Physics-aware models with SPDEs: Models and Perspectives

M. PEREIRA
Geosciences and Geoengineering Department, Mines Paris - PSL University
mike.pereira@minesparis.psl.eu

Joint work with L. Clarotto, N. Desassis, A. Lang, C. Sire, G. Victorino Cardoso

Spatial Statistics 2025
July 17th, 2025

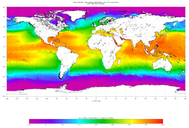
GEOLEARNING
CHAIRE /// Data Science for the Environment

DATA FROM PHYSICAL PROCESSES

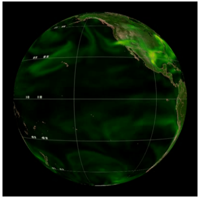
Many applications of Geostatistics to data linked to physical processes

- Climate
  - Interpolation of climate/weather variables (Temperature, Precipitation, Geotrophic winds,...)
  - Global scale data assimilation
  - Stochastic Generators
- Environment
  - Diffusion of pollutant in soils
  - Transport of pollutant in atmosphere, Air quality monitoring
- And other fields
  - Disease mapping
  - Seismic wave mapping

→ Simple (PDE-based) physical can be used to describe the corresponding variables



Daily Sea Surface Temperature (Source: NOAA)



Sulfate dispersion in the atmosphere (Source: NASA)

DATA FROM PHYSICAL PROCESSES

Examples of "physical" models

- Diffusion phenomena (eg. Temperature): Heat equation

$$\frac{\partial Z}{\partial t} - \theta \Delta Z = u$$

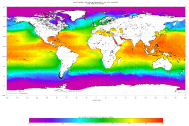
- Transport phenomena (eg. Concentration of pollutant in the air): Advection-Diffusion equation

$$\frac{\partial Z}{\partial t} + \gamma \cdot \nabla Z - \theta \Delta Z = u$$

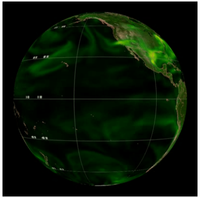
- Wave propagation : Wave equation

$$\frac{\partial^2 Z}{\partial t^2} - \theta \Delta Z = 0$$

→ *Can we incorporate this a priori knowledge into geostatistical modeling?*



Daily Sea Surface Temperature (Source: NOOA)



Sulfate dispersion in the atmosphere (Source: NASA)

SETTING

Quantity of interest (QoI) Physical variable Z over a spatial/spatio-temporal domain $\mathcal{D}$

Data (Noisy) Observations $Y(x_i)$ of Z at locations $(x_1, \dots, x_{N_D}) \in \mathcal{D}$

SETTING

Quantity of interest (QoI) Physical variable Z over a spatial/spatio-temporal domain $\mathcal{D}$

Data (Noisy) Observations $Y(x_i)$ of Z at locations $(x_1, \dots, x_{N_D}) \in \mathcal{D}$

Goal (Probabilistic) Predictions of the QoI Z at unobserved locations

$\rightarrow$ Need a probabilistic model on Y that accounts for spatial "structure" of Y

SETTING

Quantity of interest (QoI) Physical variable Z over a spatial/spatio-temporal domain $\mathcal{D}$

Data (Noisy) Observations $Y(x_i)$ of Z at locations $(x_1, \dots, x_{N_D}) \in \mathcal{D}$

Goal (Probabilistic) Predictions of the QoI Z at unobserved locations

→ Need a probabilistic model on Y that accounts for spatial "structure" of Y

→ If possible, account for the fact that the QoI arises from a physical model

$$\begin{cases} \mathcal{L}Z = u & \text{on } \mathcal{D} \\ \mathcal{B}Z = u_B & \text{on } \partial\mathcal{D} \end{cases}$$

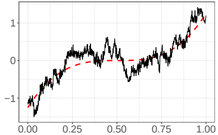
where $\mathcal{L}$ is a *linear* (differential) operator, and u is a source term, and $\mathcal{B}$ is a *linear* operator defining boundary/initial conditions

Assumption: Linear well-posed PDE problems

■ SPATIO-TEMPORAL MODELING WITH GAUSSIAN PROCESSES

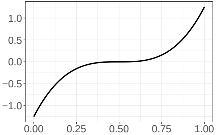
Gaussian Process (GP)		QoI across space
$\mathcal{Z} : \{\mathcal{Z}(x) : x \in \mathcal{D}\}$	$\underrightarrow{\text{Realization}}$	$Z : \{Z(x) : x \in \mathcal{D}\}$
High correlation		High "similarity"

$\rightarrow$ Allows to model non i.i.d data while only specifying the first two moments



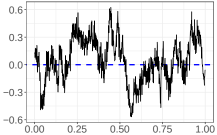
Gaussian Process $\mathcal{Z}$

=



Expectation / Deterministic effects
Deterministic trend $\mu(t, \boldsymbol{s})$
(eg. Regression over some covariates)

+



Residuals / Random effects
Zero-mean and characterized by a **covariance kernel** $C_{\mathcal{Z}}$:

$$\boxed{\mathrm{Cov}(Z(x), Z(x')) = C_{\mathcal{Z}}(x, x')}$$

SETTING

QoI Physical variable Z defined across time $[0,T]$ and/or $\mathcal{D}$

Data (Noisy) Observations $\boldsymbol{Y} = (Y(x_1),\dots,Y(x_{N_D}))^T$ of Z at $x_1,\dots,x_{N_D} \in \mathcal{D}$

Model QoI $Z \to$ GP $\mathcal{Z}$ and Data $Y \to \mathcal{Y}$ where

$$\boxed{\mathcal{Y}(x_i) = \mathcal{Z}(x_i) + \tau\varepsilon_i}, \quad i \in \{1,\dots,k\}$$

