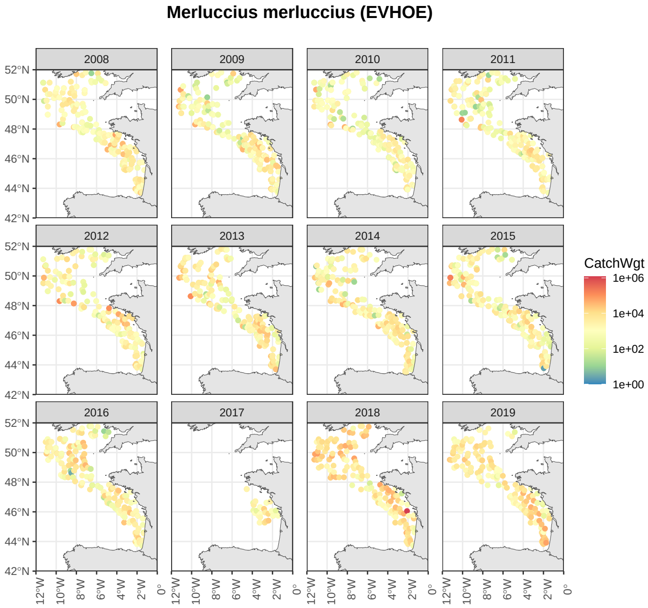
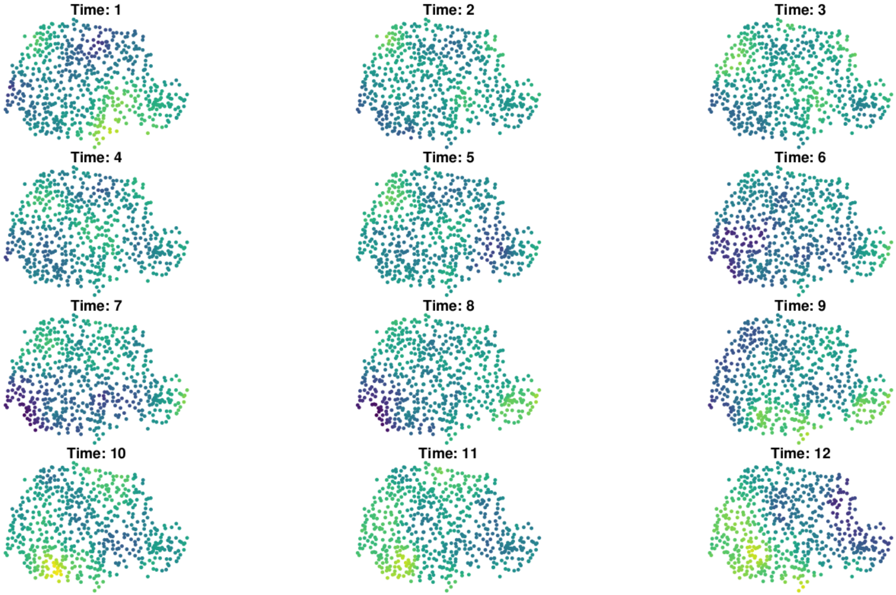
Introduction to spatial and spatio-temporal statistics for ecology

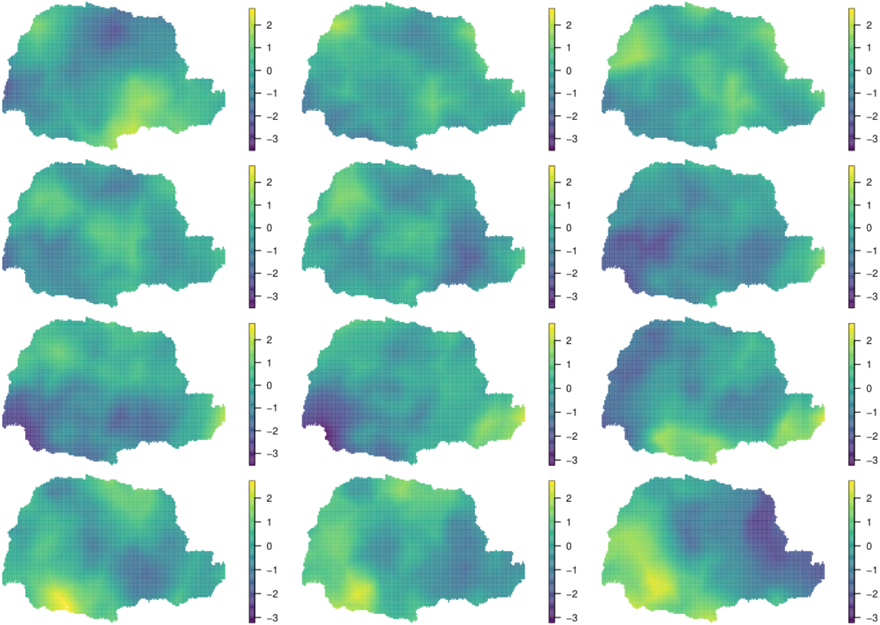
Baptiste Alglave, Jean-Baptiste Lecomte, Maxime Olmos

