# Kriging with MK distance

F. Bachoc, F. Gamboa, J-M. LOUBES & N. Venet

Institut de Mathématiques de Toulouse

Oaxaca : Optimal Transport meets Probability, Statistics and Machine Learning

Gaussian Process Models

- Fitting a proper model
- Gaussian process prediction

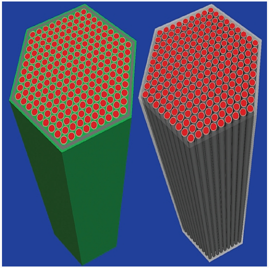
GP indexed by distributions

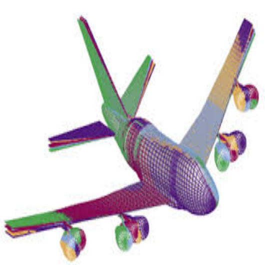
Estimation of outputs of a Gaussian process

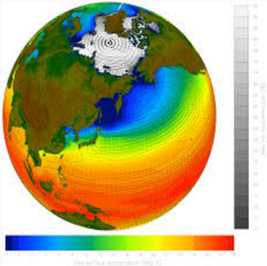
Simulations

# Motivation: computer models

Computer models have become essential in science and industry!







For clear reasons: cost reduction, possibility to explore hazardous or extreme scenarios...

# Computer models as expensive functions

A computer model can be seen as a deterministic function

$$\begin{aligned} f : \mathbb{X} \subset \mathbb{R}^d &\to \mathbb{R} \\ x &\mapsto f(x) \end{aligned}$$

- $x$: tunable simulation parameter (e.g. geometry)
- $f(x)$: scalar quantity of interest (e.g. energetic efficiency)

The function $f$ is usually

- continuous (at least)
- non-linear
- only available through evaluations $x \mapsto f(x)$

$\Longrightarrow$ black box model

# Gaussian Process Models

# Outline

Gaussian Process Models

Fitting a proper model

Gaussian process prediction

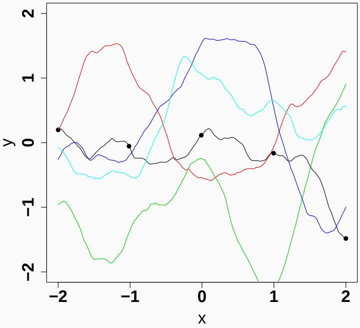
GP indexed by distributions

Estimation of outputs of a Gaussian process

Simulations

# Gaussian process (Kriging model)

Modeling the **black box function** as a **single realization** of a Gaussian process $\xi(x)$ on the domain $\mathbb{X} \subset \mathbb{R}^d$




**Usefulness :Predicting the continuous realization function, from a finite number of observation points**

# Gaussian processes

**Definition**: A stochastic process $\xi : \mathbb{X} \to \mathbb{R}$ is Gaussian if for any $x_1, ..., x_n \in \mathbb{X}$, the vector $(\xi(x_1), ..., \xi(x_n))$ is a Gaussian process

The distribution of a Gaussian process is characterized by

- Its mean function: $x \mapsto m(x) = \mathbb{E}(\xi(x))$. Can be any function $\mathbb{X} \to \mathbb{R}$
- Its covariance function $(x_1, x_2) \mapsto k(x_1, x_2) = Cov(\xi(x_1), \xi(x_2))$

**The covariance function**

- The function $k : \mathbb{X}^2 \to \mathbb{R}$, defined by $k_1(x_1, x_2) = cov(\xi(x_1), \xi(x_2))$

In most classical cases:

- Stationarity: $k(x_1, x_2) = k(x_1 - x_2)$
- Continuity: $k(x)$ is continuous $\Rightarrow$ continuous realizations