$\to \varepsilon_1,\dots,\varepsilon_{N_D} \sim \mathcal{N}(0,1)$ iid noise

Goal (Probabilistic) Predictions of Y at unobserved locations

$\to$ Vector of observations $\boldsymbol{Y}$ can be written

$$\boxed{\boldsymbol{Y} = \boldsymbol{Z} + \tau\boldsymbol{\varepsilon}}, \quad \boldsymbol{\varepsilon} \sim \mathcal{N}(0,\boldsymbol{I})$$

where $\boldsymbol{Z} = (\mathcal{Z}(x_1),\dots,\mathcal{Z}(x_n))^T$ is a Gaussian vector $\Rightarrow$ Explicit likelihood:

$$L(\boldsymbol{\theta}) = \frac{1}{2}\log|\boldsymbol{\Sigma}(\boldsymbol{\theta}) + \tau^2\boldsymbol{I}| - \frac{1}{2}(\boldsymbol{Y} - \boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{\Sigma}(\boldsymbol{\theta}) + \tau^2\boldsymbol{I}\right)^{-1}(\boldsymbol{Y} - \boldsymbol{\mu}(\boldsymbol{\theta})) + \text{ Constant}$$

where $\boldsymbol{\mu} = \mathbb{E}[\boldsymbol{Z}]$ and $\boldsymbol{\Sigma} = \left[C_{\mathcal{Z}}(x_k,x_l)\right]_{1\le k,l\le N_D}$

SETTING

QoI Physical variable Z defined across time $[0,T]$ and/or $\mathcal{D}$

Data (Noisy) Observations $\boldsymbol{Y} = (Y(x_1), \dots, Y(x_{N_D}))^T$ of Z at $x_1, \dots, x_{N_D} \in \mathcal{D}$

Model QoI $Z \to$ GP $\mathcal{Z}$ and Data $Y \to \mathcal{Y}$ where

$$\boxed{\mathcal{Y}(x_i) = \mathcal{Z}(x_i) + \tau\varepsilon_i}, \quad i \in \{1, \dots, k\}$$

$\to \varepsilon_1, \dots, \varepsilon_{N_D} \sim \mathcal{N}(0,1)$ iid noise

Goal (Probabilistic) Predictions of Y at unobserved locations

$\to$ Closed-form conditional distributions $\big(\mathcal{Z}(x')|\boldsymbol{Y}\big) \sim \mathcal{N}(m(x'), \sigma^2(x'))$ where

$$m(x') = \sum_{k=1}^{N_D} \left[\left(\boldsymbol{\Sigma} + \tau^2 \boldsymbol{I}\right)^{-1} \boldsymbol{Y}\right]_k C_{\mathcal{Z}}(x_k, x'), \quad \text{with } \boldsymbol{\Sigma} = \left[C_{\mathcal{Z}}(x_k, x_l)\right]_{1 \le k,l \le N_D}$$

$$\sigma^2(x') = C_{\mathcal{Z}}(x', x') - \sum_{k=1}^{N_D} \sum_{l=1}^{N_D} C_{\mathcal{Z}}(x', x_k) \left[\left(\boldsymbol{\Sigma} + \tau^2 \boldsymbol{I}\right)^{-1}\right]_{kl} C_{\mathcal{Z}}(x_l, x')$$

OUTLINE

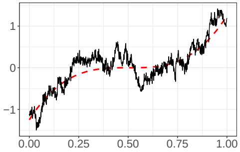
I. Covariance-based approaches

II. SPDE approach

III. Challenges in inference

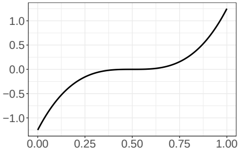
PHYSICALLY-MOTIVATED MEAN

- First possibility: Set mean of GP as output of physical model



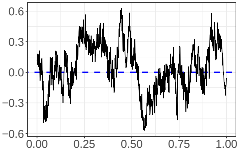
Variable of interest $Y(s,t)$

=



Mean depending on physical model $\mu(s,t)$

+

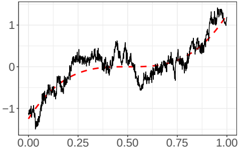


Residuals $Z(s,t)$

- Examples: Regression over several physical model outputs
- Residuals often chosen with simple structure: fully independent or independent in time, stationary in space
- Ex: Applications in predictive maintenance (Zhang et al., 2022; Jin et al., 2019)

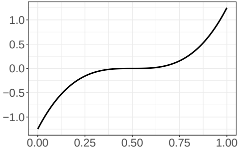
■ PHYSICALLY-CONSTRAINED COVARIANCE

- Second possibility: Constrain the covariance kernel to yield physically realistic GPs



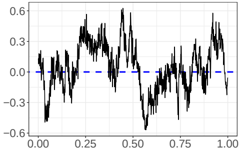
Variable of interest $Y(s,t)$

=



Simple mean (often zero) $\mu(s,t)$

+



Residuals with complex spatio-temporal correlations $Z(s,t)$

- Question: How do we pick the covariance kernel?

■ KERNEL MAGIC

Consider a GP $\mathcal{Z}$ with covariance kernel

$$C_{\mathcal{Z}}(x,y) = \mathrm{Cov}(\mathcal{Z}(x), \mathcal{Z}(y))$$

- (Under some regularity assumptions) The derivative of a GP is a GP, and

$$\mathrm{Cov}(\frac{\partial}{\partial x}\mathcal{Z}(x), \frac{\partial}{\partial y}\mathcal{Z}(y)) = \frac{\partial}{\partial x}\frac{\partial C}{\partial y}(x,y)$$

- More generally, when applying a linear operator to a GP, we get a GP :

$$\mathrm{Cov}(\mathcal{L}_x\mathcal{Z}(x), \mathcal{L}_y\mathcal{Z}(y)) = (\mathcal{L}_x\mathcal{L}_y C)(x,y)$$

  Moreover, the "vector" $(\mathcal{L}\mathcal{Z}, \mathcal{Z})$ is also a GP

$\rightarrow$ Use this to "constraint" kernel into incorporating physical constraints

■ COVARIANCE KERNEL WITH PHYSICAL CONSTRAINTS

Recall the physical constraint on the QoI

$$\mathcal{L}Z = u$$

with $\mathcal{L}$ a linear differential operator, u a source term, and additional boundary/initial conditions

- Idea: Replace Z by a GP $\mathcal{Z}$ with a fixed kernel $C_{\mathcal{Z}}$ (eg. $C_{\mathcal{Z}}(x,y) = \exp(-\|x-y\|^2)$) :

$$\mathcal{L}\mathcal{Z} = \mathcal{U}$$

$\Rightarrow$ The kernels of $\mathcal{U}$ and of the pair $(\mathcal{U},\mathcal{Z})$ (cross-covariance) can be computed from $\mathcal{L}$ and $C_{\mathcal{Z}}$:

$$C_{\mathcal{U}}(x,y) = \mathcal{L}_x\mathcal{L}_y C_{\mathcal{Z}}(x,y), \quad C_{\mathcal{U},\mathcal{Z}}(x,y) = \mathcal{L}_x C_{\mathcal{Z}}(x,y)$$

