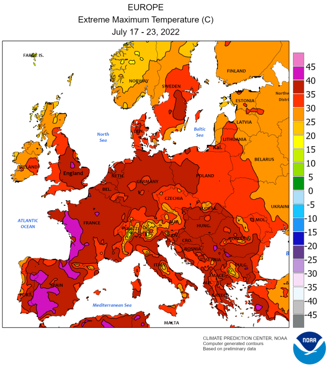
Spatio-temporal weather generator for the temperature over France

Présentation au séminaire de statistiques de Rochebrune

Caroline Cognot (MIA-PS & EDF)
Supervisors : Liliane Bel (MIA-PS) et Sylvie Parey (EDF)

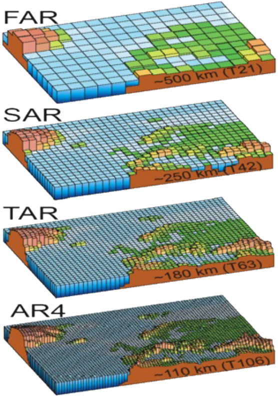
Electricity and climate studies

- Climate extremes, risks : impact on agriculture, health, energy production and demand
- Climate change : impacts frequency and intensity of spatial = multivariate meteorological hazards
- In particular, both on electricity generation and system balance



Studies are necessary and we use models to do this.

Physical climate modelling



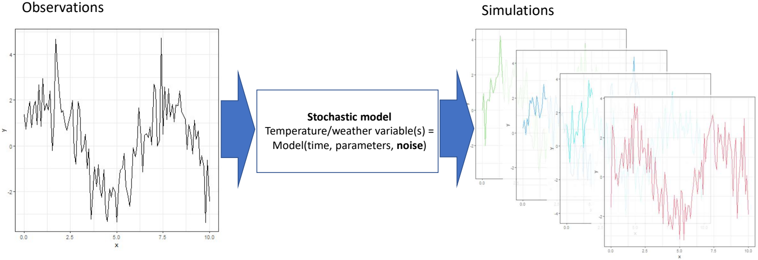
Climate models grid sizes

- Used in RCP modelling (regional models)
- Complex phenomena, well-known from physics

BUT

- Computationally expensive
- Not adapted to extremes

Why stochastic modelling



- Simulations = same statistical properties as the observations
- Reproduce the properties important to the user (e.g. heat waves)
- Computationally efficient
- Can be used to sample the distribution, including extremes
- Can be used with real data or debiased climate model output

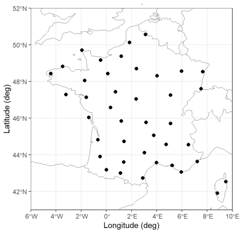
"Stochastic Weather Generator"

A temperature generator

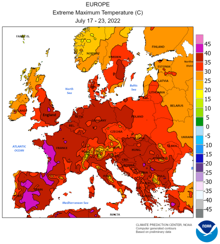
Existing work :

- Single-site models : models for $n \geq 1$ variables $X^1(t), ..., X^n(t)$
- Multi-site models, but viewing sites as many different variables : models for $n \geq 1$ variables at $n_S \geq 1$ sites $X_1^1(t), ..., X_1^n(t), ..., X_{n_S}^1(t), ..., X_{n_S}^n(t)$
- Spatial models, focused mostly on precipitations : models for $n \geq 1$ variables on a spatial extent $\{s \in \mathcal{D}\}$ $X^1(s,t), ..., X^n(s,t)$

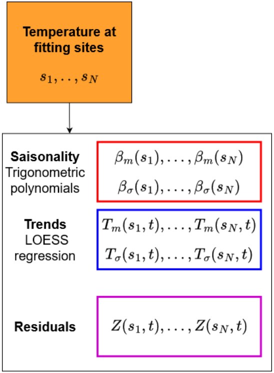
My objective :



Build a SWG for the temperature, reproducing the spatial and temporal structure and allowing for sampling in the climate variability.