# Role of the covariance function

Covariance on metric space We say that a random process $X$ indexed by a metric space $(E, d)$ is *stationary* if it has constant mean and for every isometry $g$ of the metric space we have

$$\mathrm{Cov}(X_{g(x)}, X_{g(y)}) = \mathrm{Cov}(X_x, X_y). \tag{1}$$

We will say that $X$ has *stationary increments* starting in $o \in E$ if $X$ is centred, $X_o = 0$ almost surely, and for every isometry $g$ we have

$$\mathrm{Cov}\left(X_{g(x)} - X_{g(o)}\right) = \mathrm{Cov}\left(X_x - X_o\right). \tag{2}$$

# The covariance function : some conditions are required

The covariance function

$k : (x_1, x_2) \to k(x_1, x_2) = cov(\xi(x_1), \xi(x_2))$

$k$ must me symmetric non-negative definite

$\forall n \in \mathbb{N}, \forall x_1, ..., x_n \in \mathbb{R}^d, \forall \lambda_1, ..., \lambda_n \in \mathbb{R} : \ \sum_{i,j=1}^{n} \lambda_i \lambda_j k(x_i, x_j) \geq 0$

$\Longrightarrow$ the covariance matrix $[k(x_i, x_j)]_{i,j=1,...,n}$ must be non-negative definite

Often, we require the covariance function to be positive definite:

if $(x_1, ..., x_n)$ are 2-by-2 distinct and $(\lambda_1, ..., \lambda_n) \neq (0, ..., 0)$:

$\sum_{i,j=1}^{n} \lambda_i \lambda_j k(x_i, x_j) > 0$

$\Longrightarrow$ the covariance matrix $[k(x_i, x_j)]_{i,j=1,...,n}$ must be positive definite

# Role of the covariance function

Covariance function characterizes the correlations between values of the process at different observation points. As the notion of similarity between data points is crucial, *i.e.* close location inputs are likely to have similar target values, covariance functions are the key ingredient in using Gaussian processes, since they define nearness or similarity. In order to obtain a satisfying model one need to chose a covariance function (*i.e.* a positive definite kernel) that respects the structure of the index space of the dataset.
Huge litterature(Cuturi et al. , Kolouri et al . ....)

# Outline

# Conditional mean as a predictor

Consider a partitioned random vector $(Y_1, Y_2)^t$ of size $(n_1 + 1) \times 1$, with conditional probability density function of $Y_2$ given $Y_1 = y_1$ given by $f_{Y_2|Y_1=y_1}(y_2)$.

Then the conditional mean of $Y_2$ given $Y_1 = y_1$ is

$$\mathbb{E}(Y_2|Y_1 = y_1) = \int_{\mathbb{R}} y_2 f_{Y_2|Y_1=y_1}(y_2) dy_2$$

**Optimality**

The function $y_1 \to \mathbb{E}(Y_2|Y_1 = y_1)$ is the best prediction of $Y_2$ we can make, when observing only $Y_1$. That is, for any function $f : \mathbb{R}^{n_1} \to \mathbb{R}$:

$$\mathbb{E}\left\{(Y_2 - f(Y_1))^2\right\} \geq \mathbb{E}\left\{(Y_2 - \mathbb{E}(Y_2|Y_1))^2\right\}$$

# Conditional variance

Let a random vector $(Y_1, Y_2)^t$ of size $(n_1 + 1) \times 1$, with conditional density $Y_2|Y_1 = y_1$ given by $f_{Y_2|Y_1=y_1}(y_2)$.

