

INSA – Gaussian processes

Introduction

Andrés F. López-Lopera

The French Aerospace Lab ONERA, France
Information Processing and Systems Department (DTIS)
Multidisciplinary Methods, Integrated Concepts (M2CI) Research Unit

Who am I?



Andrés F. López-Lopera

Colombia

2008-2013

Electrical Eng., Universidad Tecnológica de Pereira

- Machine learning and signal processing

2014-2015

M.Sc. in Electrical Eng., Universidad Tecnológica de Pereira

- Probabilistic modelling using Gaussian processes (GPs)

France

2016-2019

PhD in Applied Mathematics, Mines Saint-Étienne

- Joint supervision: *Institut de Mathématiques de Toulouse*
- GPs under inequality constraints
- Applications: nuclear risk assessment, coastal flooding

2019-2020

Postdoctoral Research, Institut de Mathématiques de Toulouse

- Joint supervision: *The French Geological Survey BRGM*
- Multi-output GPs & coastal flooding

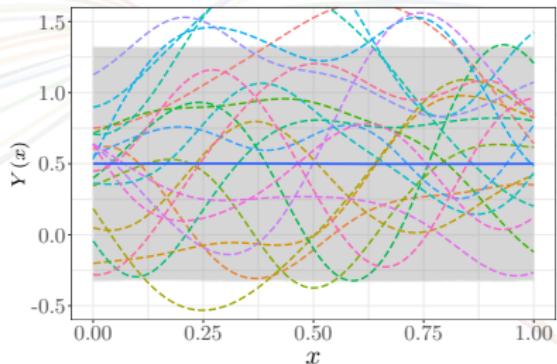
2020-2021

Postdoctoral Research, The French Aerospace Lab ONERA

- Multi-fidelity GPs & aerodynamics (wind tunnel tests)

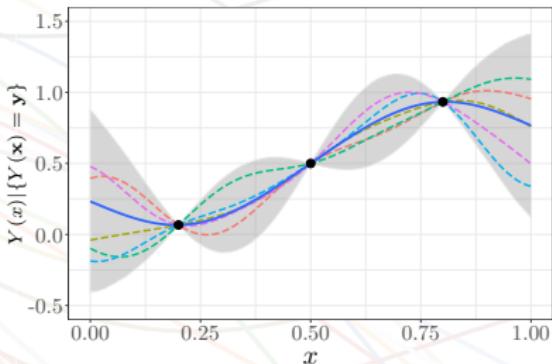
Gaussian processes (GPs) as flexible priors over functions