Decomposition in trend and seasonality[1]



Objective : separate deterministic from stochastic components. Deterministic = climate, stochastic = climate variability.
For each site s and time t, we decompose the temperature $X(s,t)$:

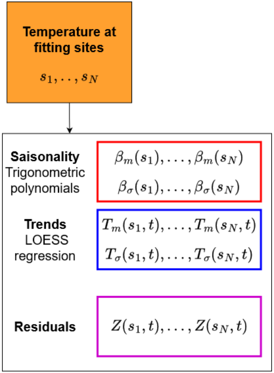
$$X(s,t) = T_m(s,t) + S_m(s,t) + T_\sigma(s,t) S_\sigma(s,t) Z(s,t)$$

where

- $T_m(s,.)$ is the long-term mean trend;
- $T_\sigma(s,.)$ is the long-term variance trend ;

[1] Hoang 2010.

Decomposition in trend and seasonality[1]



Objective : separate deterministic from stochastic components. Deterministic = climate, stochastic = climate variability.
For each site s and time t, we decompose the temperature $X(s,t)$:

$$X(s,t) = T_m(s,t) + S_m(s,t) + T_\sigma(s,t) S_\sigma(s,t) Z(s,t)$$

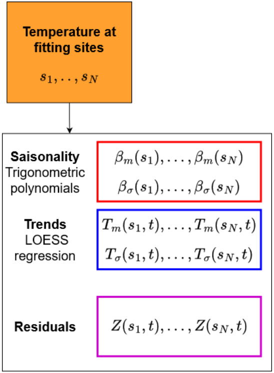
where

- $T_m(s,.)$ is the long-term mean trend;
- $T_\sigma(s,.)$ is the long-term variance trend ;
- $S_m(s,.)$ is the mean seasonality;
- $S_\sigma(s,.)$ is the variance seasonality;

$$S_{m \text{ or } \sigma}(s,t) = \beta_1(s) + \sum_{i=1}^{d_{m \text{ or } \sigma}} \left[\beta_{2i}(s) cos\left(\frac{2i\pi t}{365}\right) + \beta_{2i+1}(s) sin\left(\frac{2i\pi t}{365}\right) \right]$$

[1] Hoang 2010.

Decomposition in trend and seasonality[1]



Objective : separate deterministic from stochastic components. Deterministic = climate, stochastic = climate variability.
For each site s and time t, we decompose the temperature $X(s,t)$:

$$X(s,t) = T_m(s,t) + S_m(s,t) + T_\sigma(s,t) S_\sigma(s,t) Z(s,t)$$

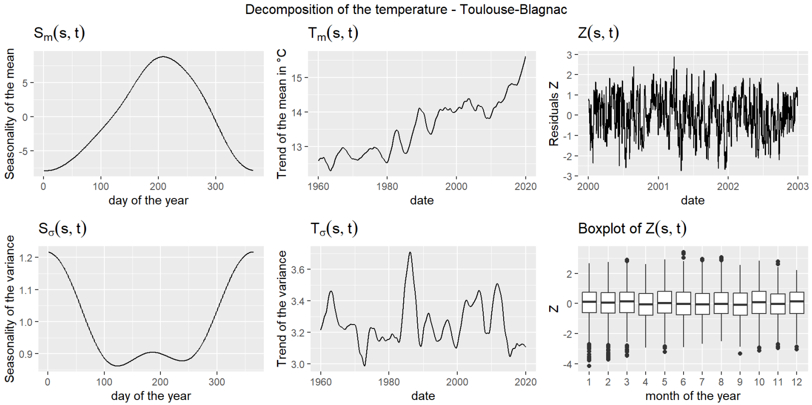
where

- $T_m(s,.)$ is the long-term mean trend;
- $T_\sigma(s,.)$ is the long-term variance trend ;
- $S_m(s,.)$ is the mean seasonality;
- $S_\sigma(s,.)$ is the variance seasonality;
- $Z(s,.)$ are the residuals at site s. They have variance 1 and mean 0.

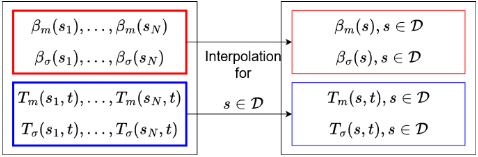
[1] Hoang 2010.

