMITATION

Separability means that for different fixed time points, the marginal spatial covariances are just proportional.

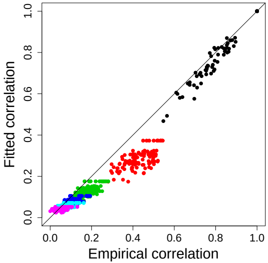


Plot of $\rho(s,t) = \frac{C(s,t)}{C(0,0)}$ for separable and nonseparable covariance functions.

EXAMPLE: IRISH WIND DATA

Separable function:

$$C(s,t) = \sigma^2 \left\{ 1 + \frac{||s/a||^{\alpha}}{\delta} \right\}^{-\lambda_1/2} \frac{K_{\lambda_1}\left(2\delta\sqrt{1 + \frac{||s/a||^{\alpha}}{\delta}} \right)}{K_{\lambda_1}(2\delta)} \left\{ 1 + |t/b|^{\beta} \right\}^{-\lambda_2}.$$



Plot: posterior median of the correlation function against empirical correlation at temporal lags zero until five, with black corresponding to lag 0, red to lag 1, green to lag 2, dark blue to lag 3, light blue to lag 4 and pink to lag 5.

SOME MODELS PROPOSED IN THE LITERATURE

- [Cressie and Huang, 1999] proposed a model that is not always valid;
- [Gneiting, 2002] proposed a model that might have lack of smoothness away from the origin;
- [Ma, 2002] proposed an approach based on the mixture of separable covariance functions;
- [Fuentes et al, 2005] proposed covariance functions in $d+1$ dimensions.
- I will consider the mixture approach.

SOME MODELS PROPOSED IN THE LITERATURE

- [Cressie and Huang, 1999] proposed a model that is not always valid;
- [Gneiting, 2002] proposed a model that might have lack of smoothness away from the origin;
- [Ma, 2002] proposed an approach based on the mixture of separable covariance functions;
- [Fuentes et al, 2005] proposed covariance functions in $d + 1$ dimensions.
- I will consider the mixture approach.

CONTINUOUS MIXTURE

MIXTURE MODELS

$$C(s,t) = \int C_1(s;u)C_2(t;v)dF(u,v) \quad (2)$$

- Idea: convex combinations of valid separable covariance functions are valid and nonseparable functions.
- $C_1(s;u)$ is a valid spatial covariance in D and $C_2(t;v)$ is a valid temporal covariance in T.
- (U,V) is a bivariate nonnegative random vector with cumulative distribution function F.

CONTINUOUS MIXTURE

MIXTURE MODELS

$$C(s,t) = \int C_1(s;u)C_2(t;v)dF(u,v) \tag{2}$$

- Idea: convex combinations of valid separable covariance functions are valid and nonseparable functions.
- $C_1(s;u)$ is a valid spatial covariance in D and $C_2(t;v)$ is a valid temporal covariance in T.
- (U,V) is a bivariate nonnegative random vector with cumulative distribution function F.

MAIN ADVANTAGES

1. One may take advantage of the whole available literature of spatial statistics and time series; C is the unconditional covariance of

$$Z(s, t; U, V) = Z_1(s; U) Z_2(t; V)$$

2. It is natural to make separate modeling decisions regarding the spatial and temporal components, eg. smoothness and long range dependence can be different across space and time.

MAIN ADVANTAGES

1. One may take advantage of the whole available literature of spatial statistics and time series; C is the unconditional covariance of

$$Z(s,t;U,V) = Z_1(s;U)Z_2(t;V)$$

2. It is natural to make separate modeling decisions regarding the spatial and temporal components, eg. smoothness and long range dependence can be different across space and time.

DEFINING F

MIXTURE MODELS

If $C_1(s;u) = \sigma_1 \exp\{-\gamma_1(s)u\}$ and $C_2(t;v) = \sigma_2 \exp\{-\gamma_2(t)v\}$ then

$$\begin{aligned} C(s,t) &= \int C_1(s;u)C_2(t;v)dF(u,v) \\ &= \sigma^2 M(-\gamma_1(s), -\gamma_2(t)) \end{aligned}$$

where $\gamma_1(s) = ||s/a||^\alpha$ and $\gamma_2(t) = |t/b|^\beta$. And $M(.,.)$ is the joint moment generating function of (U, V).

INTERACTION IN SPACE AND TIME

- If (U, V) independent then

$$C(s,t) = \sigma^2 M(-\gamma_1(s), -\gamma_2(t)) = \sigma^2 M_1(-\gamma_1(s)) M_2(-\gamma_2(t)),$$

that is, C is separable.

- The dependence between U and V will define the interaction between spatial and temporal components.

- Thus, the definition of the joint distribution F is crucial in the spatiotemporal covariance modelling.

INTERACTION IN SPACE AND TIME

- If (U, V) independent then

$$C(s,t) = \sigma^2 M(-\gamma_1(s), -\gamma_2(t)) = \sigma^2 M_1(-\gamma_1(s)) M_2(-\gamma_2(t)),$$

  that is, C is separable.

- The dependence between U and V will define the interaction between spatial and temporal components.

- Thus, the definition of the joint distribution F is crucial in the spatiotemporal covariance modelling.

INTERACTION IN SPACE AND TIME

- If (U, V) independent then

$$C(s,t) = \sigma^2 M(-\gamma_1(s), -\gamma_2(t)) = \sigma^2 M_1(-\gamma_1(s)) M_2(-\gamma_2(t)),$$

that is, C is separable.

- The dependence between U and V will define the interaction between spatial and temporal components.

- Thus, the definition of the joint distribution F is crucial in the spatiotemporal covariance modelling.

DEFINING THE JOINT DISTRIBUTION OF (U, V)

Simple way to generate dependence between U and V:

$U = X_0 + X_1$ and $V = X_0 + X_2$

MIXTURE MODELS

$$\begin{aligned} C(s,t) &= \sigma^2 M(-\gamma_1(s), -\gamma_2(t)) \\ &= \sigma^2 M_0(-\gamma_1(s) - \gamma_2(t)) M_1(-\gamma_1(s)) M_2(-\gamma_2(t)), \end{aligned} \qquad (3)$$

where $M_k(.)$ is the joint moment generating function of $X_k,\ k = 0, 1, 2$.