Then the conditional variance of $Y_2$ given $Y_1 = y_1$ is

$$var(Y_2|Y_1 = y_1) = \int_{\mathbb{R}} (y_2 - \mathbb{E}(Y_2|Y_1 = y_1))^2 f_{Y_2|Y_1=y_1}(y_2) dy_2$$

**Summary**

- The conditional mean $\mathbb{E}(Y_2|Y_1)$ is the best possible prediction of $Y_2$ given $Y_1$
- The conditional probability density function $y_2 \to f_{Y_2|Y_1=y_1}(y_2)$ can give the probability density function of the corresponding error ($\Rightarrow$ most probable value, probability of threshold exceedance...)
- The conditional variance $var(Y_2|Y_1 = y_1)$ summarizes the order of magnitude of the prediction error

# Gaussian conditioning theorem

**Theorem**

Let $(Y_1, Y_2)^t$ be a $(n_1+1)\times 1$ Gaussian vector with mean vector $(m_1^t, \mu_2)^t$ and covariance matrix

$$\begin{pmatrix} R_1 & r_{1,2} \\ r_{1,2}^t & \sigma_2^2 \end{pmatrix}$$

Then, conditionaly on $Y_1 = y_1$, $Y_2$ is a Gaussian vector with mean

$$\mathbb{E}(Y_2|Y_1 = y_1) = \mu_2 + r_{1,2}^t R_1^{-1}(y_1 - m_1)$$

and variance

$$var(Y_2|Y_1 = y_1) = \sigma_2^2 - r_{1,2}^t R_1^{-1} r_{1,2}$$

# Kriging prediction

We let $Y$ be the Gaussian process, on $\mathbb{R}^d$. $Y$ is observed at $x_1, ..., x_n \in \mathbb{R}^d$. We consider here that we know the covariance function $C$ of $Y$, and that the mean function of $Y$ is zero

**Notations**

- Let $Y_n = (Y(x_1), ..., Y(x_n))^t$ be the observation vector. It is a Gaussian vector
- Let $R$ be the $n \times n$ covariance matrix of $Y_n$: $(R)_{i,j} = C(x_i, x_j)$.
- Let $x_{new} \in \mathbb{R}^d$ be a new input point for the Gaussian process $Y$. We want to predict $Y(x_{new})$.
- Let $r$ be the $n \times 1$ covariance vector between $y$ and $Y(x_{new})$: $r_i = C(x_i, x_{new})$

# Kriging prediction

Then the Gaussian conditioning theorem gives the conditional mean of $Y(x_{new})$ given the observed values in $Y_n$:

$$\hat{y}(x_{new}) := \mathbb{E}(Y(x_{new})|Y_n) = r^t R^{-1} Y_n$$

We also have the conditional variance:

$$\hat{\sigma}^2(x_{new}) := var(Y(x_{new})|Y_n) = C(x_{new}, x_{new}) - r^t R^{-1} r$$

# GP indexed by distributions

# Distribution as entries of a numeric code

Data:

$$(\mu_i, y_i)_{i=1}^n,$$

where the $\mu_i$ are distributions on $\mathbb{R}$.

# Distribution as entries of a numeric code

Data:

$$(\mu_i, y_i)_{i=1}^{n},$$

where the $\mu_i$ are distributions on $\mathbb{R}$.

**Motivations**

- functional entries.
- code to model different kind of variations : probabilities as entries

Model different kind of uncertainties

# Distribution as entries of a numeric code

Data:

$$(\mu_i, y_i)_{i=1}^{n},$$

where the $\mu_i$ are distributions on $\mathbb{R}$.

**Motivations**

- functional entries.
- code to model different kind of variations : probabilities as entries

Model different kind of uncertainties

- Choice of a proper distance through the choice of the kernel

# Monge Kantorovich a.k.a Wasserstein distance

- Assumptions : second moment $\mathcal{W}^2(\mathbb{R})$.

# Monge Kantorovich a.k.a Wasserstein distance

- Assumptions : second moment $\mathcal{W}^2(\mathbb{R})$.
- Quadratic transportation cost between $\mu$ and $\nu$ is defined by

$$W_2(\mu,\nu) := \left( \inf_{\pi \in \Pi(\mu,\nu)} \int |x-y|^2 \, d\pi(x,y) \right)^{1/2}, \tag{3}$$

where $\Pi(\mu,\nu)$ is the set of probabilities on $\mathbb{R}^2$ with marginals distributions $\mu$ and $\nu$.