Example : s fixed, Toulouse

$$X(s,t) = T_m(s,t) + S_m(s,t) + T_\sigma(s,t)S_\sigma(s,t)Z(s,t)$$

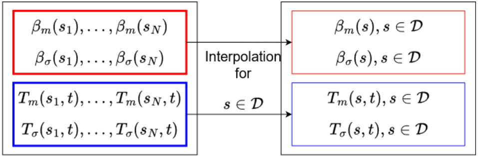


Decomposition : spatial extension



Objective :
Extend decomposition parameters from discrete to continuous
$\mathcal{D}$ is a high resolution grid

Decomposition : spatial extension

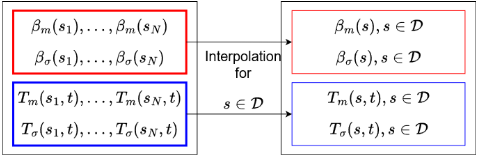


Objective :
Extend decomposition parameters from discrete to continuous
$\mathcal{D}$ is a high resolution grid

- For trends : Trends are "smooth" in time and slow-varying : reasonable to use constant weights, using Inverse Distance Weighting (IDW) interpolation $\longrightarrow$ gives many maps of slow-varying mean temperature and standard deviation.

$$u(s) = \frac{\sum_{i=1}^{n} w_i(s) u_i}{\sum_{i=1}^{n} w_i(s)}, \text{with } w_i(s) = \frac{1}{(d(s_i, s))^p}$$

Decomposition : spatial extension



Objective :
Extend decomposition parameters from discrete to continuous
$\mathcal{D}$ is a high resolution grid

- For trends : Trends are "smooth" in time and slow-varying : reasonable to use constant weights, using Inverse Distance Weighting (IDW) interpolation $\longrightarrow$ gives many maps of slow-varying mean temperature and standard deviation.
- For seasonality : kriging on the coefficients $\longrightarrow$ 2 maps of spatial coefficients, to multiply by corresponding sines and cosines to obtain the cycle.

Model of the residuals

Recall : decomposition

$$X(s,t) = T_m(s,t) + S_m(s,t) + T_\sigma(s,t)S_\sigma(s,t)\mathbf{Z(s,t)}$$

We want to model the stochastic part $Z(s,t)$, which is already supposed to be zero-mean, with variance close to 1. :

Proposal : 2nd-order spatio-temporal model

A 2nd-order spatiotemporal model $Z(\mathbf{s},t)$ is defined by its 2 components :

1. The mean function $\mathbb{E}[Z(\mathbf{s},t)] = \mu(\mathbf{s},t) = 0$ for us
2. The covariance function $C(\mathbf{s}_1,\mathbf{s}_2,t_1,t_2) = Cov[Z(\mathbf{s}_1,t_1),Z(\mathbf{s}_2,t_2)]$

Model of the residuals

Recall : decomposition

$$X(s,t) = T_m(s,t) + S_m(s,t) + T_\sigma(s,t)S_\sigma(s,t)\mathbf{Z(s,t)}$$

We want to model the stochastic part $Z(s,t)$, which is already supposed to be zero-mean, with variance close to 1. :

Proposal : 2nd-order stationnary spatio-temporal model

A 2nd-order stationnary spatiotemporal model $Z(\mathbf{s},t)$ is defined by its 2 components :

1. The mean function $\mathbb{E}[Z(\mathbf{s},t)] = \mu(\mathbf{s},t) = 0$ for us
2. The covariance function $C(\mathbf{s}_1,\mathbf{s}_2,t_1,t_2) = C(\mathbf{s}_2 - \mathbf{s}_1, t_2 - t_1)$

Model of the residuals

We want to model the stochastic part $Z(s,t)$, which is already supposed to be zero-mean, with variance close to 1. :

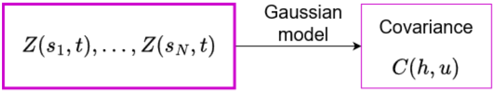
Proposal : 2nd-order stationnary isotropic spatio-temporal model

A 2nd-order stationnary isotropic spatiotemporal model $Z(\mathbf{s},t)$ is defined by its 2 components :

1. The mean function $\mathbb{E}[Z(\mathbf{s},t)] = \mu(\mathbf{s},t) = 0$ for us
2. The covariance function $C(\mathbf{s}_1,\mathbf{s}_2,t_1,t_2) = C(||\mathbf{s}_2 - \mathbf{s}_1||, t_2 - t_1)$

This reduces our problem to fitting a covariance function

$$\begin{aligned} C : \mathbb{R}^{+} \times \mathbb{R} &\longrightarrow \mathbb{R} \\ (h,u) &\longmapsto C(h,u) \end{aligned}$$



Model of the residuals : Possible covariance functions

- Separable space-time functions : $C(h,u) = C_S(h)C_T(u)$:

[2] Gneiting 2002.