COVARIANCE KERNEL WITH PHYSICAL CONSTRAINTS

Recall the physical constraint on the QoI

$$\mathcal{L}Z = u$$

with $\mathcal{L}$ a linear differential operator, u a source term, and additional boundary/initial conditions

- Idea: Replace Z by a GP $\mathcal{Z}$ with a fixed kernel $C_{\mathcal{Z}}$ (eg. $C_{\mathcal{Z}}(x,y) = \exp(-\|x-y\|^2)$) :

$$\mathcal{L}\mathcal{Z} = \mathcal{U}$$

$\Rightarrow$ The kernels of $\mathcal{U}$ and of the pair $(\mathcal{U}, \mathcal{Z})$ (cross-covariance) can be computed from $\mathcal{L}$ and $C_{\mathcal{Z}}$:

$$C_{\mathcal{U}}(x,y) = \mathcal{L}_x\mathcal{L}_y C_{\mathcal{Z}}(x,y), \quad C_{\mathcal{U},\mathcal{Z}}(x,y) = \mathcal{L}_x C_{\mathcal{Z}}(x,y)$$

- Then, we can "enforce" the PDE by adequately conditioning the pair $(\mathcal{Z}, \mathcal{U})$
  - $\rightarrow$ Create "data" to enforce **(or not)** the right-hand side : $\mathcal{U}(y_i) = u(y_i)$, $y_1, \dots, y_p \in \mathcal{D}$
  - $\rightarrow$ Create "data" to enforce potential initial/boundary conditions

■ COVARIANCE KERNEL WITH PHYSICAL CONSTRAINTS

Recall the physical constraint on the QoI

$$\mathcal{L}Z = u$$

with $\mathcal{L}$ a linear differential operator, u a source term, and additional boundary/initial conditions

- Idea: Replace Z by a GP $\mathcal{Z}$ with a fixed kernel $C_{\mathcal{Z}}$ (eg. $C_{\mathcal{Z}}(x,y) = \exp(-\|x-y\|^2)$) :

$$\mathcal{L}\mathcal{Z} = \mathcal{U}$$

  $\Rightarrow$ The kernels of $\mathcal{U}$ and of the pair $(\mathcal{U},\mathcal{Z})$ (cross-covariance) can be computed from $\mathcal{L}$ and $C_{\mathcal{Z}}$:

$$C_{\mathcal{U}}(x,y) = \mathcal{L}_x\mathcal{L}_y C_{\mathcal{Z}}(x,y), \quad C_{\mathcal{U},\mathcal{Z}}(x,y) = \mathcal{L}_x C_{\mathcal{Z}}(x,y)$$

- Then, we can "enforce" the PDE by adequately conditioning the pair $(\mathcal{Z},\mathcal{U})$
  $\rightarrow$ Create "data" to enforce **(or not)** the right-hand side : $\mathcal{U}(y_i) = u(y_i)$, $y_1,\dots,y_p \in \mathcal{D}$
  $\rightarrow$ Create "data" to enforce potential initial/boundary conditions

- Finally, compute a posteriori distribution of $\mathcal{Z}$ (and of $\mathcal{U}$!) given "real" data $\boldsymbol{Y}$ and artificial (PDE-enforcing) data

COVARIANCE KERNEL WITH PHYSICAL CONSTRAINTS

Recall the physical constraint on the QoI

$$\mathcal{L}Z = u$$

with $\mathcal{L}$ a linear differential operator, u a source term, and additional boundary/initial conditions
Assumption : We can compute the solution operator $\mathcal{L}^{-1}$: Source term $\mapsto$ Solution of the PDE (eg. Green's function, spectral approach)

- Alternative approach: Fix the kernel of $\mathcal{U}$ instead (eg. $C_{\mathcal{U}}(x, y) = \exp(-\|x-y\|^2)$) and use

$$\mathcal{L}\mathcal{Z} = \mathcal{U} \quad \leftrightarrow \quad \mathcal{Z} = \mathcal{L}^{-1}\mathcal{U}$$

  where $\mathcal{L}^{-1}$ denotes the solution operator associated $\Rightarrow$ The kernels of $\mathcal{Z}$ and of the pair $(\mathcal{U}, \mathcal{Z})$ (cross-covariance) can be computed from $\mathcal{L}^{-1}$ and $C_{\mathcal{Z}}$

COVARIANCE KERNEL WITH PHYSICAL CONSTRAINTS

Recall the physical constraint on the QoI

$$\mathcal{L}Z = u$$

with $\mathcal{L}$ a linear differential operator, u a source term, and additional boundary/initial conditions
Assumption : We can compute the solution operator $\mathcal{L}^{-1}$: Source term $\mapsto$ Solution of the PDE (eg. Green's function, spectral approach)

- Alternative approach: Fix the kernel of $\mathcal{U}$ instead (eg. $C_{\mathcal{U}}(x,y) = \exp(-\|x-y\|^2)$) and use

$$\mathcal{L}\mathcal{Z} = \mathcal{U} \quad \leftrightarrow \quad \mathcal{Z} = \mathcal{L}^{-1}\mathcal{U}$$

  where $\mathcal{L}^{-1}$ denotes the solution operator associated $\Rightarrow$ The kernels of $\mathcal{Z}$ and of the pair $(\mathcal{U}, \mathcal{Z})$ (cross-covariance) can be computed from $\mathcal{L}^{-1}$ and $C_{\mathcal{Z}}$

- Note that
  - PDE parameters can now be part of the inference: Part of kernel parameters
  - Putting a prior on $\mathcal{U}$ can be seen as a way of dealing with unknown source terms
  - Enforce PDE constraint on $\mathcal{Z}$ or on the pair $(\mathcal{Z}, \mathcal{U})$ ($\rightarrow$ Physics-based multivariate model)

COVARIANCE KERNEL WITH PHYSICAL CONSTRAINTS

Recall the physical constraint on the QoI

$$\mathcal{L}Z = u$$

with $\mathcal{L}$ a linear differential operator, u a source term, and additional boundary/initial conditions
Assumption : We can compute the solution operator $\mathcal{L}^{-1}$: Source term $\mapsto$ Solution of the PDE (eg. Green's function, spectral approach)