GP prior



$$Y \sim \mathcal{GP}(m, k_{\theta})$$

GP regression

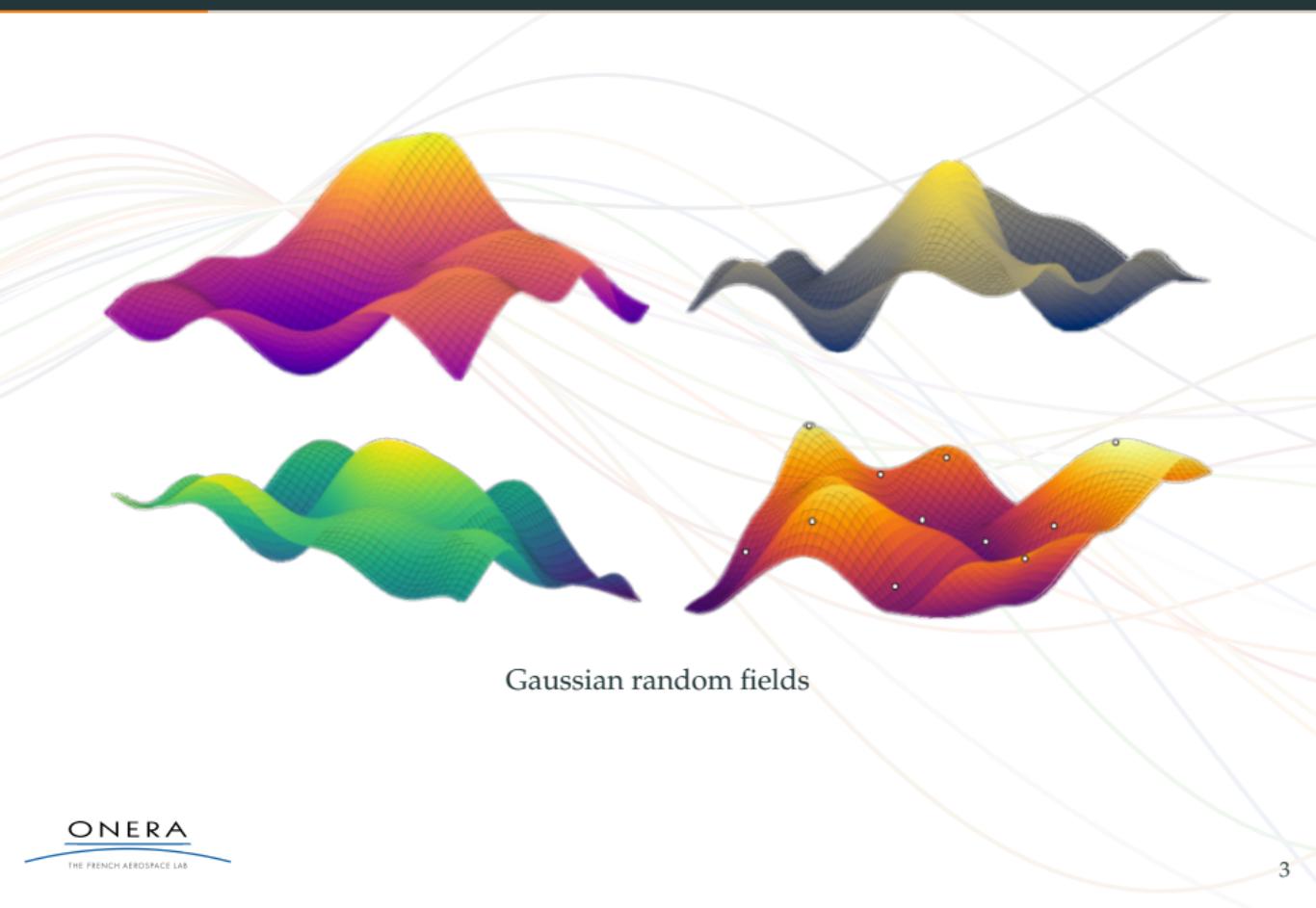


$$Y | \{Y(x) = y\} \sim \mathcal{GP}(m_{\text{cond}}, c_{\theta})$$

■ mean function ■ prediction intervals ■ ■ ... ■ samples

- Interpolation conditions: $(x, y) = (x_i, y_i)_{i=1}^n$

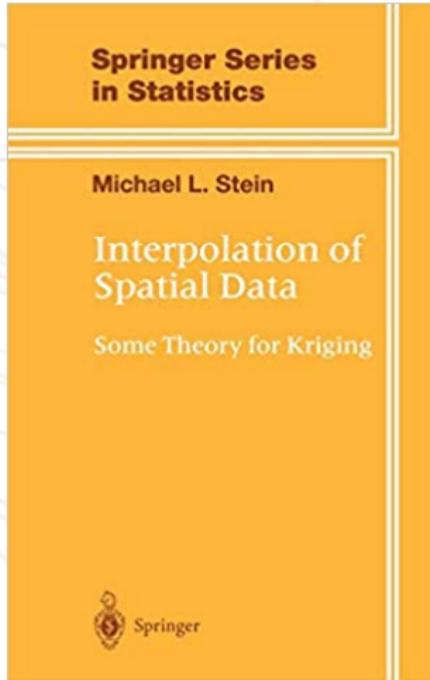
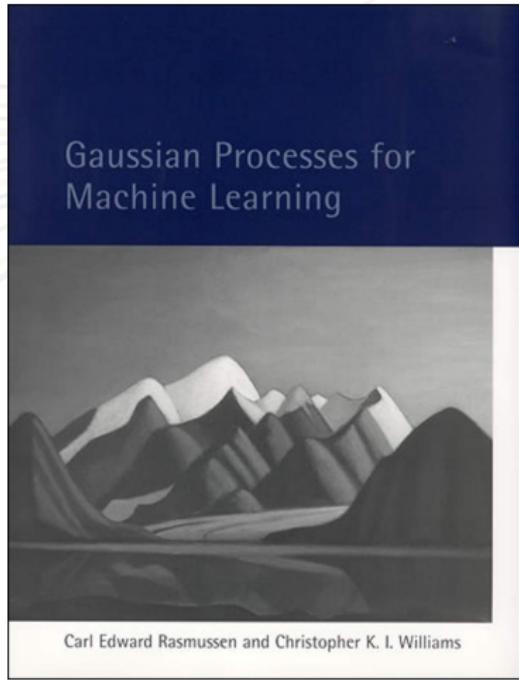
Gaussian processes (GPs) as flexible priors over functions