Model of the residuals : Possible covariance functions

- Separable space-time functions : $C(h,u) = C_S(h)C_T(u)$: no space-time interactions !

[2] Gneiting 2002.

Model of the residuals : Possible covariance functions

- Separable space-time functions : $C(h,u) = C_S(h)C_T(u)$: no space-time interactions !
- Non - separable functions : many classes of Gneiting type[2].
  For "nice" $\phi(t)$ and $\psi(t)$

$$C(h,u) = \frac{\sigma^2}{\psi(|u|^2)^{d/2}} \phi\left(\frac{||h||^2}{\psi(|u|^2)}\right)$$

is a covariance function.

[2] Gneiting 2002.

Model of the residuals : Possible covariance functions

- Separable space-time functions : $C(h, u) = C_S(h)C_T(u)$: no space-time interactions !
- Non - separable functions : many classes of Gneiting type[2].
  For "nice" $\phi(t)$ and $\psi(t)$

$$C(h, u) = \frac{\sigma^2}{\psi(|u|^2)^{d/2}}\phi\left(\frac{||h||^2}{\psi(|u|^2)}\right)$$

is a covariance function. An usual $\psi(t)$ is of the form $(at^\alpha + 1)^\beta$; $\phi(t)$ can be covariance kernels.

[2] Gneiting 2002.

Model of the residuals : Possible covariance functions

- Separable space-time functions : $C(h, u) = C_S(h)C_T(u)$: no space-time interactions !
- Non - separable functions : many classes of Gneiting type[2].
  For "nice" $\phi(t)$ and $\psi(t)$

$$C(h, u) = \frac{\sigma^2}{\psi(|u|^2)^{d/2}} \phi\left(\frac{||h||^2}{\psi(|u|^2)}\right)$$

  is a covariance function. An usual $\psi(t)$ is of the form $(at^{\alpha} + 1)^{\beta}$; $\phi(t)$ can be covariance kernels.
  Matérn kernel : $C(h, u) = \frac{\sigma^2}{(\alpha u^{2a}+1)^b} \mathcal{M}\left(\frac{h}{\sqrt{\alpha u^{2a}+1}^b}; r; \nu\right)$

[2] Gneiting 2002.

Model of the residuals : Possible covariance functions

- Separable space-time functions : $C(h,u) = C_S(h)C_T(u)$: no space-time interactions !
- Non - separable functions : many classes of Gneiting type[2]. For "nice" $\phi(t)$ and $\psi(t)$

$$C(h,u) = \frac{\sigma^2}{\psi(|u|^2)^{d/2}}\phi\left(\frac{||h||^2}{\psi(|u|^2)}\right)$$

is a covariance function. An usual $\psi(t)$ is of the form $(at^\alpha + 1)^\beta$; $\phi(t)$ can be covariance kernels.

Matérn kernel : $C(h,u) = \frac{\sigma^2}{(\alpha u^{2a}+1)^b}\mathcal{M}\left(\frac{h}{\sqrt{\alpha u^{2a}+1}^b}; r; \nu\right)$

In practice, multiplied by a purely temporal covariance function for the Gneiting-Matérn class :

Gneiting-Matérn covariance function

$$C(h,u) = \frac{\sigma^2}{(\alpha u^{2a}+1)^{b+\delta}}\mathcal{M}\left(\frac{h}{\sqrt{\alpha u^{2a}+1}^b}; r; \nu\right)$$

[2] Gneiting 2002.