# Monge Kantorovich a.k.a Wasserstein distance

- Assumptions : second moment $\mathcal{W}^2(\mathbb{R})$.
- Quadratic transportation cost between $\mu$ and $\nu$ is defined by

$$W_2(\mu,\nu) := \left( \inf_{\pi \in \Pi(\mu,\nu)} \int |x-y|^2 \, d\pi(x,y) \right)^{1/2}, \qquad (3)$$

  where $\Pi(\mu,\nu)$ is the set of probabilities on $\mathbb{R}^2$ with marginals distributions $\mu$ and $\nu$.
- Problem : finding *stationnary kernels non negative* on $\mathcal{W}^2(\mathbb{R})$.

# Negative Kernels

**Theorem**

*For all* $H \in [0,1]$,

$$K : (\mu,\nu) \mapsto W_2(\mu,\nu))^{2H} \tag{4}$$

*is a negative definite kernel if and only if* $0 \leq H \leq 1$ *:*
$\forall \mu_1, \cdots, \mu_n \in \mathcal{W}_2(\mathbb{R})$, $\forall c_1, \cdots, c_n \in \mathbb{R}$ *t.q.* $\sum_{i=1}^n c_i = 0$,

$$\sum_{i,j=1}^{n} c_i c_j W_2(\mu_i,\mu_j))^{2H} \leq 0. \tag{5}$$

- The fractional exponent $\beta_{\mathcal{W}_2(\mathbb{R})}$ of the Wasserstein space is equal to 2.

# Applications

**Theorem (Fractionnary Brownian Field)**

*For all* $0 \leq H \leq 1$ $\mu$ *and* $\sigma \in \mathcal{W}_2(\mathbb{R})$,

$$K^{H,\sigma}(\mu,\nu) = \frac{1}{2}\left(W_2^{2H}(\sigma,\mu) + W_2^{2H}(\sigma,\nu) - W_2^{2H}(\mu,\nu)\right) \tag{6}$$

*is a proper covariance function on* $\mathcal{W}_2(\mathbb{R})$. *It is non degenerate if and only if* $0 < H < 1$.

- We can define with this covariance function a fractional brownian field $\mathcal{W}_2(\mathbb{R})$. non stationary but stationary increments.

The centered Gaussian process $(X_\mu)_{\mu \in \mathcal{W}_2(\mathbb{R})}$ such that

$$\begin{cases} \mathbb{E} X_\mu = 0, \\ \mathrm{Cov}(X_\mu, X_\nu) = K^{H,\sigma}(\mu,\nu) \end{cases} \tag{7}$$

is the $H$-fractional Brownian motion with index space $\mathcal{W}_2(\mathbb{R})$ and origin in $\sigma$. It is the only Gaussian random process such that

$$\begin{cases} \mathbb{E} X_\mu = 0, \\ \mathbb{E}(X_\mu - X_\nu)^2 = W_2^{2H}(\mu,\nu), \\ X_\sigma = 0 \text{ almost surely.} \end{cases} \tag{8}$$

It is a generalization of the seminal fractional Brownian motion on the real line.

$$Y_\mu := X_{(\mu - m(\mu))} + m(\mu), \quad m(\mu) := \int x d\mu(x).$$

# Applications 2

**Theorem (Schoenberg)**

*Let $F : \mathbb{R}^+ \to \mathbb{R}^+$ be a completely monotone function, and $K$ a negative definite kernel. Then $(x, y) \mapsto F(K(x, y))$ is a positive definite kernel.*

- Recall that $F : \mathbb{R}^+ \to \mathbb{R}^+$ is fully monotone if and only if it is indefinitely derivable such that $(-1)^n F^{(n)}$ is positive for any $n \in \mathbb{N}$.