Outline

- In this course:
 1. A recap of Gaussian processes
 2. Spectral representation and Bochner's theorem
 3. Regularity conditions (e.g. continuity, differentiability)
 4. An introduction to reproducing kernel Hilbert-spaces (RKHS)
- Material of the lectures can be found at: <https://anfelopera.github.io/>

Main references



<http://www.gaussianprocess.org/gpml/>

<https://www.springer.com/gp/book/9780387986296>

Additional references

- Alain Berlinet and Christine Thomas-Agnan. *Reproducing kernel Hilbert spaces in probability and statistics*. Springer Science & Business Media, 2011.
- C. Chatfield. *The Analysis of Time Series: An Introduction, Sixth Edition*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press, 2016.
- Marc G. Genton. Classes of kernels for machine learning: A statistics perspective. *Journal of Machine Learning Research*, 2001.
- Jochen Görtler, Rebecca Kehlbeck, and Oliver Deussen. A visual exploration of Gaussian processes. *Distill*, 2019.
<https://distill.pub/2019/visual-exploration-gaussian-processes>.
- Kevin P. Murphy. *Probabilistic Machine Learning: An introduction*. MIT Press, 2021.
- Carl E. Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. MIT Press, 2005.
- Arno Solin. *Machine learning with signal processing*. ICML – TUTORIAL, 2020.
- Michael L. Stein. *Interpolation of Spatial Data: Some Theory for Kriging*. Springer, 1999.

Gaussian processes

Gaussian processes

- Let $\{Y(x); x \in \mathbb{R}^d\}$ be a GP
- Y is completely defined by its mean m and covariance (kernel) k functions:

$$Y \sim \mathcal{GP}(m, k), \quad (1)$$

where

$$\begin{array}{ll} \text{(trend)} & m(x) = \mathbb{E} \{ Y(x) \}, \\ \text{(correlation, p.s.d.)} & k(x, x') = \text{cov} \{ Y(x), Y(x') \}, \quad \text{for } x, x' \in \mathbb{R}^d. \end{array} \quad (2)$$

- The operator \mathbb{E} denotes the expectation of random variables (r.v's), and the covariance operator is given by

$$\text{cov} \{ Y(x), Y(x') \} = \mathbb{E} \{ [Y(x) - m(x)][Y(x') - m(x')] \}.$$

- It is common to assume that Y has mean zero, i.e. $m(\cdot) = 0$. Then,

$$k(x, x') = \text{cov} \{ Y(x), Y(x') \} = \mathbb{E} \{ Y(x)Y(x') \}. \quad (3)$$

- If $m(\cdot) = 0$, then Y is known as a centred GP.

Exercise. Show that $Z \sim \mathcal{GP}(m, k)$ can be written in terms of $Y \sim \mathcal{GP}(0, k)$:

$$Z(x) = m(x) + Y(x). \quad (4)$$

- If Y is a centred GP, then it is completely defined by its kernel k .
- Regularity assumptions are then encoded in k [Genton, 2001]:
 - stationarity
 - isotropy
 - periodicity
 - smoothness

Definition (Stationary kernel functions)

A kernel function $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$, with $\mathcal{X} \subset \mathbb{R}^d$, is **stationary** if, for all $x, x' \in \mathcal{X}$, $k(x, x')$ only depends on $x - x'$.

Definition (Isotropic kernel functions)

A kernel k is **isotropic** (or homogeneous) if $k(x, x')$ only depends on $|x - x'|$.