February 2026



Merluccius merluccius (EVHOE)
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
52°N
50°N
48°N
46°N
44°N
42°N
12°W
10°W
8°W
6°W
4°W
2°W
0°
CatchWgt
1e+06
1e+04
1e+02
1e+00


Time: 1
Time: 2
Time: 3
Time: 4
Time: 5
Time: 6
Time: 7
Time: 8
Time: 9
Time: 10
Time: 11
Time: 12


2 1 0 −1 −2 −3
2 1 0 −1 −2 −3
2 1 0 −1 −2 −3
2 1 0 −1 −2 −3
2 1 0 −1 −2 −3
2 1 0 −1 −2 −3
2 1 0 −1 −2 −3
2 1 0 −1 −2 −3
2 1 0 −1 −2 −3
2 1 0 −1 −2 −3
2 1 0 −1 −2 −3
2 1 0 −1 −2 −3

Data and ecological processes

What are the characteristics of these data?

- They arise from the ecological process of interest (= they are **conditionnal** on the ecological process)
- **Noisy** (= they are not perfect observations of the ecological process)
- **Sparse** (= they do not cover the full area and some hypothesis need to be set to predict the ecological process between the sampled locations)

What are the characteristics of the ecological process we want to infer?

- **Hidden** or **latent**
- **Structured** (relations that structure the process?)

How to relate the data to the ecological process?

➠ **Hierarchical modeling**

Data and ecological processes

What are the characteristics of these data?

- They arise from the ecological process of interest (= they are **conditionnal** on the ecological process)
- **Noisy** (= they are not perfect observations of the ecological process)
- **Sparse** (= they do not cover the full area and some hypothesis need to be set to predict the ecological process between the sampled locations)

What are the characteristics of the ecological process we want to infer?

- **Hidden** or **latent**
- **Structured** (relations that structure the process?)

How to relate the data to the ecological process?

➠ **Hierarchical modeling**

Data and ecological processes

What are the characteristics of these data?

- They arise from the ecological process of interest (= they are **conditionnal** on the ecological process)
- **Noisy** (= they are not perfect observations of the ecological process)
- **Sparse** (= they do not cover the full area and some hypothesis need to be set to predict the ecological process between the sampled locations)

What are the characteristics of the ecological process we want to infer?

- **Hidden** or **latent**
- **Structured** (relations that structure the process?)

How to relate the data to the ecological process?

➠ **Hierarchical modeling**

Data and ecological processes

What are the characteristics of these data?

- They arise from the ecological process of interest (= they are **conditionnal** on the ecological process)
- **Noisy** (= they are not perfect observations of the ecological process)
- **Sparse** (= they do not cover the full area and some hypothesis need to be set to predict the ecological process between the sampled locations)

What are the characteristics of the ecological process we want to infer?

- **Hidden** or **latent**
- **Structured** (relations that structure the process?)

How to relate the data to the ecological process?

➠ **Hierarchical modeling**

Hierarchical models

Let's define observations $\boldsymbol{Y}$:

$$\boldsymbol{Y}|\boldsymbol{S}, \boldsymbol{\theta} \sim \mathcal{L}_Y(\boldsymbol{S}, \boldsymbol{\theta}_{obs})$$

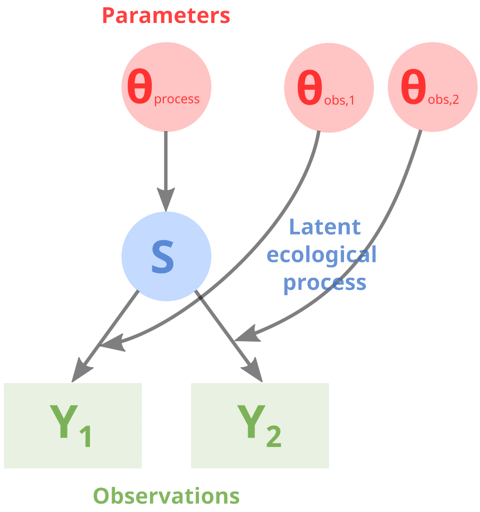
a latent field $\boldsymbol{S}$:

$$\boldsymbol{S}|\boldsymbol{\theta}_{process} \sim \mathcal{L}_S(\boldsymbol{\theta}_{process})$$

and parameters:

$$\boldsymbol{\theta} = (\boldsymbol{\theta}_{obs}, \boldsymbol{\theta}_{process})$$