Possible choices for $X_k, k = 0, 1, 2$

Cauchy covariance

If $X_k \sim Ga(\lambda_k, 1)$ then $M_k(x) = (1 - x)^{-\lambda_k}$;

Matérn covariance

If $X_k \sim InvGa(\nu, 1)$ then $M_k(x) = \frac{(2\sqrt{x})^{\nu}}{2^{\nu-1}\Gamma(\nu)} K_{\nu}(2\sqrt{x})$;

Generalized Matérn covariance

If $X_k \sim GIG(\lambda_k, \delta, \delta)$ then $M_k(x) = \left\{1 - \frac{x}{\delta}\right\}^{-\lambda_1/2} \frac{K_{\lambda_1}\left(2\delta\sqrt{1-\frac{x}{\delta}}\right)}{K_{\lambda_1}(2\delta)}$;

POSSIBLE CHOICES FOR $X_k, k = 0, 1, 2$

CAUCHY COVARIANCE

If $X_k \sim Ga(\lambda_k, 1)$ then $M_k(x) = (1-x)^{-\lambda_k}$;

MATÉRN COVARIANCE

If $X_k \sim InvGa(\nu, 1)$ then $M_k(x) = \frac{(2\sqrt{x})^{\nu}}{2^{\nu-1}\Gamma(\nu)} K_\nu\left(2\sqrt{x}\right)$;

GENERALIZED MATÉRN COVARIANCE

If $X_k \sim GIG(\lambda_k, \delta, \delta)$ then $M_k(x) = \left\{1 - \frac{x}{\delta}\right\}^{-\lambda_1/2} \frac{K_{\lambda_1}\left(2\delta\sqrt{1-\frac{x}{\delta}}\right)}{K_{\lambda_1}(2\delta)}$;

POSSIBLE CHOICES FOR $X_k, k = 0, 1, 2$

CAUCHY COVARIANCE

If $X_k \sim Ga(\lambda_k, 1)$ then $M_k(x) = (1-x)^{-\lambda_k}$;

MATÉRN COVARIANCE

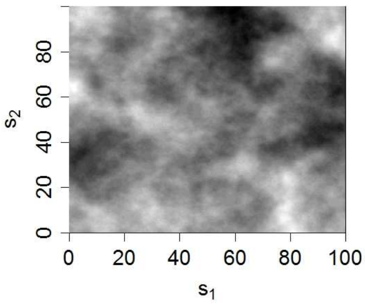
If $X_k \sim InvGa(\nu, 1)$ then $M_k(x) = \frac{(2\sqrt{x})^{\nu}}{2^{\nu-1}\Gamma(\nu)} K_{\nu}\left(2\sqrt{x}\right)$;

GENERALIZED MATÉRN COVARIANCE

If $X_k \sim GIG(\lambda_k, \delta, \delta)$ then $M_k(x) = \left\{1 - \frac{x}{\delta}\right\}^{-\lambda_1/2} \frac{K_{\lambda_1}\left(2\delta\sqrt{1-\frac{x}{\delta}}\right)}{K_{\lambda_1}(2\delta)}$;

MATÉRN COVARIANCE FUNCTION

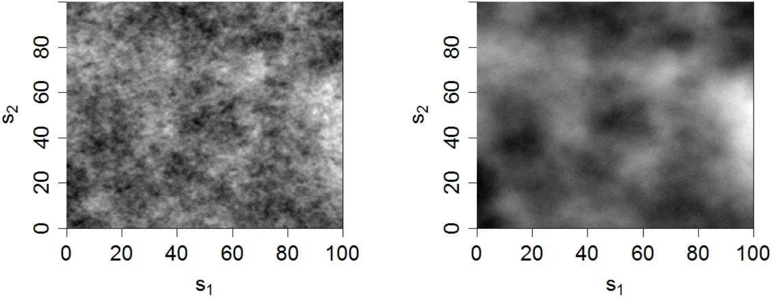
REALIZATION OF A GAUSSIAN RANDOM FUNCTION WITH $s = (s_1, s_2)$.



(a) $\nu = 1$. (b) $\nu = 2$.

CAUCHY COVARIANCE FUNCTION

REALIZATION OF A GAUSSIAN RANDOM FUNCTION WITH $s = (s_1, s_2)$.



(a) $\beta = 0.6$. (b) $\beta = 1.8$.

EXAMPLES OF SPATIOTEMPORAL FUNCTIONS

- Model 1: $X_0 \sim Ga(\lambda_0, 1)$, $X_1 \sim GIG(\lambda_1, \delta, \delta)$ and $X_2 \sim Ga(\lambda_2, 1)$

$$C(s,t) = \sigma^2 \left\{1 + ||s/a||^\alpha + |t/b|^\beta\right\}^{-\lambda_0} \left\{1 + \frac{||s/a||^\alpha}{\delta}\right\}^{-\lambda_1/2} \frac{K_{\lambda_1}\left(2\delta\sqrt{1 + \frac{||s/a||^\alpha}{\delta}}\right)}{K_{\lambda_1}(2\delta)} \left\{1 + |t/b|^\beta\right\}$$

- Model 2: $X_0 \sim Ga(\lambda_0, 1)$, $X_1 \sim InvGa(\nu, 1)$ and $X_2 \sim Ga(\lambda_2, 1)$

$$C(s,t) = \sigma^2 \left\{1 + ||s/a||^\alpha + |t/b|^\beta\right\}^{-\lambda_0} \frac{(2||s/a||^{\alpha/2})^\nu}{2^{\nu-1}\Gamma(\nu)} K_\nu\left(2||s/a||^{\alpha/2}\right) \left\{1 + |t/b|^\beta\right\}^{-\lambda_2}.$$

- Model 3: $X_0 \sim Ga(\lambda_0, 1)$, $X_1 \sim Ga(\lambda_1, 1)$ and $X_2 \sim Ga(\lambda_2, 1)$

$$C(s,t) = \sigma^2 \left\{1 + ||s/a||^\alpha + |t/b|^\beta\right\}^{-\lambda_0} \left\{1 + ||s/a||^\alpha\right\}^{-\lambda_1} \left\{1 + |t/b|^\beta\right\}^{-\lambda_2}.$$

EXAMPLES OF SPATIOTEMPORAL FUNCTIONS

- Model 1: $X_0 \sim Ga(\lambda_0, 1)$, $X_1 \sim GIG(\lambda_1, \delta, \delta)$ and $X_2 \sim Ga(\lambda_2, 1)$

$$C(s,t) = \sigma^2 \left\{1 + ||s/a||^\alpha + |t/b|^\beta\right\}^{-\lambda_0} \left\{1 + \frac{||s/a||^\alpha}{\delta}\right\}^{-\lambda_1/2} \frac{K_{\lambda_1}\left(2\delta\sqrt{1 + \frac{||s/a||^\alpha}{\delta}}\right)}{K_{\lambda_1}(2\delta)} \left\{1 + |t/b|^\beta\right\}$$

- Model 2: $X_0 \sim Ga(\lambda_0, 1)$, $X_1 \sim InvGa(\nu, 1)$ and $X_2 \sim Ga(\lambda_2, 1)$

$$C(s,t) = \sigma^2 \left\{1 + ||s/a||^\alpha + |t/b|^\beta\right\}^{-\lambda_0} \frac{(2||s/a||^{\alpha/2})^\nu}{2^{\nu-1}\Gamma(\nu)} K_\nu\left(2||s/a||^{\alpha/2}\right) \left\{1 + |t/b|^\beta\right\}^{-\lambda_2}.$$

- Model 3: $X_0 \sim Ga(\lambda_0, 1)$, $X_1 \sim Ga(\lambda_1, 1)$ and $X_2 \sim Ga(\lambda_2, 1)$

$$C(s,t) = \sigma^2 \left\{1 + ||s/a||^\alpha + |t/b|^\beta\right\}^{-\lambda_0} \left\{1 + ||s/a||^\alpha\right\}^{-\lambda_1} \left\{1 + |t/b|^\beta\right\}^{-\lambda_2}.$$