Examples of 1D stationary kernels

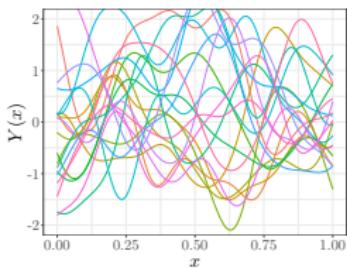
- Denote $k(\tau) := k(x, x + \tau)$ (abuse of notation)
- Some classic 1D stationary kernels are [Genton, 2001]:

Squared Exponential (SE): $k_{\sigma^2, \ell}(\tau) = \sigma^2 \exp \left\{ -\frac{1}{2} \frac{\tau^2}{\ell^2} \right\},$

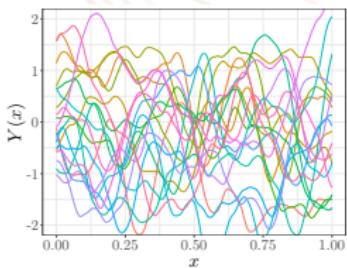
Matérn 5/2: $k_{\sigma^2, \ell}(\tau) = \sigma^2 \left(1 + \sqrt{5} \frac{|\tau|}{\ell} + \frac{5}{3} \frac{\tau^2}{\ell^2} \right) \exp \left\{ -\sqrt{5} \frac{|\tau|}{\ell} \right\},$

Exponential: $k_{\sigma^2, \ell}(\tau) = \sigma^2 \exp \left\{ -\frac{|\tau|}{\ell} \right\},$

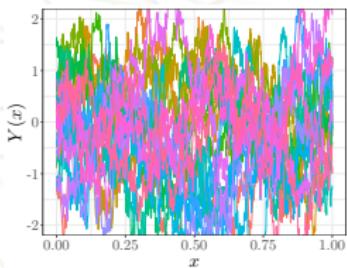
with variance parameter σ^2 and length-scale parameter ℓ .



(a) Squared Exponential



(b) Matérn 5/2

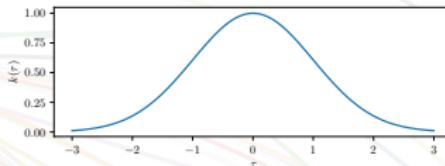


(c) Exponential

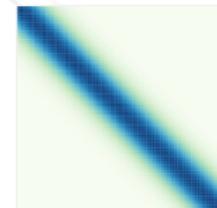
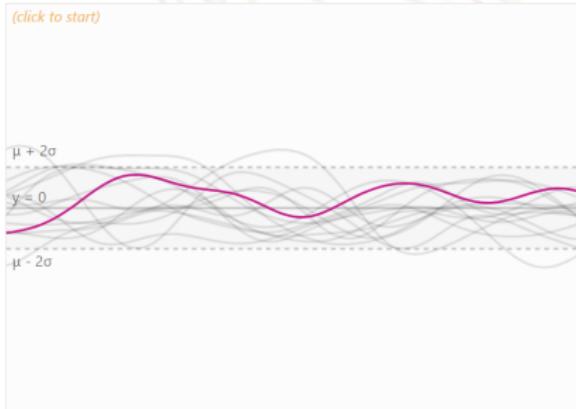
█ █ ... █ samples

Examples of 1D stationary kernels

Squared Exponential (SE): $k_{\sigma^2, \ell}(\tau) = \sigma^2 \exp \left\{ -\frac{1}{2} \frac{\tau^2}{\ell^2} \right\}$,



(click to start)



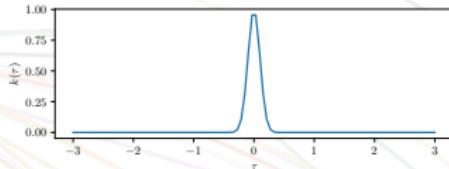
Variance $\sigma^2 = 1$

Length $\ell = 1$

Effect of the variance σ^2 and the length-scale ℓ [Görtler et al., 2019]

Examples of 1D stationary kernels

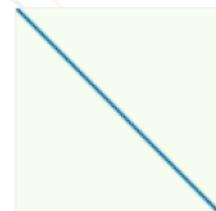
Squared Exponential (SE): $k_{\sigma^2, \ell}(\tau) = \sigma^2 \exp \left\{ -\frac{1}{2} \frac{\tau^2}{\ell^2} \right\}$,



(click to start)