- Examples of applications
  - Approximate solutions of PDEs (Raissi et al., 2017)
  - Estimation of geostrophic and tropical winds (Berliner, 2003; Wikle et al., 2001) $\rightarrow$ Gradient constraints on components
  - Concentrations of proteins in drosophila (López-Lopera et al., 2019) $\rightarrow$ Reaction diffusion equation

■ COMMENTS ABOUT THE COVARIANCE-BASED APPROACH

Starting from the relation $\mathcal{L}\mathcal{Z} = \mathcal{U}$

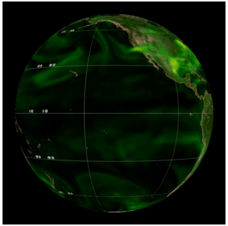
Challenges when putting the GP prior on $\mathcal{Z}$

- The regularity of $\mathcal{Z}$ depends on the choice of $C_{\mathcal{Z}}$ (eg. Gaussian kernel $\rightarrow$ Infinitely differentiable paths)
- Need to come up with sensible kernels for the QoI: Challenging when
  - Non-stationarity / Transport phenomena
  - Non-Euclidean geometries for the spatial domain

Challenges when putting the GP prior on $\mathcal{U}$

- The solution operator is not rarely available in closed-form

$\rightarrow$ Alternative approach: Solve the Stochastic PDE numerically



(Source: NASA)

OUTLINE

I. Covariance-based approaches

II. SPDE approach

III. Challenges in inference

THE SPDE APPROACH

Basic idea: if $\mathcal{Z}$ is an isotropic Markovian field over $\mathbb{R}^d$, then it is **equivalently** characterized by (Whittle, 1954; Rozanov, 1977):

Spectral density $\Gamma_{\mathcal{Z}} = \mathscr{F}[C_{\mathcal{Z}}]$

$$\Gamma_{\mathcal{Z}} : \boldsymbol{\xi} \in \mathbb{R}^d \mapsto \frac{1}{P(\|\boldsymbol{\xi}\|^2)}$$

Stochastic partial differential equation (SPDE)

$$\boxed{P(-\Delta)^{1/2}\mathcal{Z} = \mathcal{W}}$$

- $\mathcal{W}$: Gaussian white noise
- $P(-\Delta)^{1/2}\mathcal{Z} := \mathscr{F}^{-1}\left[\boldsymbol{\xi} \mapsto \sqrt{P(\|\boldsymbol{\xi}\|^2)} \times \mathscr{F}[\mathcal{Z}](\boldsymbol{\xi})\right]$

where P is a **polynomial**, strictly positive over $\mathbb{R}_+$

THE SPDE APPROACH

Basic idea: if $\mathcal{Z}$ is an isotropic Markovian field over $\mathbb{R}^d$, then it is **equivalently** characterized by (Whittle, 1954; Rozanov, 1977):

Spectral density $\Gamma_{\mathcal{Z}} = \mathscr{F}[C_{\mathcal{Z}}]$

$$\Gamma_{\mathcal{Z}} : \boldsymbol{\xi} \in \mathbb{R}^d \mapsto \frac{1}{P(\|\boldsymbol{\xi}\|^2)}$$

Stochastic partial differential equation (SPDE)

$$\boxed{P(-\Delta)^{1/2}\mathcal{Z} = \mathcal{W}}$$

- $\mathcal{W}$: Gaussian white noise
- $P(-\Delta)^{1/2}\mathcal{Z} := \mathscr{F}^{-1}\left[\boldsymbol{\xi} \mapsto \sqrt{P(\|\boldsymbol{\xi}\|^2)} \times \mathscr{F}[\mathcal{Z}](\boldsymbol{\xi})\right]$

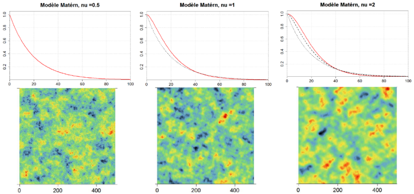
where P is a **polynomial**, strictly positive over $\mathbb{R}_+$

→ In particular, if $P(x) = (\kappa^2 + x)^\alpha$, i.e. with the SPDE

$$(\kappa^2 - \Delta)^{\alpha/2}\mathcal{Z} = \mathcal{W}$$

then $\mathcal{Z}$ has a Matérn covariance function

$$C_{\mathcal{Z}}(x, x+h) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}\big(\kappa\|h\|\big)^\nu \mathcal{K}_\nu\big(\kappa\|h\|\big), \ (\nu = \alpha - d/2)$$



THE SPDE APPROACH

Basic idea: if $\mathcal{Z}$ is an isotropic Markovian field over $\mathbb{R}^d$, then it is **equivalently** characterized by (Rozanov, 1977):

Spectral density

$$\Gamma_{\mathcal{Z}} : \boldsymbol{\xi} \in \mathbb{R}^d \mapsto \frac{1}{P(\|\boldsymbol{\xi}\|^2)}$$

Stochastic partial differential equation (SPDE)

$$\boxed{P(-\Delta)^{1/2}\mathcal{Z} = \mathcal{W}}$$

- $\mathcal{W}$: Gaussian white noise
- $P(-\Delta)^{1/2}\mathcal{Z} := \mathscr{F}^{-1}\left[\boldsymbol{\xi} \mapsto \sqrt{P(\|\boldsymbol{\xi}\|^2)} \times \mathscr{F}[\mathcal{Z}](\boldsymbol{\xi})\right]$

where P is a **polynomial**, strictly positive over $\mathbb{R}_+$

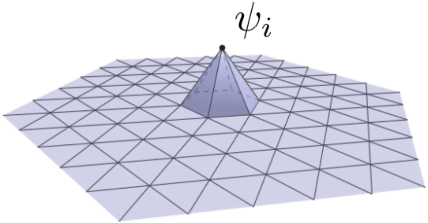
SPDE approach (Lindgren et al., 2011, 2022)

- See GPs as solutions of SPDEs (rather than through their covariance kernel)
- Solve SPDE numerically with finite elements (or finite volume) method to sample GP / characterize its distribution

FINITE ELEMENT APPROXIMATION OF SPDES

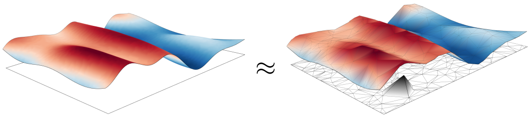
Goal: Find an approximation of the solution to: $P(-\Delta)^{1/2}\mathcal{Z} = \mathcal{W}$