EXAMPLES OF SPATIOTEMPORAL FUNCTIONS

- Model 1: $X_0 \sim Ga(\lambda_0, 1)$, $X_1 \sim GIG(\lambda_1, \delta, \delta)$ and $X_2 \sim Ga(\lambda_2, 1)$

$$C(s,t) = \sigma^2 \left\{1 + ||s/a||^\alpha + |t/b|^\beta\right\}^{-\lambda_0} \left\{1 + \frac{||s/a||^\alpha}{\delta}\right\}^{-\lambda_1/2} \frac{K_{\lambda_1}\left(2\delta\sqrt{1 + \frac{||s/a||^\alpha}{\delta}}\right)}{K_{\lambda_1}(2\delta)} \left\{1 + |t/b|^\beta\right\}$$

- Model 2: $X_0 \sim Ga(\lambda_0, 1)$, $X_1 \sim InvGa(\nu, 1)$ and $X_2 \sim Ga(\lambda_2, 1)$

$$C(s,t) = \sigma^2 \left\{1 + ||s/a||^\alpha + |t/b|^\beta\right\}^{-\lambda_0} \frac{(2||s/a||^{\alpha/2})^\nu}{2^{\nu-1}\Gamma(\nu)} K_\nu\left(2||s/a||^{\alpha/2}\right) \left\{1 + |t/b|^\beta\right\}^{-\lambda_2}.$$

- Model 3: $X_0 \sim Ga(\lambda_0, 1)$, $X_1 \sim Ga(\lambda_1, 1)$ and $X_2 \sim Ga(\lambda_2, 1)$

$$C(s,t) = \sigma^2 \left\{1 + ||s/a||^\alpha + |t/b|^\beta\right\}^{-\lambda_0} \left\{1 + ||s/a||^\alpha\right\}^{-\lambda_1} \left\{1 + |t/b|^\beta\right\}^{-\lambda_2}.$$

DEGREE OF SEPARABILITY

It is defined by the correlation between (U, V).

MEASURE OF SEPARABILITY

$$c = corr(U, V) = \frac{Var(X_0)}{\sqrt{(Var(X_0) + Var(X_1))(Var(X_0) + Var(X_2))}}, \quad (4)$$

$0 \leq c \leq 1$.

- 0 means separability and 1 means strong nonseparability.

IRISH WIND DATA: MODEL COMPARISON

TABLE: Natural logarithm of the Bayes factor in favour of the nonseparable Model 1 using Newton-Raftery ($d = 0.01$), Bridge-sampling and Shifted-Gamma ($\lambda = 0.98$) estimators for the marginal likelihood. $log(BF) > 5$ suggests strong evidence.

	Newton-Raftery	Bridge-Sampling	Shifted-Gamma
Separable Model 1	49	46	50
Nonseparable Model 2	57	68	42
Nonseparable Model 3	6	9	7

POSTERIOR DISTRIBUTION OF c

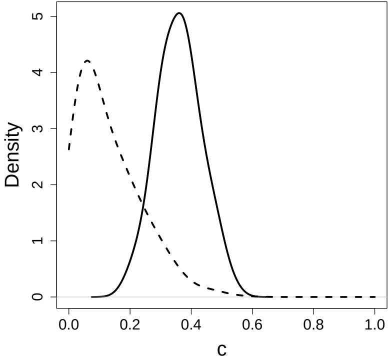
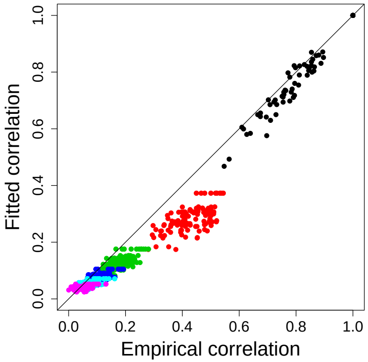
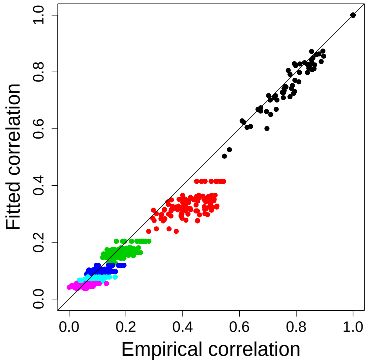


FIGURE: Nonseparable model 1: Posterior (solid line) and prior (dashed line) densities for c in (4).

EMPIRICAL FIT



(a) Separable model



(b) Nonseparable model

Plot: Posterior median of the correlation function against empirical correlation at temporal lags zero until five, with black corresponding to lag 0, red to lag 1, green to lag 2, dark blue to lag 3, light blue to lag 4 and pink to lag 5.

REALISTIC MODELS

- The model just presented can be easily extended to accommodate realistic features of space-time data as decisions regarding time and space can be taken separately;

$$Z(s, t; U, V) = Z_1(s; U)Z_2(t; V)$$

- The following extensions were considered:
  - Nongaussianity;
  - Nonstationarity;
  - Asymmetry.

ASYMMETRIC MODEL

- Notice the clear lack of fit at lag one.
- This is due to asymmetry of the covariance function at lag one.
- Simple way to address this problem:

$$C^*(s,t) = C(s - \epsilon t w, t),$$

where ϵ is a parameter to be estimated and w is a unit vector.

- As the asymmetries in this example are mainly functions of differences in longitude, we take $w = (0,1)$ as suggested by [Stein, 2005].
- In our framework, this is equivalent to replacing the variogram $\gamma_1(s)$ by $\gamma_1(s - \epsilon t w)$.

ASYMMETRIC MODEL

- Notice the clear lack of fit at lag one.
- This is due to asymmetry of the covariance function at lag one.
- Simple way to address this problem:

$$C^*(s, t) = C(s - \epsilon t w, t),$$

  where ϵ is a parameter to be estimated and w is a unit vector.
- As the asymmetries in this example are mainly functions of differences in longitude, we take $w = (0, 1)$ as suggested by [Stein, 2005].
- In our framework, this is equivalent to replacing the variogram $\gamma_1(s)$ by $\gamma_1(s - \epsilon t w)$.

ASYMMETRIC MODEL

- Notice the clear lack of fit at lag one.
- This is due to asymmetry of the covariance function at lag one.
- Simple way to address this problem:

$$C^*(s,t) = C(s - \epsilon t w, t),$$

where ϵ is a parameter to be estimated and w is a unit vector.

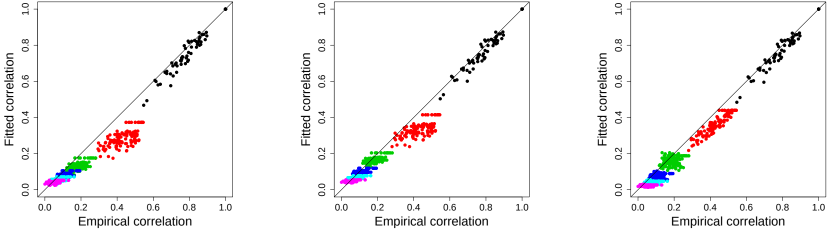
- As the asymmetries in this example are mainly functions of differences in longitude, we take $w = (0, 1)$ as suggested by [Stein, 2005].
- In our framework, this is equivalent to replacing the variogram $\gamma_1(s)$ by $\gamma_1(s - \epsilon t w)$.