$\mu + 2\sigma$

$\mu - 2\sigma$



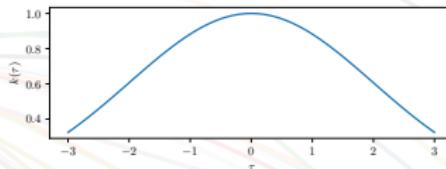
Variance $\sigma^2 = 1$

Length $\ell = 0.1$

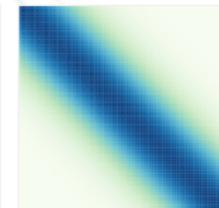
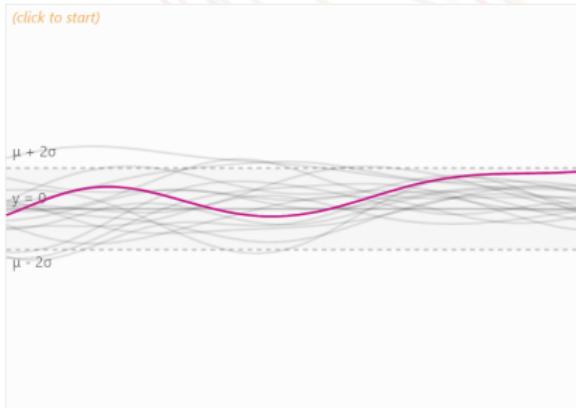
Effect of the variance σ^2 and the length-scale ℓ [Görtler et al., 2019]

Examples of 1D stationary kernels

Squared Exponential (SE): $k_{\sigma^2, \ell}(\tau) = \sigma^2 \exp \left\{ -\frac{1}{2} \frac{\tau^2}{\ell^2} \right\}$,



(click to start)



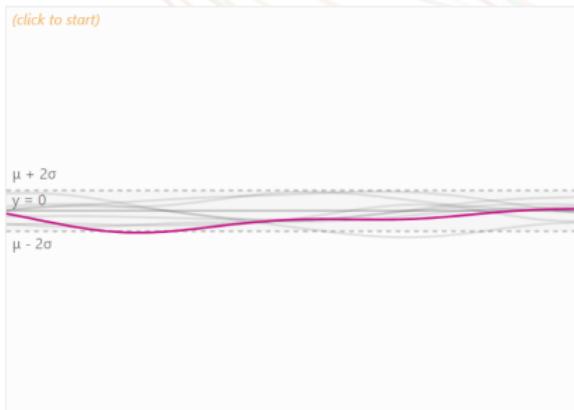
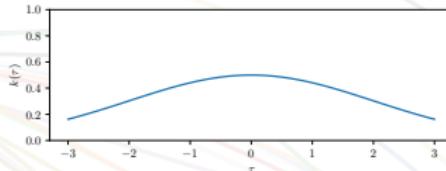
Variance $\sigma^2 = 1$

Length $\ell = 2$

Effect of the variance σ^2 and the length-scale ℓ [Görtler et al., 2019]

Examples of 1D stationary kernels

Squared Exponential (SE): $k_{\sigma^2, \ell}(\tau) = \sigma^2 \exp \left\{ -\frac{1}{2} \frac{\tau^2}{\ell^2} \right\}$,