Triangulation of the domain



Piecewise linear approximation of the solution

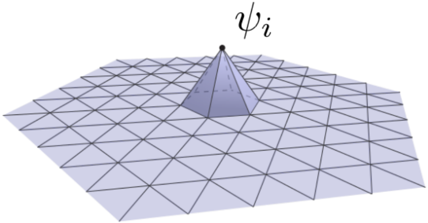
$$\mathcal{Z}(\boldsymbol{p}) \approx \widehat{\mathcal{Z}}(\boldsymbol{p}) = \sum_{i=1}^{n} Z_i \psi_i(\boldsymbol{p})$$



■ FINITE ELEMENT APPROXIMATION OF SPDES

Goal: Find an approximation of the solution to: $P(-\Delta)^{1/2}\mathcal{Z} = \mathcal{W}$

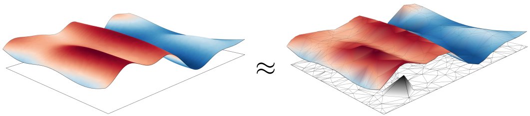
Triangulation of the domain



→

Piecewise linear approximation of the solution

$$\mathcal{Z}(\boldsymbol{p}) \approx \widehat{\mathcal{Z}}(\boldsymbol{p}) = \sum_{i=1}^{n} Z_i \psi_i(\boldsymbol{p})$$



The weights $\boldsymbol{Z} = (Z_1, \ldots, Z_n)$ form a (centered) Gaussian vector whose **precision matrix** $\boldsymbol{Q}_{\mathcal{Z}}$ has an explicit formula

→ Depends on known sparse matrices

GAUSSIAN MARKOV RANDOM FIELD APPROXIMATION

Goal: Find an approximation of the solution to: $\boxed{P(-\Delta)^{1/2}\mathcal{Z} = \mathcal{W}}$

Mass matrix (→ Diagonal)

$$\boldsymbol{C} = \begin{pmatrix} \langle \psi_1, 1\rangle & & \\ & \ddots & \\ & & \langle \psi_n, 1\rangle \end{pmatrix}$$

Stiffness matrix (→ Sparse)

$$\boldsymbol{R} = \begin{pmatrix} \ddots & \vdots & \iddots \\ & \langle \nabla\psi_i, \nabla\psi_j\rangle & \\ \iddots & \vdots & \ddots \end{pmatrix}$$

and set

$$\boxed{\boldsymbol{S} = \boldsymbol{C}^{-1/2}\boldsymbol{R}\boldsymbol{C}^{-1/2}}$$

The weights $\boldsymbol{Z} = (Z_1, \dots, Z_n)$ satisfy form a (centered) Gaussian vector whose **precision matrix** $\boldsymbol{Q}_{\mathcal{Z}}$

$$\boxed{\boldsymbol{Q}_{\mathcal{Z}} = \boldsymbol{C}^{1/2}P(\boldsymbol{S})\boldsymbol{C}^{1/2}}$$

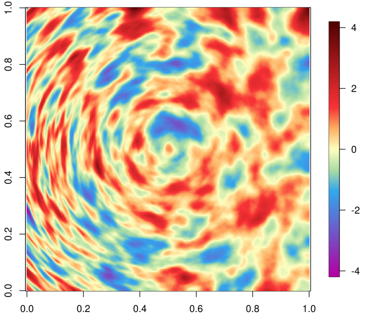
→ $\boldsymbol{Q}_{\mathcal{Z}}$ is sparse (when $\deg P$ is relatively small)

STRAIGHTFORWARD EXTENSIONS

Non-stationarity

Define SPDE with spatially-varying coefficients

$$(\kappa^2(\cdot) - \mathrm{div}(H(\cdot)\nabla))^\alpha \mathcal{Z} = \mathcal{W}$$



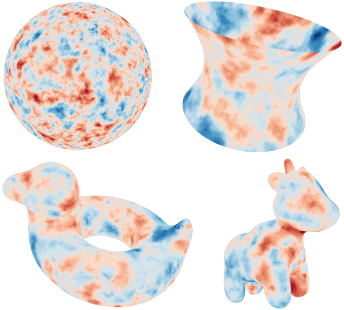
Simulation of non-stationary GP using the SPDE approach

Non-Euclidean domain

Define SPDE on Riemannian manifolds $(\mathcal{D}, g)$

$$P(-\Delta_g)\mathcal{Z} = \mathcal{W}$$

where $-\Delta_g$ is the Laplace–Beltrami operator



Simulations of GPs on smooth surfaces using the SPDE approach

STRAIGHTFORWARD EXTENSIONS

Galerkin-Chebyshev approximation of random fields (Pereira et al., 2022; Lang and Pereira, 2021)

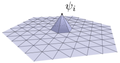
Field

$$\mathcal{Z} = \gamma(\mathcal{L})\mathcal{W},$$

where $|\gamma(\lambda)| = \mathcal{O}_{\lambda\to\infty}(|\lambda|^{-\beta})$ with $\beta > d/4$

Finite element approximation

$$\mathcal{Z} = \sum_{i=1}^{n} Z_i \psi_i$$



Weights of the approximation

$$\boxed{\boldsymbol{Z} = \boldsymbol{C}^{-1/2}\gamma(\boldsymbol{S})\boldsymbol{W}}, \qquad \text{where } \boldsymbol{S} = \boldsymbol{C}^{-1/2}\boldsymbol{R}\boldsymbol{C}^{-1/2}$$

$$\boldsymbol{C} = \left[\langle\psi_i, \psi_j\rangle_{L^2(\mathcal{M})}\right], \boldsymbol{R} = \left[\langle\mathcal{L}\psi_i, \psi_j\rangle_{L^2(\mathcal{M})}\right]$$

$\rightarrow$ Computed with polynomial approximation

BACK TO OUR INITIAL QUESTION

Recall the physical constraint on the QoI

$$\mathcal{L}Z = u$$

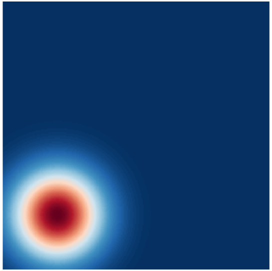
with $\mathcal{L}$ a linear differential operator, u a source term, and additional boundary/initial conditions