MODEL COMPARISON

TABLE: Natural logarithm of the Bayes factor in favour of the asymmetric Model 1 with free λ_0. Bayes factors were calculated using Newton-Raftery ($d = 0.01$), Bridge-sampling and Shifted gamma ($\lambda = 0.98$) estimators for the marginal likelihood. $log(BF) > 5$ suggests strong evidence.

	Newton-Raftery	Bridge sampling	Shifted gamma
Asym. Model 1 $\lambda_0 = 0$	149	153	148
Nonseparable Model 1	166	159	162
Separable Model 1	215	205	212
Nonseparable Model 2	223	227	205
Nonseparable Model 3	172	168	169
Model of Gneiting et al.	206	212	204

EMPIRICAL FIT



(a) Separable model (b) Nonseparable model (c) Asymmetric model

Plot: Posterior median of the correlation function against empirical correlation at temporal lags zero until five, with black corresponding to lag 0, red to lag 1, green to lag 2, dark blue to lag 3, light blue to lag 4 and pink to lag 5.

NONGAUSSIANITY

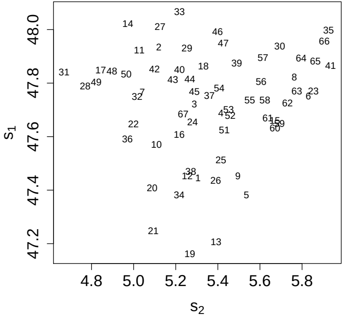
- Now I exemplify how to extend the proposed nonseparable models to accommodate nongaussianity;
- This is a problem of interest in many fields of science such as geology, hydrology and meteorology where extreme events and heterogeneity is often observed;
- I consider the approach of Palacios and Steel [2006] used in spatial data in order to account for nongaussian tail behaviour;

Outliers

- The models will account for individual outliers and regions in space with larger observational variance.
- The latter is quite common in meteorological applications where outliers are often associated with severe weather events such as tornados and hurricanes.
- These events do not usually happen in a single location but cover an extended region.

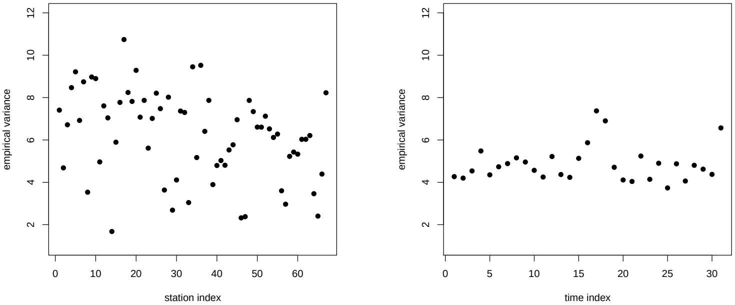
SPATIOTEMPORAL DATA - EXAMPLE

- Maximum temperature data - Spanish Basque Country (67 stations)



EXAMPLE

Maximum temperature data - Spanish Basque Country



MIXING IN SPACE AND TIME

We consider the process

$$\tilde{Z}(s,t;U,V) = \tilde{Z}_1(s;U)\tilde{Z}_2(t;V), \tag{5}$$

MIXING IN SPACE

$$\tilde{Z}_1(s;U) = \sqrt{1-\tau^2}\frac{Z_1(s;U)}{\sqrt{\lambda_1(s)}} + \tau\frac{\epsilon(s)}{\sqrt{h(s)}} \tag{6}$$

MIXING IN TIME

$$\tilde{Z}_2(t;V) = \frac{Z_2(t;V)}{\sqrt{\lambda_2(t)}} \tag{7}$$

MIXING IN SPACE AND TIME

We consider the process

$$\tilde{Z}(s,t;U,V) = \tilde{Z}_1(s;U)\tilde{Z}_2(t;V), \tag{5}$$

MIXING IN SPACE

$$\tilde{Z}_1(s;U) = \sqrt{1-\tau^2}\frac{Z_1(s;U)}{\sqrt{\lambda_1(s)}} + \tau\frac{\epsilon(s)}{\sqrt{h(s)}} \tag{6}$$

MIXING IN TIME

$$\tilde{Z}_2(t;V) = \frac{Z_2(t;V)}{\sqrt{\lambda_2(t)}} \tag{7}$$

Mixing in space and time

We consider the process

$$\tilde{Z}(s, t; U, V) = \tilde{Z}_1(s; U)\tilde{Z}_2(t; V), \tag{5}$$

Mixing in space

$$\tilde{Z}_1(s; U) = \sqrt{1 - \tau^2}\frac{Z_1(s; U)}{\sqrt{\lambda_1(s)}} + \tau\frac{\epsilon(s)}{\sqrt{h(s)}} \tag{6}$$

Mixing in time

$$\tilde{Z}_2(t; V) = \frac{Z_2(t; V)}{\sqrt{\lambda_2(t)}} \tag{7}$$

PROCESS $\lambda_1(s)$

MIXING IN SPACE

$$\tilde{Z}_1(s;U) = \sqrt{1-\tau^2}\frac{Z_1(s;U)}{\sqrt{\lambda_1(s)}} + \tau\frac{\epsilon(s)}{\sqrt{h(s)}}$$

- $\lambda_1(s)$ accounts for regions in space with larger observational variance.
- $\tilde{Z}$ is multivariate Gaussian with covariance matrix

$$\text{Cov}(\tilde{Z}_{ij},\tilde{Z}_{i'j'}) = \sigma^2 M_0(-\gamma_1-\gamma_2)\left[(1-\tau^2)\frac{M_1(-\gamma_1)}{\sqrt{\lambda_{1i}\lambda_{1i'}}} + \tau^2\frac{I(s_i=s_{i'})}{\sqrt{h_i h_{i'}}}\right] M_2(-\gamma_2), \quad (8)$$

  where $\lambda_{1i} = \lambda_1(s_i)$.
- Note that Gaussian behaviour is only assumed given λ_1 and h. Integrating out with respect to these mixing variables leads to non-Gaussian distributions.

Process $\lambda_1(s)$

Mixing in space

$$\tilde{Z}_1(s;U) = \sqrt{1-\tau^2}\frac{Z_1(s;U)}{\sqrt{\lambda_1(s)}} + \tau\frac{\epsilon(s)}{\sqrt{h(s)}}$$

- $\lambda_1(s)$ accounts for regions in space with larger observational variance.
- $\tilde{Z}$ is multivariate Gaussian with covariance matrix

$$\mathrm{Cov}(\tilde{Z}_{ij}, \tilde{Z}_{i'j'}) = \sigma^2 M_0(-\gamma_1-\gamma_2)\left[(1-\tau^2)\frac{M_1(-\gamma_1)}{\sqrt{\lambda_{1i}\lambda_{1i'}}} + \tau^2\frac{I(s_i = s_{i'})}{\sqrt{h_i h_{i'}}}\right] M_2(-\gamma_2), \quad (8)$$

  where $\lambda_{1i} = \lambda_1(s_i)$.
- Note that Gaussian behaviour is only assumed given λ_1 and h. Integrating out with respect to these mixing variables leads to non-Gaussian distributions.