Model of the residuals : Parameter estimation

Gneiting-Matérn covariance function

$$C(h,u) = \frac{\sigma^2}{(\alpha u^{2a}+1)^{b+\delta}} \mathcal{M}\left(\frac{h}{\sqrt{\alpha u^{2a}+1}^{b}}; r; \nu\right) \longrightarrow \text{ 7 parameters.}$$

log-likelihood : for $(X1, ..., Xn) \sim \mathcal{N}(0, \Sigma)$

$$\begin{aligned}\ell(x,\theta) &= log(f(X_1 = x_1, ..., X_n = x_n)) \\ &= -\frac{1}{2} log(|\Sigma|) - \frac{1}{2}((x_1 ... x_n)\Sigma^{-1}(x_1 ... x_n)^T)\end{aligned}$$

Model of the residuals : Parameter estimation

Gneiting-Matérn covariance function

$$C(h,u) = \frac{\sigma^2}{(\alpha u^{2a}+1)^{b+\delta}} \mathcal{M}\left(\frac{h}{\sqrt{\alpha u^{2a}+1}^{b}}; r; \nu \right) \longrightarrow \text{ 7 parameters.}$$

~~log-likelihood : for $(X1, ..., Xn) \sim \mathcal{N}(0, \Sigma)$~~

$$\ell(x,\theta) = log(f(X1 = x1, ..., X_n = x_n))$$
$$= -\frac{1}{2} log(|\Sigma|) - \frac{1}{2}((x_1 ... x_n)\Sigma^{-1}(x_1 ... x_n)^T)$$

Here, $n = 41 \times 365$ for a year of data ! Inversion is not recommended.

Model of the residuals : Parameter estimation

Gneiting-Matérn covariance function

$$C(h,u) = \frac{\sigma^2}{(\alpha u^{2a}+1)^{b+\delta}} \mathcal{M}\left(\frac{h}{\sqrt{\alpha u^{2a}+1}^{b}}; r; \nu\right) \longrightarrow \text{ 7 parameters.}$$

log-likelihood : for $(X1, ..., Xn) \sim \mathcal{N}(0, \Sigma)$

$$\begin{aligned} \ell(x,\theta) &= log(f(X1 = x1, ..., X_n = x_n)) \\ &= -\frac{1}{2} log(|\Sigma|) - \frac{1}{2}((x_1 ... x_n)\Sigma^{-1}(x_1 ... x_n)^T) \end{aligned}$$

Here, $n = 41 \times 365$ for a year of data ! Inversion is not recommended.
Composite pairwise log-likelihood :

$$\begin{aligned} \ell_C(x,\theta) &= \sum_{\text{pairs } i,j} log(f(X_i = x_i, X_j = x_j)) \\ &= \sum_{\text{pairs } i,j} (-\frac{1}{2} log \begin{vmatrix} \sigma^2 & \Sigma_{ij} \\ \Sigma_{ij} & \sigma^2 \end{vmatrix} - \frac{1}{2}((x_i, x_j) \begin{pmatrix} \sigma^2 & \Sigma_{ij} \\ \Sigma_{ij} & \sigma^2 \end{pmatrix}^{-1} (x_i, x_j)^T)) \end{aligned}$$

Model of the residuals : Parameter estimation

Gneiting-Matérn covariance function

$$C(h,u) = \frac{\sigma^2}{(\alpha u^{2a}+1)^{b+\delta}}\mathcal{M}\left(\frac{h}{\sqrt{\alpha u^{2a}+1}^{b}};r;\nu\right) \longrightarrow \text{ 7 parameters.}$$

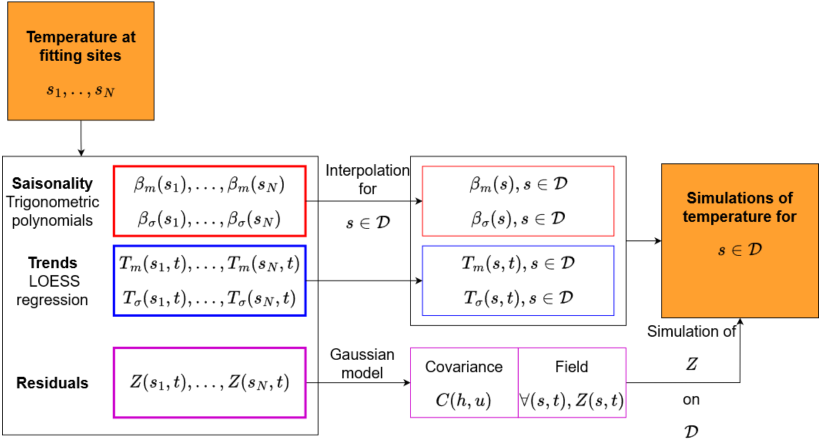
log-likelihood : for $(X1, ..., Xn) \sim \mathcal{N}(0,\Sigma)$

$$\ell(x,\theta) = log(f(X1 = x1, ..., X_n = x_n))$$
$$= -\frac{1}{2}log(|\Sigma|) - \frac{1}{2}((x_1 ... x_n)\Sigma^{-1}(x_1 ... x_n)^T)$$