- Idea: Replace the right-hand side with a GP $\mathcal{U}$ (usually Gaussian white noise) and solve the SPDE using finite element / volume method
- Result: Finite element approximation of the SPDE solution

$$\boxed{\mathcal{Z}(\boldsymbol{p}) \approx \widehat{\mathcal{Z}}(\boldsymbol{p}) = \sum_{i=1}^{n} Z_i \psi_i(\boldsymbol{p})}$$

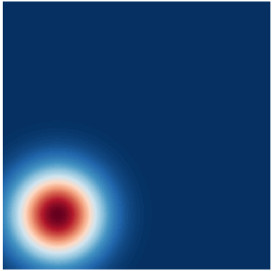
weights $\boldsymbol{Z} = (Z_1, \dots, Z_n)$ form a (centered) Gaussian vector with precision matrix $\boldsymbol{Q}_{\mathcal{Z}}$ depending on the $\mathcal{L}$

EXAMPLE: SOME SIMPLE TRANSPORT PHENOMENA



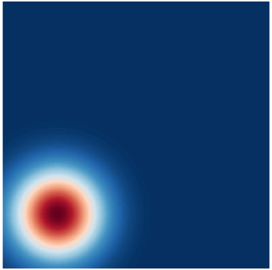
Advection

$$\frac{\partial z}{\partial t} + \vec{v} \cdot \nabla z = 0$$



Diffusion

$$\frac{\partial z}{\partial t} - \Delta z = 0$$



Advection + Diffusion

$$\frac{\partial z}{\partial t} + \vec{v} \cdot \nabla z - \Delta z = 0$$

EXAMPLE: DIFFUSION PROCESS

Physical constraint on the QoI:

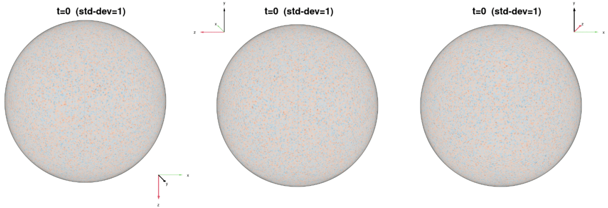
$$\mathcal{L}Z = \frac{\partial Z}{\partial t} + (\kappa^2 - \Delta)^{\alpha/2} Z = u$$

with u a source term, and additional boundary/initial conditions

→ Consider the diffusion SPDE (Bakka et al., 2020; Rayner et al., 2020):

$$\frac{\partial \mathcal{Z}}{\partial t} + (\kappa^2 - \Delta)^{\alpha/2} \mathcal{Z} = \tau \mathcal{W}_T \otimes \mathcal{W}_S,$$

where $\mathcal{W}_T \otimes \mathcal{Y}_S$ is a white noise in time and colored noise in space



■ EXAMPLE: ADVECTION-DIFFUSION PROCESS

Physical constraint on the QoI:

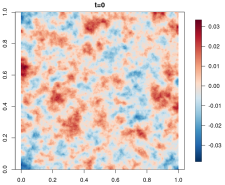
$$\mathcal{L}Z = \frac{\partial Z}{\partial t} + \frac{1}{c}\left((\kappa^2 - \Delta)^\alpha Z + \gamma \cdot \nabla \mathcal{Z}\right) = u$$

with u a source term, and additional boundary/initial conditions

→ Consider the advection-diffusion SPDE:

$$\frac{\partial \mathcal{Z}}{\partial t} + \frac{1}{c}\left((\kappa^2 - \Delta)^\alpha \mathcal{Z} + \gamma \cdot \nabla \mathcal{Z}\right) = \frac{\tau}{\sqrt{c}} \mathcal{W}_T \otimes \mathcal{Y}_S$$

where $\mathcal{W}_T \otimes \mathcal{Y}_S$ is a white noise in time and colored noise in space, and γ an advection vector



- Clouding covering estimation (Clarotto et al., 2022)
- Precipitation and tropical thunderstorm modeling (Sigrist et al., 2015; Liu et al., 2022)
- Air pollution (Chen et al., 2023)
- Ocean salinity data (Berild and Fuglstad, 2024)

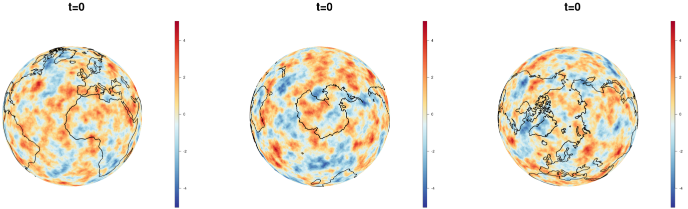
EXAMPLE: ADVECTION-DIFFUSION PROCESS

Physical constraint on the QoI:

$$\mathcal{L}Z = \frac{\partial Z}{\partial t} + \frac{1}{c}\left((\kappa^2 - \Delta)^\alpha Z + \gamma \cdot \nabla \mathcal{Z}\right) = u$$

with u a source term, and additional boundary/initial conditions

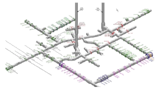
$\rightarrow$ Application: Model pollutant dispersion at a global scale : *L. Clarotto's talk [O.A4.1]*

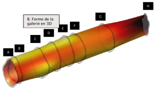


Solution of an advection-diffusion SPDE with advection defined by wind data (ECMWF Atmospheric Composition Reanalysis 4, every 3 hours, 01/12/2024 - 03/12/2024)

■ EXAMPLE: COUPLED SYSTEM OF SPDE

Project with the French National Agency for Nuclear Waste : Study of deformations underground nuclear waste storage galleries





- QoI : Temperature and Deformations over gallery surface
- Low number of captors (32 per section, max)

Physical constraints on the QoIs : Temperature T, Radial Deformation Z

$$\begin{cases} \mathcal{L}_T T &= \frac{\partial Z}{\partial t} - \alpha \Delta T = u_T \\ \mathcal{L}_Z Z &= \Delta^2 Z \beta_1 \Delta Z - \beta_2 \Delta T = u_Z \end{cases}$$

with u_T, u_Z (unknown) source terms, and additional boundary/initial conditions

Example of joint simulation of Temperature and Deformation with the SPDE approach