PROCESS $\lambda_1(s)$

MIXING IN SPACE

$$\tilde{Z}_1(s;U) = \sqrt{1-\tau^2}\frac{Z_1(s;U)}{\sqrt{\lambda_1(s)}} + \tau\frac{\epsilon(s)}{\sqrt{h(s)}}$$

- $\lambda_1(s)$ accounts for regions in space with larger observational variance.
- $\tilde{Z}$ is multivariate Gaussian with covariance matrix

$$\mathrm{Cov}(\tilde{Z}_{ij}, \tilde{Z}_{i'j'}) = \sigma^2 M_0(-\gamma_1-\gamma_2)\left[(1-\tau^2)\frac{M_1(-\gamma_1)}{\sqrt{\lambda_{1i}\lambda_{1i'}}} + \tau^2\frac{I(s_i = s_{i'})}{\sqrt{h_i h_{i'}}}\right] M_2(-\gamma_2), \quad (8)$$

where $\lambda_{1i} = \lambda_1(s_i)$.

- Note that Gaussian behaviour is only assumed given λ_1 and h. Integrating out with respect to these mixing variables leads to non-Gaussian distributions.

PROCESS $\lambda_1(s)$

MIXING IN SPACE

$$\tilde{Z}_1(s;U) = \sqrt{1-\tau^2}\frac{Z_1(s;U)}{\sqrt{\lambda_1(s)}} + \tau\frac{\epsilon(s)}{\sqrt{h(s)}}$$

- $\lambda_1(s)$ accounts for regions in space with larger observational variance.
- $\tilde{Z}$ is multivariate Gaussian with covariance matrix

$$\text{Cov}(\tilde{Z}_{ij}, \tilde{Z}_{i'j'}) = \sigma^2 M_0(-\gamma_1 - \gamma_2)\left[(1-\tau^2)\frac{M_1(-\gamma_1)}{\sqrt{\lambda_{1i}\lambda_{1i'}}} + \tau^2\frac{I(s_i = s_{i'})}{\sqrt{h_i h_{i'}}}\right] M_2(-\gamma_2), \quad (8)$$

  where $\lambda_{1i} = \lambda_1(s_i)$.
- Note that Gaussian behaviour is only assumed given λ_1 and h. Integrating out with respect to these mixing variables leads to non-Gaussian distributions.

PROCESS $\lambda_1(s)$

- Scale mixing introduces a potential problem with the continuity of the resulting random function Z.
- Thus $\lambda_1(s)$ needs to be correlated to induce m.s. continuity of $\tilde{Z}_1(s;U)$, this is equivalent to $E[\lambda_1^{-1/2}(s_i)\lambda_1^{-1/2}(s_{i'})] \rightarrow E[\lambda_1^{-1}(s_i)]$ as $s_i \rightarrow s_{i'}$.
- Example: $\lambda_1(s) = \lambda, \forall s \Rightarrow$ Student-t process.
- But this does not account for regions with larger variance.
- We want to account for different variances in different regions.
- Solution: glg process where $\{ln(\lambda_1(s)); s \in D\}$ is a gaussian process with mean $-\frac{\nu}{2}$ and covariance structure $\nu C_1(.)$. [Palacios and Steel, 2006]

PROCESS $\lambda_1(s)$

- Scale mixing introduces a potential problem with the continuity of the resulting random function Z.
- Thus $\lambda_1(s)$ needs to be correlated to induce m.s. continuity of $\tilde{Z}_1(s; U)$, this is equivalent to $E[\lambda_1^{-1/2}(s_i)\lambda_1^{-1/2}(s_{i'})] \to E[\lambda_1^{-1}(s_i)]$ as $s_i \to s_{i'}$.
- Example: $\lambda_1(s) = \lambda, \forall s \Rightarrow$ Student-t process.
- But this does not account for regions with larger variance.
- We want to account for different variances in different regions.
- Solution: glg process where $\{ln(\lambda_1(s)); s \in D\}$ is a gaussian process with mean $-\frac{\nu}{2}$ and covariance structure $\nu C_1(.)$.
[Palacios and Steel, 2006]

PROCESS $\lambda_1(s)$

- Scale mixing introduces a potential problem with the continuity of the resulting random function Z.
- Thus $\lambda_1(s)$ needs to be correlated to induce m.s. continuity of $\tilde{Z}_1(s; U)$, this is equivalent to $E[\lambda_1^{-1/2}(s_i)\lambda_1^{-1/2}(s_{i'})] \rightarrow E[\lambda_1^{-1}(s_i)]$ as $s_i \rightarrow s_{i'}$.
- Example: $\lambda_1(s) = \lambda, \forall s \Rightarrow$ Student-t process.
- But this does not account for regions with larger variance.
- We want to account for different variances in different regions.
- Solution: glg process where $\{ln(\lambda_1(s)); s \in D\}$ is a gaussian process with mean $-\frac{\nu}{2}$ and covariance structure $\nu C_1(.)$. [Palacios and Steel, 2006]

PROCESS $\lambda_1(s)$

- Scale mixing introduces a potential problem with the continuity of the resulting random function Z.
- Thus $\lambda_1(s)$ needs to be correlated to induce m.s. continuity of $\tilde{Z}_1(s;U)$, this is equivalent to $E[\lambda_1^{-1/2}(s_i)\lambda_1^{-1/2}(s_{i'})] \to E[\lambda_1^{-1}(s_i)]$ as $s_i \to s_{i'}$.
- Example: $\lambda_1(s) = \lambda, \forall s \Rightarrow$ Student-t process.
- But this does not account for regions with larger variance.
- We want to account for different variances in different regions.
- Solution: glg process where $\{ln(\lambda_1(s)); s \in D\}$ is a gaussian process with mean $-\frac{\nu}{2}$ and covariance structure $\nu C_1(.)$.
  [Palacios and Steel, 2006]

PROCESS $h(s)$

MIXING IN SPACE

$$\tilde{Z}_1(s;U) = \sqrt{1-\tau^2}\frac{Z_1(s;U)}{\sqrt{\lambda_1(s)}} + \tau\frac{\epsilon(s)}{\sqrt{h(s)}}$$

- $h(s)$ accounts for traditional outliers (different nugget effects).
- We consider the detection of outliers jointly in the estimation procedure and the variable $h_i = h(s_i), i = 1, \ldots, I$ are considered latent variables
- Their posterior distribution indicate outlying observations (h_i close to 0).
- We consider

Process $h(s)$

Mixing in space

$$\tilde{Z}_1(s;U) = \sqrt{1-\tau^2}\frac{Z_1(s;U)}{\sqrt{\lambda_1(s)}} + \tau\frac{\epsilon(s)}{\sqrt{h(s)}}$$

- $h(s)$ accounts for traditional outliers (different nugget effects).
- We consider the detection of outliers jointly in the estimation procedure and the variable $h_i = h(s_i), i = 1, \dots, I$ are considered latent variables
- Their posterior distribution indicate outlying observations (h_i close to 0).
- We consider
  - $log(h_i) \sim \mathrm{N}(-\nu_h/2, \nu_h)$.
  - $h_i \sim \mathrm{Ga}(1/\nu_h, 1/\nu_h)$.