# Parametric Gaussian mode

**Theorem (Stationary Processes )**

*For any* $F : \mathbb{R}^+ \to \mathbb{R}^+$ *fully monotone and for* $0 < H \leq 1$*,*

$$K : (\mu, \nu) \mapsto F\left(W_2^{2H}(\mu, \nu)\right) \tag{9}$$

*is the covariance function of a staionnary Gaussian process indexed by* $\mathcal{W}_2(\mathbb{R})$*.*

**In very particular**

$$K_{\sigma^2,\ell,H}(\nu_1, \nu_2) = \sigma^2 \exp\left(-\frac{W_2(\nu_1, \nu_2)^{2H}}{\ell}\right), \tag{10}$$

$$H \in [0, 1], \sigma > 0, l > 0,$$

provides a parametric model of stationary Gaussian processes indexed by $\mathcal{W}_2(\mathbb{R})$.

# Estimation of outputs of a Gaussian process

# Model estimation

$$L_\theta = \frac{1}{n}\ln(\det R_\theta) + \frac{1}{n} y^t R_\theta^{-1} y, \tag{11}$$

where $R_\theta = [K_\theta(\mu_i, \mu_j)]_{1 \le i,j \le n}$

Consistency of maximum likelihood estimator

$$\hat{\theta}_{ML} \in \arg\min_{\theta \in \Theta} L_\theta$$

**Theorem**

*Under the conditions 2 to 5*

$$\hat{\theta}_{ML} \xrightarrow[n \to \infty]{\mathbb{P}} \theta_0.$$

# Ideas of proof

$$\sup_{\theta \in \Theta} \| L_\theta - \mathbb{E}(L_\theta) \| = o_{\mathbb{P}}(1). \tag{12}$$

So we obtain the existence of a positive $a$ such that

$$\mathbb{E}(L_\theta) - \mathbb{E}(L_{\theta_0}) \geq c \frac{1}{n} \| R_\theta - R_{\theta_0} \|^2.$$

Hence we have $\forall \alpha > 0$,

$$\mathbb{P}\left( \left\| \hat{\theta}_{ML} - \theta_0 \right\| \geq \alpha \right) \underset{n \to \infty}{\longrightarrow} 0$$

and so

$$\hat{\theta}_{ML} \xrightarrow[n \to \infty]{\mathbb{P}} \theta_0.$$

# Assumptions

**Condition (1)**

*Data is a triangular array* $\mathcal{W}_2(\mathbb{R})$ $\{\mu_1, ..., \mu_n\} = \{\mu_1^{(n)}, ..., \mu_n^{(n)}\}$ *such that for all* $n \in \mathbb{N}$ *and* $1 \leq i \leq n$, $\mu_i$ *as support in* $[i, i+K]$, *where* $K < \infty$.

**Condition (2)**

*The covariance functions* $\{K_\theta, \theta \in \Theta \subset \mathbb{R}^p\}$ *are such that*

$$\forall \theta \in \Theta, K_\theta(\mu, \nu) = F_\theta\left(W_2(\mu, \nu)\right) \text{and} \sup_{\theta \in \Theta} |F_\theta(t)| \leq \frac{A}{1 + |t|^{1+\tau}}$$

*where* $A < \infty$ *and* $\tau > 1$ *are constant.*

**Condition (3)**

*Observations $y_i = Y(\mu_i)$, $i = 1, \cdots, n$ are drawn from a Gaussian process $Y$, centered with covariance $K_{\theta_0}$ for a $\theta_0 \in \Theta$.*

**Condition (4)**

*The sequence of matrices $R_\theta = (K_\theta(\mu_i, \mu_j))_{1 \leq i,j \leq n}$ is such that $\lambda_{\inf}(R_\theta) \geq c$ for a constant $c > 0$, with $\lambda_{\inf}(R_\theta)$ the smallest eigenvalue of $R_\theta$.*

**Condition (5)**

$$\forall \alpha > 0, \liminf_{n \to \infty} \inf_{\|\theta - \theta_0\| \geq \alpha} \frac{1}{n} \sum_{i,j=1}^{n} \left[K_\theta(\mu_i, \mu_j) - K_{\theta_0}(\mu_i, \mu_j)\right]^2 > 0.$$

**Lemma (expansion model)**

$$\sup_{\mu \in W_2(\mathbb{R})} \sup_{\theta \in \Theta} \sum_{i=1}^{n} |K_\theta(\mu_i, \mu_j)|$$

*is bounded as* $n \to \infty$.