OUTLINE

I. Covariance-based approaches

II. SPDE approach

III. Challenges in inference

■ COVARIANCE BASED APPROACHES

Vector of observations $\boldsymbol{Y}$ can be expressed as

$$\boldsymbol{Y} = \boldsymbol{Z} + \tau\boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim \mathcal{N}(0, \boldsymbol{I})$$

where $\boldsymbol{Z}$ is a Gaussian vector $\Rightarrow$ Gaussian likelihood expensive too compute!

$$L(\boldsymbol{\theta}) = \frac{1}{2}\log|\boldsymbol{\Sigma}(\boldsymbol{\theta}) + \tau^2\boldsymbol{I}| - \frac{1}{2}(\boldsymbol{Y} - \boldsymbol{\mu}(\boldsymbol{\theta}))^T(\boldsymbol{\Sigma}(\boldsymbol{\theta}) + \tau^2\boldsymbol{I})^{-1}(\boldsymbol{Y} - \boldsymbol{\mu}(\boldsymbol{\theta})) + \text{ Constant}$$

Several approaches to speed up computations

- Covariance tapering $\rightarrow$ Sparse matrices (Kaufman et al., 2008)
- Low-rank approximations / Inducing points methods (Hensman et al., 2013; Datta et al., 2016)
- Vecchia approximation (Katzfuss and Guinness, 2021)

$\rightarrow$ The GP distribution is approximated to allow inference

SPDE-BASED APPROACHES

Vector of observations $\boldsymbol{Y}$ can be expressed as

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{Z} + \tau\boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim \mathcal{N}(0, \boldsymbol{I})$$

where $\boldsymbol{Z}$ is a Gaussian vector, $\boldsymbol{A}$ a design matrix $\Rightarrow$ Gaussian likelihood expensive too compute!

$$L(\boldsymbol{\theta}) = \frac{1}{2}\log|\boldsymbol{A}\boldsymbol{Q}(\boldsymbol{\theta})\boldsymbol{A}^T + \tau^2\boldsymbol{I}| - \frac{1}{2}(\boldsymbol{Y} - \boldsymbol{\mu}(\boldsymbol{\theta}))^T\left(\boldsymbol{A}\boldsymbol{Q}(\boldsymbol{\theta})\boldsymbol{A}^T + \tau^2\boldsymbol{I}\right)^{-1}(\boldsymbol{Y} - \boldsymbol{\mu}(\boldsymbol{\theta})) + \text{ Constant}$$

Several approaches to speed up computations

- Restriction on parameters / SPDEs considered $\rightarrow$ Sparse matrices (Lindgren et al., 2011)
- Stochastic trace estimation (Pereira et al., 2022)
- Design of preconditionners (Antil and Saibaba, 2024)

$\rightarrow$ The GP distribution is approximated to allow inference

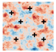
MACHINE LEARNING TO THE RESCUE?

Setting: Likelihood evaluations / Inference can be cumbersome, but simulating is often easy
$\rightarrow$ Can we "learn", with neural networks, ways to infer model parameters from spatial data?

Two useful remarks

- By sampling from some prior $\theta \sim \pi(\theta)$ and then simulating data $\boldsymbol{Y} \sim \pi(\boldsymbol{Y}|\theta)$
  $\rightarrow$ We can create (many!) samples $(\boldsymbol{Y}; \theta)$ from the joint distribution $\pi(\boldsymbol{Y}, \theta)$

$$(\boldsymbol{Y}\,;\,\theta\,)$$



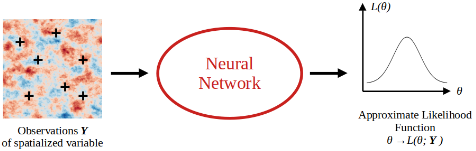
- We can approximate (expectations of) KL-divergences involving the posterior distribution using these samples

MACHINE LEARNING TO THE RESCUE?

Setting: Likelihood evaluations / Inference can be cumbersome, but simulating is often easy
$\rightarrow$ Can we "learn", with neural networks, ways to infer model parameters from spatial data?

First approach: Learn a way to approximate the likelihood of the model

(Dutta et al., 2016; Walchessen et al., 2024)



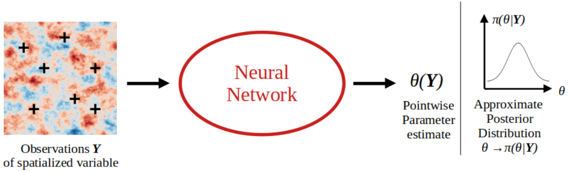
- Use Likelihood-to-Evidence ratio $\pi(\boldsymbol{Y}|\theta)/\pi(\boldsymbol{Y})$ to link likelihoood to a classification probability
  $\rightarrow$ Only need to learn a classifier
- Learn variational approximation of likelihood $L(\theta; \boldsymbol{Y}) \approx q(\theta; \kappa(\boldsymbol{Y}))$

■ MACHINE LEARNING TO THE RESCUE?

Setting: Likelihood evaluations / Inference can be cumbersome, but simulating is often easy
$\rightarrow$ Can we "learn", with neural networks, ways to infer model parameters from spatial data?

Second approach: Learn a way to get approximate inferred parameters

(Mnih and Gregor, 2014; Lenzi et al., 2023; Sainsbury-Dale et al., 2024)

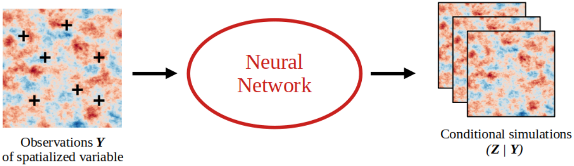


- Given the pairs $(\boldsymbol{Y}, \theta)$, learn a functional relationship between the pair by minimizing a loss function
- Learn variational approximation of the posterior distribution $\pi(\theta|\boldsymbol{Y}) \approx q(\theta; \kappa(\boldsymbol{Y}))$ by minimizing expected KL-divergences

MACHINE LEARNING TO THE RESCUE?

Setting: Likelihood evaluations / Inference can be cumbersome, but simulating is often easy
$\rightarrow$ Can we "learn", with neural networks, ways to infer model parameters from spatial data?