PROCESS $h(s)$

MIXING IN SPACE

$$\tilde{Z}_1(s;U) = \sqrt{1-\tau^2}\frac{Z_1(s;U)}{\sqrt{\lambda_1(s)}} + \tau\frac{\epsilon(s)}{\sqrt{h(s)}}$$

- $h(s)$ accounts for traditional outliers (different nugget effects).
- We consider the detection of outliers jointly in the estimation procedure and the variable $h_i = h(s_i), i = 1, \dots, I$ are considered latent variables
- Their posterior distribution indicate outlying observations (h_i close to 0).
- We consider
  - $\log(h_i) \sim N(-\nu_h/2, \nu_h)$.
  - $h_i \sim Ga(1/\nu_h, 1/\nu_h)$.

PROCESS $h(s)$

MIXING IN SPACE

$$\tilde{Z}_1(s;U) = \sqrt{1-\tau^2}\frac{Z_1(s;U)}{\sqrt{\lambda_1(s)}} + \tau\frac{\epsilon(s)}{\sqrt{h(s)}}$$

- $h(s)$ accounts for traditional outliers (different nugget effects).
- We consider the detection of outliers jointly in the estimation procedure and the variable $h_i = h(s_i), i = 1, \ldots, I$ are considered latent variables
- Their posterior distribution indicate outlying observations (h_i close to 0).
- We consider
  - $log(h_i) \sim \mathrm{N}(-\nu_h/2, \nu_h)$.
  - $h_i \sim \mathrm{Ga}(1/\nu_h, 1/\nu_h)$.

PROCESS $\lambda_2(t)$

MIXING IN TIME

$$\tilde{Z}_2(t; V) = \frac{Z_2(t; V)}{\sqrt{\lambda_2(t)}}$$

- $\lambda_2(t)$ accounts for sections in time with larger observational variance.
- This can be seen as a way to address the issue of volatility clustering, which is common in finantial time series data.
- We consider the log gaussian process where $\{ln(\lambda_2(t)); t \in T\}$ is a gaussian process with mean $-\frac{\nu_2}{2}$ and covariance structure $\nu_2 C_2(.)$.

PROCESS $\lambda_2(t)$

MIXING IN TIME

$$\tilde{Z}_2(t;V) = \frac{Z_2(t;V)}{\sqrt{\lambda_2(t)}}$$

- $\lambda_2(t)$ accounts for sections in time with larger observational variance.
- This can be seen as a way to address the issue of volatility clustering, which is common in finantial time series data.
- We consider the log gaussian process where $\{ln(\lambda_2(t)); t \in T\}$ is a gaussian process with mean $-\frac{\nu_2}{2}$ and covariance structure $\nu_2 C_2(.)$.

PROCESS $\lambda_2(t)$

MIXING IN TIME

$$\tilde{Z}_2(t; V) = \frac{Z_2(t; V)}{\sqrt{\lambda_2(t)}}$$

- $\lambda_2(t)$ accounts for sections in time with larger observational variance.
- This can be seen as a way to address the issue of volatility clustering, which is common in finantial time series data.
- We consider the log gaussian process where $\{ln(\lambda_2(t)); t \in T\}$ is a gaussian process with mean $-\frac{\nu_2}{2}$ and covariance structure $\nu_2 C_2(.)$.

PREDICTIONS

- $(\lambda_{1i}, h_i, \lambda_{2j})$ are considered latent variables and sampled in our MCMC sampler.
- Given $(\lambda_{1i}, h_i, \lambda_{2j})$ the process is gaussian and we can predict at unobserved locations and time points.
- We compare the predictive performance using proper scoring rules [Gneiting and Raftery, 2008]:
  - $LPS(p, x) = -log(p(x))$
  - $IS(q_1, q_2; x) = (q_2 - q_1) + \frac{2}{\xi}(q_1 - x)I(x < q_1) + \frac{2}{\xi}(x - q_2)I(x > q_2)$. We use $\xi = 0.05$ resulting in a 95% credible interval.

PREDICTIONS

- $(\lambda_{1i}, h_i, \lambda_{2j})$ are considered latent variables and sampled in our MCMC sampler.
- Given $(\lambda_{1i}, h_i, \lambda_{2j})$ the process is gaussian and we can predict at unobserved locations and time points.
- We compare the predictive performance using proper scoring rules [Gneiting and Raftery, 2008]:
  - $LPS(p,x) = -log(p(x))$
  - $IS(q_1, q_2; x) = (q_2 - q_1) + \frac{2}{\xi}(q_1 - x)I(x < q_1) + \frac{2}{\xi}(x - q_2)I(x > q_2)$. We use $\xi = 0.05$ resulting in a 95% credible interval.

PREDICTIONS

- $(\lambda_{1i}, h_i, \lambda_{2j})$ are considered latent variables and sampled in our MCMC sampler.
- Given $(\lambda_{1i}, h_i, \lambda_{2j})$ the process is gaussian and we can predict at unobserved locations and time points.
- We compare the predictive performance using proper scoring rules [Gneiting and Raftery, 2008]:
  - $LPS(p, x) = -log(p(x))$
  - $IS(q_1, q_2; x) = (q_2 - q_1) + \frac{2}{\xi}(q_1 - x)I(x < q_1) + \frac{2}{\xi}(x - q_2)I(x > q_2)$. We use $\xi = 0.05$ resulting in a 95% credible interval.

PREDICTIONS

- $(\lambda_{1i}, h_i, \lambda_{2j})$ are considered latent variables and sampled in our MCMC sampler.
- Given $(\lambda_{1i}, h_i, \lambda_{2j})$ the process is gaussian and we can predict at unobserved locations and time points.
- We compare the predictive performance using proper scoring rules [Gneiting and Raftery, 2008]:
  - $LPS(p, x) = -log(p(x))$
  - $IS(q_1, q_2; x) = (q_2 - q_1) + \frac{2}{\xi}(q_1 - x)I(x < q_1) + \frac{2}{\xi}(x - q_2)I(x > q_2)$. We use $\xi = 0.05$ resulting in a 95% credible interval.

SIMULATED EXAMPLE: DATA CONTAMINATION

- This data set has $I = 30$ locations and $J = 30$ time points generated from a Gaussian model with no nugget effect ($\tau^2 = 0$).
- The covariance model is nonseparable Cauchy ($X_i \sim \text{Ga}(\lambda_i, 1)$, $i = 0, 1, 2$) in space and time with $c = 0.5$.
- We contaminated this data set with different kinds of "outliers" in order to see the performance of the proposed models in each situation.