Here, $n = 41 \times 365$ for a year of data ! Inversion is not recommended.
Composite pairwise log-likelihood :

$$\ell_C(x,\theta) = \sum_{\text{pairs } i,j} log(f(X_i = x_i, X_j = x_j))$$
$$= \sum_{\text{pairs } i,j}(-\frac{1}{2}log\begin{vmatrix}\sigma^2 & \Sigma_{ij}\\ \Sigma_{ij} & \sigma^2\end{vmatrix} - \frac{1}{2}((x_i,x_j)\begin{pmatrix}\sigma^2 & \Sigma_{ij}\\ \Sigma_{ij} & \sigma^2\end{pmatrix}^{-1}(x_i,x_j)^T))$$

Outline of the model



Simulation methods for the Gaussian field

- Naive approach : simulate from multivariate Gaussian covariance matrix

[3] Allard et al. 2020.

Simulation methods for the Gaussian field

- Naive approach : simulate from multivariate Gaussian covariance matrix . Not feasible for large space and time grid : **ok for 40 points in space, but not for 1000**.

Simulation methods for the Gaussian field

- Naive approach : simulate from multivariate Gaussian covariance matrix . Not feasible for large space and time grid : **ok for 40 points in space, but not for 1000**.
- A bit more refined : simulate all spatial points for each day t from previous l days using Gaussian properties and stationnarity : for X_i the vector at all points in space at time i, there are matrices A, B, C such that

$$\begin{bmatrix} X_1 \\ X_2 \\ \dots \\ X_l \end{bmatrix} \sim \mathcal{N}(0, A)$$

$$\forall t > \ell,\ X_t | \begin{bmatrix} X_1 \\ \dots \\ X_{t-1} \end{bmatrix} = X_t \mid \begin{bmatrix} X_{t-\ell} \\ \dots \\ X_{t-1} \end{bmatrix} \sim \mathcal{N}\left(B \begin{bmatrix} X_{t-\ell} \\ \dots \\ X_{t-1} \end{bmatrix}, C \right)$$

[3] Allard et al. 2020.

Simulation methods for the Gaussian field

- Naive approach : simulate from multivariate Gaussian covariance matrix . Not feasible for large space and time grid : **ok for 40 points in space, but not for 1000**.
- A bit more refined : simulate all spatial points for each day t from previous l days using Gaussian properties and stationnarity : for X_i the vector at all points in space at time i, there are matrices A, B, C such that

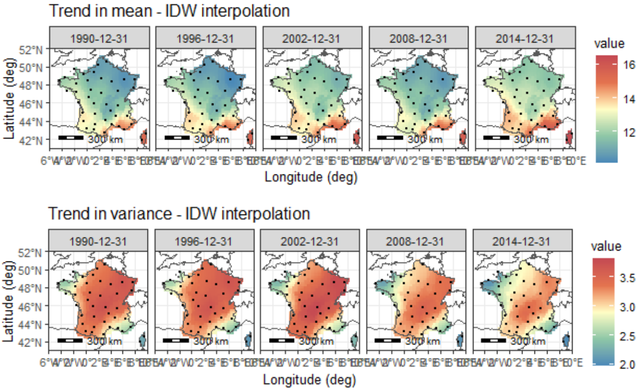
$$\begin{bmatrix} X_1 \\ X_2 \\ \ldots \\ X_l \end{bmatrix} \sim \mathcal{N}(0, A)$$

$$\forall t > \ell,\ X_t | \begin{bmatrix} X_1 \\ \ldots \\ X_{t-1} \end{bmatrix} = X_t \mid \begin{bmatrix} X_{t-\ell} \\ \ldots \\ X_{t-1} \end{bmatrix} \sim \mathcal{N}\left(B \begin{bmatrix} X_{t-\ell} \\ \ldots \\ X_{t-1} \end{bmatrix}, C \right)$$

- A specific algorithm for Gneiting-type covariance functions : use the spectral algorithm[3]

[3] Allard et al. 2020.