**Lemma**

*Under Conditions 2 to 5,*

$$\sup_{\theta \in \Theta} \lambda_{\max}(R_\theta)$$

*and*

$$\sup_{\theta \in \Theta} \max_{i=1 \cdots p} \lambda_{\max}\left(\frac{\partial}{\partial \theta_i} R_\theta\right)$$

*are bounded as* $n \to \infty$.

# Asymptotic behavious of MLE

**Theorem**

*Let $M_{ML}$ a matrix $p \times p$ defined by*

$$(M_{ML})_{i,j} = \frac{1}{2n} \mathrm{Tr} \left( K_{\theta_0}^{-1} \frac{\partial K_{\theta_0}}{\partial \theta_i} K_{\theta_0}^{-1} \frac{\partial K_{\theta_0}}{\partial \theta_j} \right).$$

*Under Conditions 2 to 9, we get*

$$\sqrt{n} M_{ML}^{1/2} \left( \hat{\theta}_{ML} - \theta_0 \right) \xrightarrow[n \to \infty]{\mathcal{L}} \mathcal{N}(0, I_p).$$

*Moreover*

$0 < \liminf_{n \to \infty} \lambda_{min}(M_{ML}) \leq \limsup_{n \to \infty} \lambda_{max}(M_{ML}) < +\infty.$

Hence the parametric process is fitted to the model.

**Condition (6)**

*$\forall t \geq 0$, $F_\theta(t)$ is of class $\mathcal{C}^1$ w.r.t $\theta$ and satisfies* $\sup_{\theta\in\Theta} \max_{i=1,\cdots,p} \left| \frac{\partial}{\partial\theta_i} F_\theta(t) \right| \leq \frac{A}{1+t^{1+\tau}}$*, with $A, \tau$ defined in Condition 3.*

**Condition (7)**

*For all $t \geq 0$, $F_\theta(t)$ is $\mathcal{C}^3$ w.r.t $\theta$ e and $\forall q \in \{2, 3\}$, $\forall i_1 \cdots i_q \in \{1, \cdots p\}$,*

$$\sup_{\theta\in\Theta} \max_{i=1,\cdots,p} \left| \frac{\partial}{\partial\theta_{i_1}} \cdots \frac{\partial}{\partial\theta_{i_q}} F_\theta(t) \right| \leq \frac{A}{1+|t|^{1+\tau}}.$$

**Condition (8)**

$\forall(\lambda_1 \cdots, \lambda_p) \neq (0, \cdots, 0)$,

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{i,j=1}^{n} \left( \sum_{k=1}^{p} \lambda_k \frac{\partial}{\partial_{\theta_k}} K_{\theta_0}(\mu_i, \mu_j) \right)^2 > 0.$$

## Forecast using GP modeling

$$\hat{Y}_\theta(\mu) = r_\theta^t(\mu) R_\theta^{-1} y \qquad (13)$$

and

$$r_\theta(\mu) = \begin{bmatrix} K_\theta(\mu, \mu_1) \\ \vdots \\ K_\theta(\mu, \mu_n) \end{bmatrix}.$$

$\hat{Y}_\theta(\mu)$ is the conditional expectation of $Y(\mu)$ given $y_1, ..., y_n$, when $Y$ is a centered Gaussian process with covariance $K_\theta$.