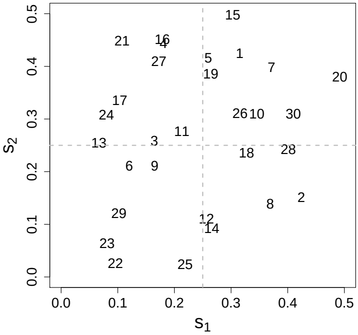
SIMULATED EXAMPLE: DATA CONTAMINATION

- This data set has $I = 30$ locations and $J = 30$ time points generated from a Gaussian model with no nugget effect ($\tau^2 = 0$).
- The covariance model is nonseparable Cauchy ($X_i \sim \text{Ga}(\lambda_i, 1)$, $i = 0, 1, 2$) in space and time with $c = 0.5$.
- We contaminated this data set with different kinds of "outliers" in order to see the performance of the proposed models in each situation.

SIMULATED EXAMPLE: DATA CONTAMINATION

- This data set has $I = 30$ locations and $J = 30$ time points generated from a Gaussian model with no nugget effect ($\tau^2 = 0$).
- The covariance model is nonseparable Cauchy ($X_i \sim \text{Ga}(\lambda_i, 1)$, $i = 0, 1, 2$) in space and time with $c = 0.5$.
- We contaminated this data set with different kinds of ”outliers” in order to see the performance of the proposed models in each situation.

SPATIAL DOMAIN

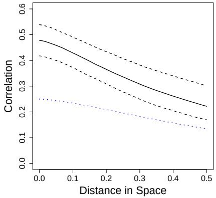


- The proposal for $\lambda_{1i}, h_i,\ i = 1, \ldots, I$ in the MCMC sampler is constructed by dividing the observations in blocks defined by position in the spatial domain.

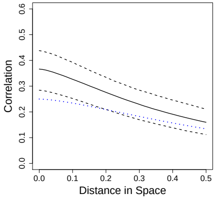
DATA 1 (TRADITIONAL OUTLIER) - DESCRIPTION

One location was selected at random (location 7) and a random increment from Unif$(1.0, 1.5)$ times the standard deviation was added to each observation for this location for the first 20 time points.

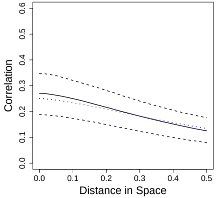
ESTIMATED CORRELATION FUNCTION - $t_0 = 1$



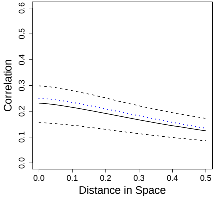
(a) Gaussian



(b) Nongaussian with λ_1

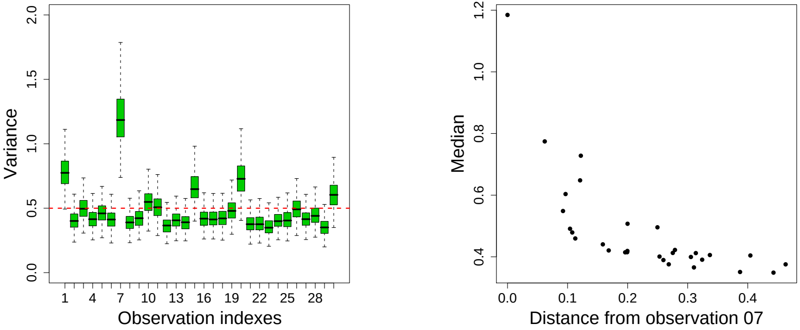


(c) Nongaussian with h and λ_1



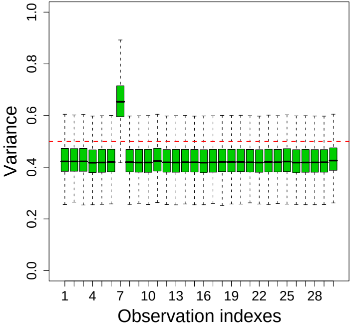
(d) Gaussian (Uncontaminated data)

NONGAUSSIAN MODEL WITH λ_1

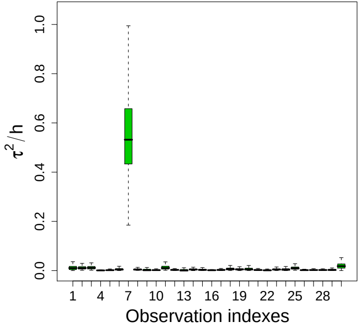


(a) Variance for each location. (b) Median of σ_i^2 vs. distance from obs. 7.

Nongaussian model with h (lognormal)

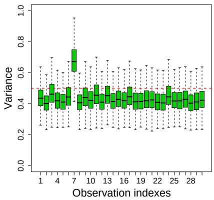


(a) Variance for each location.

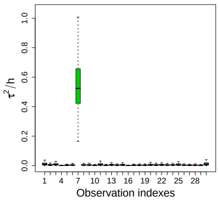


(b) Nugget for each location.

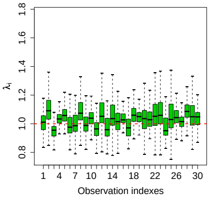
NONGAUSSIAN MODEL WITH λ_1 AND h



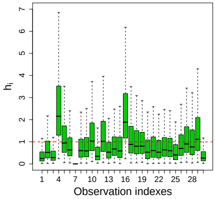
(a) Variance for each location.



(b) Nugget for each location.



(c) λ_{1i}, $i = 1, \ldots, 30$.



(d) h_i, $i = 1, \ldots, 30$.

TEMPERATURE DATA - MODEL

- Mean function:

$$\mu(s,t) = \delta_0 + \delta_1 s_1 + \delta_2 s_2 + \delta_3 h + \delta_4 t + \delta_5 t^2$$

- $\tilde{Z}$ is multivariate Gaussian with covariance matrix

$$\mathrm{Cov}(\tilde{Z}_{ij}, \tilde{Z}_{i'j'}) = \sigma^2 M_0(-\gamma_1 - \gamma_2)\left[(1-\tau^2)\frac{M_1(-\gamma_1)}{\sqrt{\lambda_{1i}\lambda_{1i'}}} + \tau^2 \frac{I(s_i = s_{i'})}{\sqrt{h_i h_{i'}}}\right] M_2(-\gamma_2), \quad (9)$$

  where $\lambda_{1i} = \lambda_1(s_i)$.

- M_0, M_1 and M_2 are Cauchy covariance functions.

TEMPERATURE DATA - MODEL

- Mean function:

$$\mu(s,t) = \delta_0 + \delta_1 s_1 + \delta_2 s_2 + \delta_3 h + \delta_4 t + \delta_5 t^2$$

- $\tilde{Z}$ is multivariate Gaussian with covariance matrix

$$\mathrm{Cov}(\tilde{Z}_{ij}, \tilde{Z}_{i'j'}) = \sigma^2 M_0(-\gamma_1 - \gamma_2)\left[(1-\tau^2)\frac{M_1(-\gamma_1)}{\sqrt{\lambda_{1i}\lambda_{1i'}}} + \tau^2 \frac{I(s_i = s_{i'})}{\sqrt{h_i h_{i'}}}\right] M_2(-\gamma_2), \quad (9)$$

where $\lambda_{1i} = \lambda_1(s_i)$.