Decomposition : trends

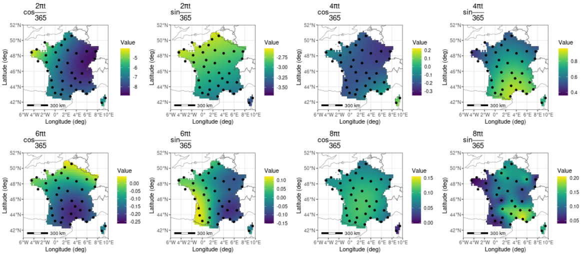


- Choice of presented results : same date each year, but there is no monotonicity (see Toulouse result earlier, especially for the variance)
- Trends show expected French climate

Decomposition : seasonality of the mean

Recall : Seasonality

$$S_{m \text{ or } \sigma}(s,t) = \beta_1(s) + \sum_{i=1}^{d_{m \text{ or } \sigma}} \left[\beta_{2i}(s) cos\left(\frac{2i\pi t}{365}\right) + \beta_{2i+1}(s) sin\left(\frac{2i\pi t}{365}\right)\right]$$



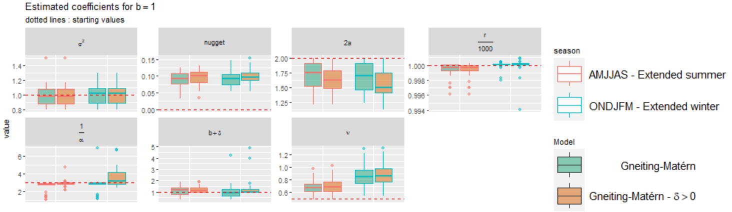
- Different "frequency" give different maps
- Points out different climates
- What could be useful : instead of $A\cos(2k\pi t/365) + B\sin(2k\pi t/365)$, have $C\cos(2k\pi t/365 + \Phi)$ can give more temporal insight
- Seasonality in variance is less interesting to show

Model of the residuals : estimated parameters

Recall : Gneiting-Matérn covariance function

$$C(h,u) = \frac{\sigma^2}{(\alpha u^{2a}+1)^{b+\delta}} \mathcal{M}\left(\frac{h}{\sqrt{\alpha u^{2a}+1}^{b}}; r; \nu\right)$$

Grid search for b in $[0,1]$ $\longrightarrow$ maximum of likelihood for 1.



- Spatial range parameter stayed at initial value : probable redundancy with α, a, b inside the Matérn kernel
- Winter and summer have different ν (smoothness parameter)

Validation

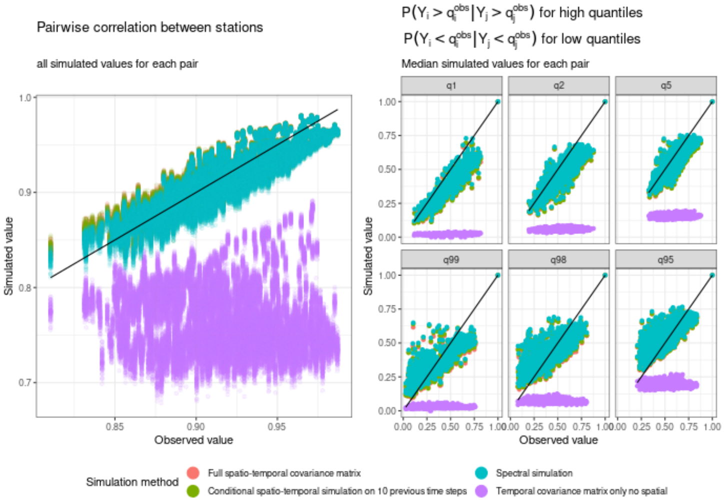
Idea : Compute indicators from the observations and compare with the same indicators from many simulations. The model is adequate if the observed values are in the range of the simulations.
Indicators of good fit :