Third approach: Generative model approach



- Idea: Some generative models are "easy" to condition to observations and naturally include variability of prior distribution
- Learn a neural approximation of whole data distribution $\pi(\boldsymbol{Z}) = \int \pi(\boldsymbol{Z}, \theta) d\theta$ and use conditioning algorithm to get samples $\pi(\boldsymbol{Y}|\theta)$
- Application with Diffusion Models (Cardoso and Pereira, 2025) $\rightarrow$ *G. V. Cardoso's talk [O.A4.3]*

TAKEAWAY MESSAGES

- Many tools to incorporate (PDE-based) physical constraints into GP models for spatial statistics
  - Through derivation of covariance kernels
  - By solving SPDEs

- Still work to do for practical application
  - Inference can be challenging
  - Hope through ML based approaches

THANK YOU FOR YOUR ATTENTION!

https://mike-pereira.github.io/

REFERENCES

Antil, H. and Saibaba, A. K. (2024). Efficient algorithms for bayesian inverse problems with whittle–matérn priors. *SIAM Journal on Scientific Computing*, 46(2):S176–S198.

Bakka, H., Krainski, E., Bolin, D., Rue, H., and Lindgren, F. (2020). The diffusion-based extension of the Matérn field to space-time.

Berild, M. O. and Fuglstad, G.-A. (2024). Non-stationary spatio-temporal modeling using the stochastic advection–diffusion equation. *Spatial Statistics*, 64:100867.

Berliner, L. M. (2003). Physical-statistical modeling in geophysics. *Journal of Geophysical Research: Atmospheres*, 108(D24).

Cardoso, G. V. and Pereira, M. (2025). Predictive posterior sampling from non-stationnary gaussian process priors via diffusion models with application to climate data. *arXiv preprint arXiv:2505.24556*.

Chen, J., Miao, C., Yang, D., Liu, Y., Zhang, H., and Dong, G. (2023). Estimation of fine-resolution pm2.5 concentrations using the inla-spde method. *Atmospheric Pollution Research*, 14(7):101781.

REFERENCES

Clarotto, L., Allard, D., Romary, T., and Desassis, N. (2022). The spde approach for spatio-temporal datasets with advection and diffusion. *arXiv preprint arXiv:2208.14015*.

Datta, A., Banerjee, S., Finley, A. O., and Gelfand, A. E. (2016). On nearest-neighbor gaussian process models for massive spatial data. *Wiley Interdisciplinary Reviews: Computational Statistics*, 8(5):162–171.

Dutta, R., Corander, J., Kaski, S., and Gutmann, M. U. (2016). Likelihood-free inference by ratio estimation. *arXiv preprint arXiv:1611.10242*.

Hensman, J., Fusi, N., and Lawrence, N. D. (2013). Gaussian processes for big data. *arXiv preprint arXiv:1309.6835*.

Jin, X., Ni, J., et al. (2019). Physics-based gaussian process for the health monitoring for a rolling bearing. *Acta astronautica*, 154:133–139.

Katzfuss, M. and Guinness, J. (2021). A general framework for vecchia approximations of gaussian processes. *Statistical Science*, 36(1):124–141.

■ REFERENCES

Kaufman, C. G., Schervish, M. J., and Nychka, D. W. (2008). Covariance tapering for likelihood-based estimation in large spatial data sets. *Journal of the American Statistical Association*, 103(484):1545–1555.

Lang, A. and Pereira, M. (2021). Galerkin–Chebyshev approximation of Gaussian random fields on compact Riemannian manifolds. *arXiv preprint arXiv:2107.02667*.

Lenzi, A., Bessac, J., Rudi, J., and Stein, M. L. (2023). Neural networks for parameter estimation in intractable models. *Computational Statistics & Data Analysis*, 185:107762.

Lindgren, F., Bolin, D., and Rue, H. (2022). The SPDE approach for Gaussian and non-Gaussian fields: 10 years and still running. *Spatial Statistics*, 50:100599.

Lindgren, F., Rue, H., and Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498.

Liu, X., Yeo, K., and Lu, S. (2022). Statistical modeling for spatio-temporal data from stochastic convection-diffusion processes. *Journal of the American Statistical Association*, 117(539):1482–1499.

REFERENCES

López-Lopera, A. F., Durrande, N., and Alvarez, M. A. (2019). Physically-inspired gaussian process models for post-transcriptional regulation in drosophila. *IEEE/ACM transactions on computational biology and bioinformatics*, 18(2):656–666.

Mnih, A. and Gregor, K. (2014). Neural variational inference and learning in belief networks. In *International Conference on Machine Learning*, pages 1791–1799. PMLR.

Pereira, M., Desassis, N., and Allard, D. (2022). Geostatistics for large datasets on riemannian manifolds: A matrix-free approach. *Journal of Data Science*, 20(4):512–532.

Raissi, M., Perdikaris, P., and Karniadakis, G. E. (2017). Machine learning of linear differential equations using gaussian processes. *Journal of Computational Physics*, 348:683–693.

Rayner, N. A., Auchmann, R., Bessembinder, J., Brönnimann, S., Brugnara, Y., Capponi, F., Carrea, L., Dodd, E. M., Ghent, D., Good, E., et al. (2020). The eustace project: Delivering global, daily information on surface air temperature. *Bulletin of the American Meteorological Society*, 101(11):E1924–E1947.

Rozanov, J. A. (1977). Markov random fields and stochastic partial differential equations. *Mathematics of the USSR-Sbornik*, 32(4):515.

REFERENCES

Sainsbury-Dale, M., Zammit-Mangion, A., and Huser, R. (2024). Likelihood-free parameter estimation with neural bayes estimators. *The American Statistician*, 78(1):1–14.

Sigrist, F., Künsch, H. R., and Stahel, W. A. (2015). Stochastic partial differential equation based modelling of large space–time data sets. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 77(1):3–33.

Walchessen, J., Lenzi, A., and Kuusela, M. (2024). Neural likelihood surfaces for spatial processes with computationally intensive or intractable likelihoods. *Spatial Statistics*, 62:100848.

Whittle, P. (1954). On stationary processes in the plane. *Biometrika*, pages 434–449.

Wikle, C. K., Milliff, R. F., Nychka, D., and Berliner, L. M. (2001). Spatiotemporal hierarchical bayesian modeling tropical ocean surface winds. *Journal of the american statistical association*, 96(454):382–397.

Zhang, J., Liu, C., and Gao, R. X. (2022). Physics-guided gaussian process for hvac system performance prognosis. *Mechanical Systems and Signal Processing*, 179:109336.