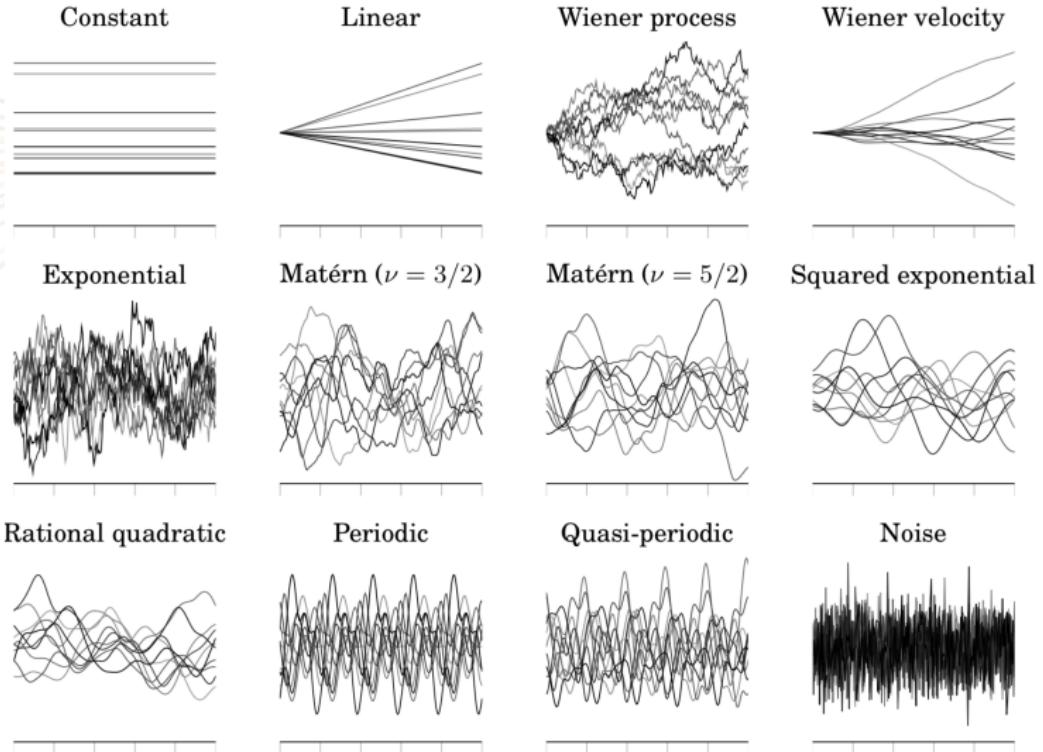


Variance $\sigma^2 = 0.5$

Length $\ell = 2$

Effect of the variance σ^2 and the length-scale ℓ [Görtler et al., 2019]

Other examples of 1D kernels



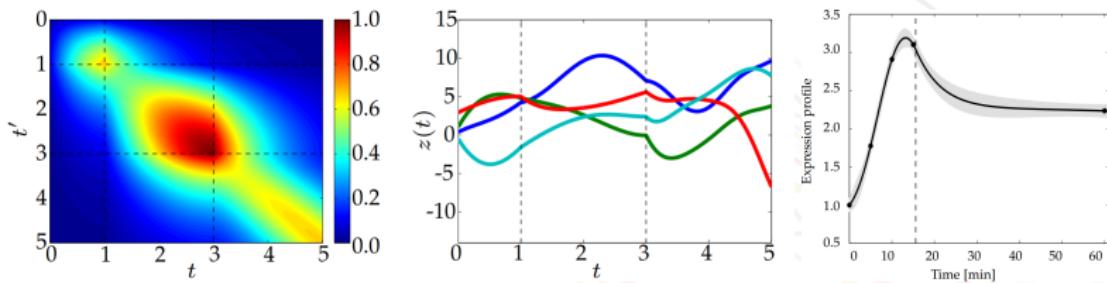
Examples of GP samples [Solin, 2020]

A visual exploration of GPs

- **Görtler et al. [2019]:** A visual exploration of Gaussian processes [[Link](#)]
- **Damianou [2016]:** A Python notebook on Gaussian processes [[Link](#)]

Applications of Gaussian processes

Biology: prediction of protein concentrations

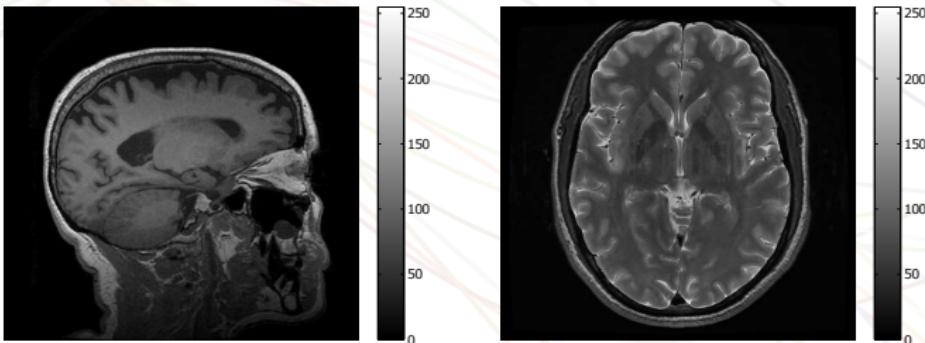


- A. F. López-Lopera and M. A. Alvarez:

Switched latent force models for reverse-engineering transcriptional regulation in genes
IEEE/ACM Transaction on Computational Biology and Bioinformatics, 2017