- **Pairwise correlation** between pairs of stations
- **Pairwise conditional threshold exceedance**
  For every pair of stations i, j, define

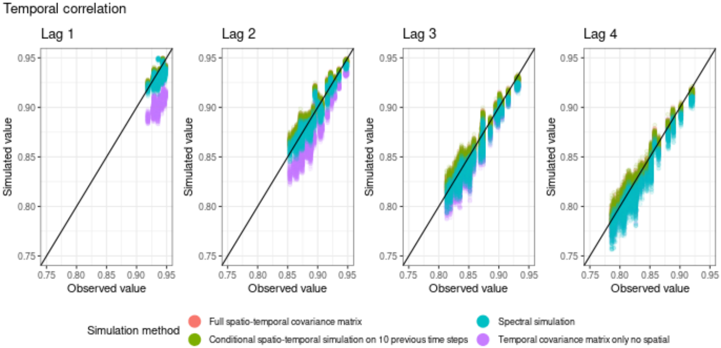
$$p_{i,j}^{\alpha} = P(X_i > q_\alpha(i) | X_j > q_\alpha(j))$$
$$\hat{p}_{i,j}^{\alpha} = \frac{\sum_{t=1}^{N_t} \mathbf{1}_{X_i > q_\alpha(i) \cap X_j > q_\alpha(j)}}{\sum_{t=1}^{N_t} \mathbf{1}_{X_j > q_\alpha(j)}}.$$

  With inverted signs in the case of low quantiles.
- Lagged temporal auto-correlation

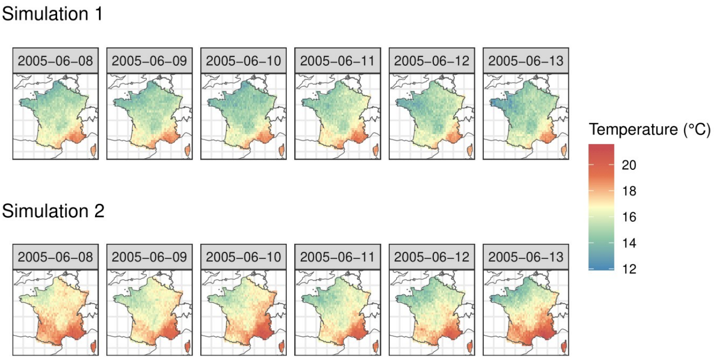
Validation : simulation at the fitting points



Validation : simulation at the fitting points



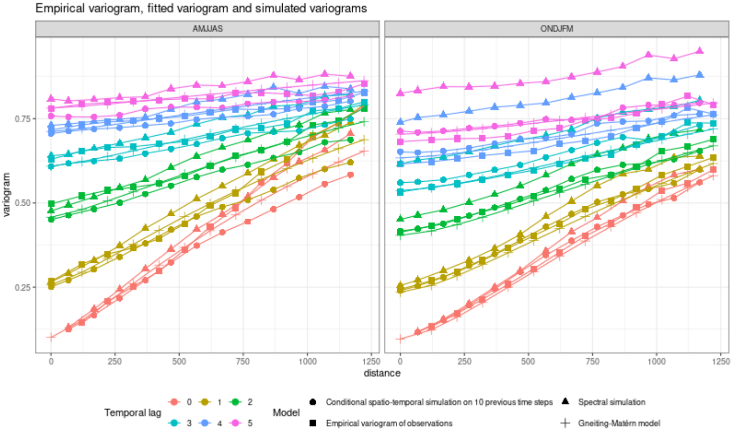
Simulation on a grid : what it looks like



180 days of simulation = 150s (spectral method, can be improved) or 10s (conditionnal iterative method, but with additional time for matrix creation)

Simulation on a grid : variograms

Variograms : $\gamma(h, u) = \frac{1}{2} Var(Z(s+h, t+u) - Z(s,t))$



- The theoretical model is close to the observations
- The simulations are close to the model (the methods worked)

Conclusion

What is done :

- A model for temperature that takes into account spatial structure
- Results : spatial correlation well reproduced, extremal dependance not so much (but not so bad). May need refining.

What I want to do next :

- Compare simulations on a grid to EOBS dataset (same grid)
- Compare interpolated decomposition with exact one
- Use this model with a precipitation model