**Theorem**

*Under Conditions 2 to 9, the Kriging estimator built using the parameter $\hat{\theta}_{ML}$ is asymptotically optimal in the sense*

$$\forall \mu \in \mathcal{W}_2(\mathbb{R}), \ \left| \hat{Y}_{\hat{\theta}_{ML}}(\mu) - \hat{Y}_{\theta_0}(\mu) \right| = o_{\mathbb{P}}(1).$$

# Simulations

## Simulated Data

- Let $m_k(\nu)$ the $k$ order moment of $\nu$ and set $F : \mathcal{W}_2(\mathbb{R}) \to \mathbb{R}$ such that

$$F(\nu) = \frac{m_1(\nu)}{0.05 + \sqrt{m_2(\nu) - m_1(\nu)^2}}, \tag{14}$$

standing for the code to be forecast

## Simulated Data

- Let $m_k(\nu)$ the $k$ order moment of $\nu$ and set $F : \mathcal{W}_2(\mathbb{R}) \to \mathbb{R}$ such that

$$F(\nu) = \frac{m_1(\nu)}{0.05 + \sqrt{m_2(\nu) - m_1(\nu)^2}}, \tag{14}$$

standing for the code to be forecast

- Entries : $\nu_1, \cdots, \nu_{100}$ random Gaussian

## Simulated Data

- Let $m_k(\nu)$ the $k$ order moment of $\nu$ and set $F : \mathcal{W}_2(\mathbb{R}) \to \mathbb{R}$ such that

$$F(\nu) = \frac{m_1(\nu)}{0.05 + \sqrt{m_2(\nu) - m_1(\nu)^2}}, \tag{14}$$

  standing for the code to be forecast
- Entries : $\nu_1, \cdots, \nu_{100}$ random Gaussian
- Maximum likelihood $\hat{\sigma}^2, \hat{\ell}, \hat{H}$ for Gaussian parametric model

$$K_{\sigma^2,\ell,H}(\nu_1, \nu_2) = \sigma^2 \exp\left(-\frac{W_2(\nu_1, \nu_2)^{2H}}{\ell}\right). \tag{15}$$

# Simulated Data

Test dataset $(\nu_{t,i})_{i=500}$

$$RMSE^2 = \frac{1}{500}\sum_{i=1}^{500}\left(F(\nu_{t,i}) - \hat{F}(\nu_{t,i})\right)^2,$$

$$CIR_\alpha = \frac{1}{500}\sum_{i=1}^{n_t} \mathbf{1}\left\{\left|F(\nu_{t,i}) - \hat{F}(\nu_{t,i})\right| \leq q_\alpha \hat{\sigma}(\nu_{t,i})\right\},$$

# Simulated Data

Test dataset $(\nu_{t,i})_{i=500}$

$$RMSE^2 = \frac{1}{500}\sum_{i=1}^{500}\Big(F(\nu_{t,i}) - \hat{F}(\nu_{t,i})\Big)^2,$$

$$CIR_\alpha = \tfrac{1}{500}\textstyle\sum_{i=1}^{n_t} \mathbf{1}\left\{\left|F(\nu_{t,i}) - \hat{F}(\nu_{t,i})\right| \leq q_\alpha \hat{\sigma}(\nu_{t,i})\right\},$$

| Model | RMSE | $CIR_{0.9}$ |
|---|---|---|
| “distribution” | 0.094 | 0.92 |
| “Legendre” ordre 5 | 0.49 | 0.92 |
| “Legendre” ordre 10 | 0.34 | 0.89 |
| “Legendre” ordre 15 | 0.29 | 0.91 |
| “PCA” ordre 5 | 0.63 | 0.82 |
| “PCA” ordre 10 | 0.52 | 0.87 |
| “PCA” ordre 15 | 0.47 | 0.93 |

## Next ...

- Not working directly in dimension $\geq 2$
- Extension using copulas ...
- Not working great in practice for the moment
- Real datasets from CEA