Applications of Gaussian processes

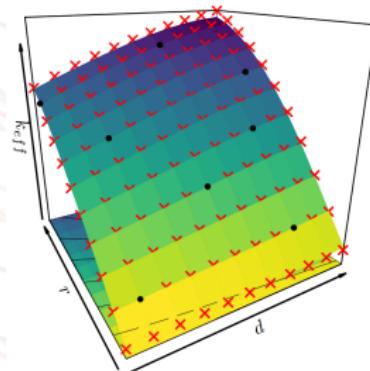
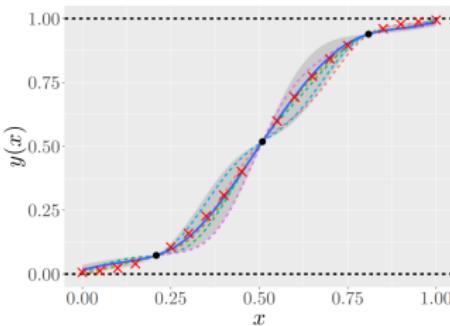
Neuroscience: magnetic resonance imaging (MRI)



- H. Vargas, A. López-Lopera, M. A. Ivarez, A. Orozco, J. Hernández and N. Malpica:
Gaussian processes for slice-based super-resolution MR images
Lecture Notes in Computer Science (LNCC), 2015

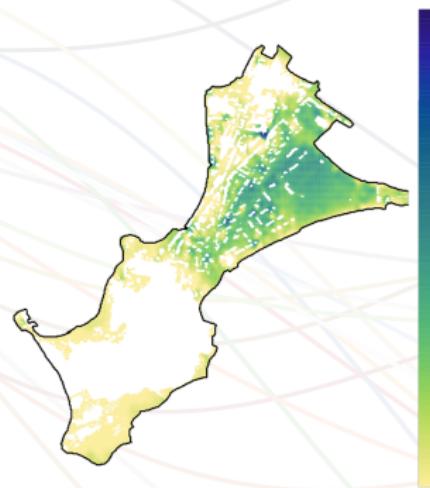
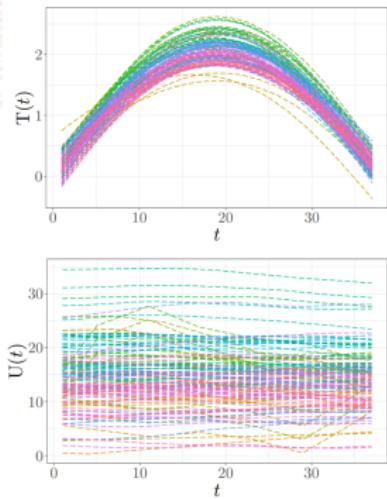
Applications of Gaussian processes

Risk assessment: nuclear safety



- A. F. López-Lopera, N. Durrande, F. Bachoc and O. Roustant:
Finite-dimensional Gaussian approximation with linear inequality constraints
SIAM/ASA Journal on Uncertainty Quantification, 2018

Risk assessment: coastal flooding

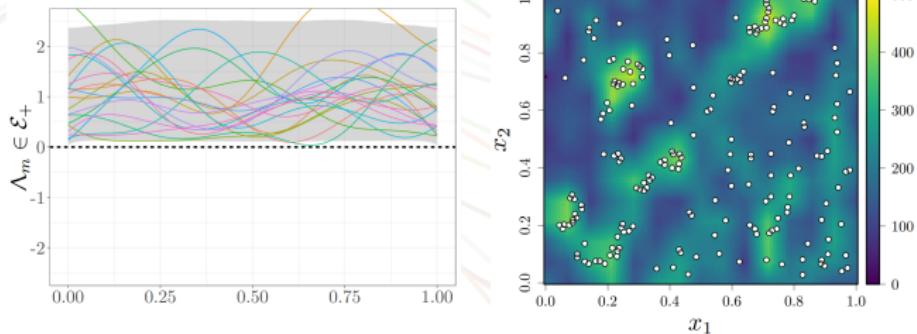


- A. F. López-Lopera, D. Idier, J. Rohmer and F. Bachoc:

Multi-output Gaussian processes with functional data: A study on coastal flood hazard assessment
Submitted, 2020

Applications of Gaussian processes

Geostatistics: spatial distribution of tree species



- A. F. López-Lopera, S. John and N. Durrande:

Gaussian process modulated Cox processes under linear inequality constraints
International Conference on Artificial Intelligence and Statistics (AISTATS), 2019

Conclusions

Conclusions

- GPs provide a well-founded non-parametric (Bayesian) framework
- They have been successfully applied in diverse applications:
 - Geostatistics, physics, chemistry
 - Neuroscience, biology and medicine
 - Engineering fields
 - Econometrics
 - ...
- Regularity assumptions are encoded in kernel functions
 - smoothness, periodicity, stationarity, isotropy, ...