- M_0, M_1 and M_2 are Cauchy covariance functions.

LIKELIHOOD

- In order to calculate the likelihood function we need to invert a matrix with dimension 2077×2077.
- We approximate the likelihood by using conditional distributions.
- We consider a partition of Z into subvectors $Z_1, ..., Z_{31}$ where $Z_j = (Z(s_1, t_j), \ldots, Z(s_{67}, t_j))'$ and we define $Z_{(j)} = (Z_{j-L+1}, ..., Z_j)$. Then

$$p(z|\phi) \approx p(z_1|\phi) \prod_{j=2}^{31} p(z_j|z_{(j-1)}, \phi). \tag{10}$$

- This means the distribution of Z_j will only depend on the observations in space for the previous L time points.
- In this application we used $L = 5$ to make the MCMC feasible.

BAYES FACTOR

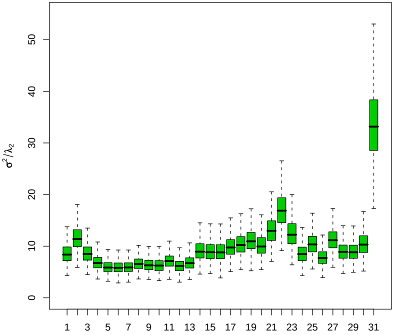
	h	λ_1	λ_1 & h	λ_2	λ_2 & h	λ_1 & λ_2	λ_1, h & λ_2
Shifted gamma	172	148	345	138	279	417	547

TABLE: The natural logarithm of the Bayes factor in favor of the model in the column versus Gaussian model using Shifted-Gamma ($\lambda = 0.98$) estimator for the predictive density of z.

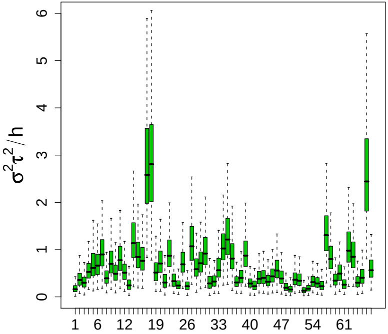
MODEL COMPARISON

model	Average width	$\bar{IS}$	LPS
Gaussian	3.78	4.35	97.25
h	3.83	4.34	112.56
λ_1	3.74	4.36	107.43
λ_1 & h	3.75	4.48	117.20
λ_2	3.73	3.94	76.73
λ_2 & h	3.73	3.87	77.60
λ_1 & λ_2	4.51	4.65	96.35
λ_1, h & λ_2	3.84	4.02	90.30

MODEL WITH h AND λ_2

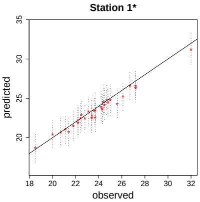


(a) $\sigma^2(1-\tau^2)/\lambda_2$.

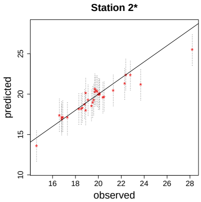


(b) $\sigma^2\tau^2/h$.

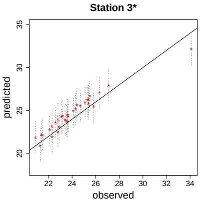
PREDICTED TEMPERATURE AT THE OUT-OF-SAMPLE STATIONS



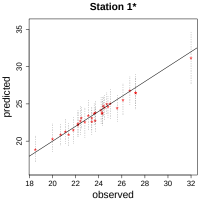
(a) Gaussian Model.



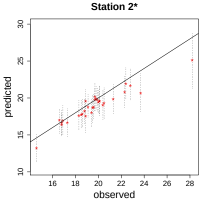
(b) Gaussian Model.



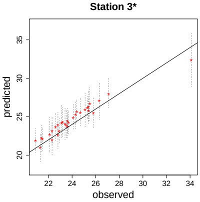
(c) Gaussian Model.



(d) Model with λ_2 & h.



(e) Model with λ_2 & h.



(f) Model with λ_2 & h.

CONCLUSIONS AND FUTURE WORK

- The resulting model has very useful theoretical properties [Fonseca and Steel, 2011];
- For practical modelling purposes, I suggest a number of different parameterisations, leading to a variety of special cases;
- The examples clearly show the overwhelming data support for our proposed covariance functions.
- We have extended the model to accommodate nongaussianity, nonstationarity (not presented here) and intend to extend it to multivariate processes.

CONCLUSIONS AND FUTURE WORK

- The resulting model has very useful theoretical properties [Fonseca and Steel, 2011];
- For practical modelling purposes, I suggest a number of different parameterisations, leading to a variety of special cases;
- The examples clearly show the overwhelming data support for our proposed covariance functions.
- We have extended the model to accommodate nongaussianity, nonstationarity (not presented here) and intend to extend it to multivariate processes.

CONCLUSIONS AND FUTURE WORK

- The resulting model has very useful theoretical properties [Fonseca and Steel, 2011];
- For practical modelling purposes, I suggest a number of different parameterisations, leading to a variety of special cases;
- The examples clearly show the overwhelming data support for our proposed covariance functions.
- We have extended the model to accommodate nongaussianity, nonstationarity (not presented here) and intend to extend it to multivariate processes.

CONCLUSIONS AND FUTURE WORK

- The resulting model has very useful theoretical properties [Fonseca and Steel, 2011];
- For practical modelling purposes, I suggest a number of different parameterisations, leading to a variety of special cases;
- The examples clearly show the overwhelming data support for our proposed covariance functions.
- We have extended the model to accommodate nongaussianity, nonstationarity (not presented here) and intend to extend it to multivariate processes.

REFERENCES

N Cressie and H-C Huang
Classes of Nonseparable, Spatio-Temporal Stationary Covariance Functions
Journal of the American Statistical Association. (94) 1330–1340, 1999.

T Fonseca and M F J Steel
A General Class of Nonseparable Space-time Covariance Models
Environmetrics. (22) 224–242, 2011.

M Fuentes and J M Davis and G M Lackmann
Modeling and predicting complex space-time structures and patterns of coastal wind fields.
Environmetrics. (16) 449–464, 2005.

T Gneiting
Nonseparable, Stationary Covariance Functions for Space-Time Data
Journal of the American Statistical Association. (97) 590–600, 2002.

T Gneiting and A E Raftery
Strictly proper scoring rules, prediction and estimation
JASA. (102) 360–378, 2007.

J Haslett and A E Raftery
Space-Time Modeling With Long-Memory Dependence: Assessing Ireland's Wind-Power Resource
Applied Statistics. (38) 1–50, 1989.

C Ma
Spatio-temporal covariance functions generated by mixtures
Mathematical geology. (34) 965–975, 2002.

M B Palacios and M F J Steel
Non-Gaussian Bayesian Geostatistical Modeling
Journal of the American Statistical Association. (101) 604–618, 2006.

M L Stein
Space-Time Covariance Functions
Journal of the American Statistical Association. (100) 310–321, 2005.