There can be several data sources $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$, in which case each has its own probability distribution ($\mathcal{L}_{Y_1}$,$\mathcal{L}_{Y_2}$) and observation parameters ($\boldsymbol{\theta}_{obs,1}$, $\boldsymbol{\theta}_{obs,2}$).



➠ **How to specify such model for space-time applications?**

1. Univariate spatial models

2. Multivariate spatio-temporal models

3. Model inference

Univariate spatial models

$$\boldsymbol{Y}|\delta, \beta \sim \mathcal{L}(\boldsymbol{S}, \sigma^2)$$

$$f(\boldsymbol{S}) = \boldsymbol{X} \cdot \beta + \boldsymbol{\delta}$$

$$\boldsymbol{\delta} \sim \mathcal{MG}(0, \boldsymbol{\Sigma})$$

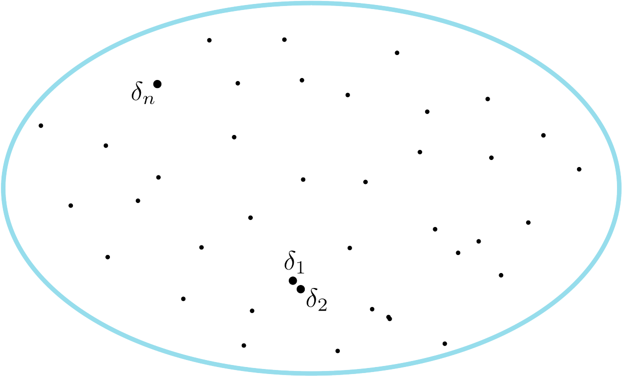
The key element is the covariance matrix $\boldsymbol{\Sigma}_{(n \times n)}$

or its inverse, the precision matrix $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1}$

Notations:

$\boldsymbol{Y} = (Y_i, i \in \{1, \cdots, n\})$ the observation vector
$(x_1, \cdots, x_n)$ the related locations
$\boldsymbol{S} = (S(x_i), i \in \{1, \cdots, n\})$ the latent field variables
$\boldsymbol{\delta} = (\delta(x_i), i \in \{1, \cdots, n\})$ the (latent) random effects
$\boldsymbol{\beta} = (\beta_j, j \in \{1, \cdots, p\})$ the parameter vector
$\boldsymbol{X}_{(n \times p)}$ the covariate matrix
f a link function.

$$\boldsymbol{\Sigma} = \begin{pmatrix} Var(\delta(x_1)) & Cov(\delta(x_1), \delta(x_2)) & \cdots & Cov(\delta(x_1), \delta(x_n)) \\ Cov(\delta(x_2), \delta(x_1)) & Var(\delta(x_2)) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ Cov(\delta(x_n), \delta(x_1)) & \cdots & \cdots & Var(\delta(x_n)) \end{pmatrix}$$

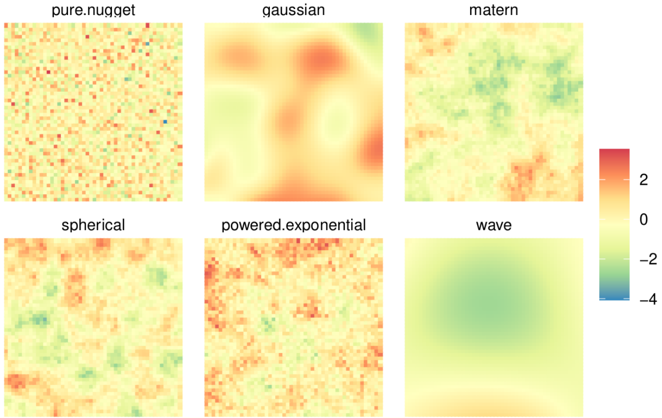


➠ **The covariance is stationary (and isotropic)**

$$Cov(\delta(x_i), \delta(x_j)) = \mathcal{C}(h_{ij}) \text{ with } h_{ij} \in \mathbb{R}$$

h_{ij} is the distance between x_i and x_j

$\mathcal{C}(h_{ij})$ is the covariance function



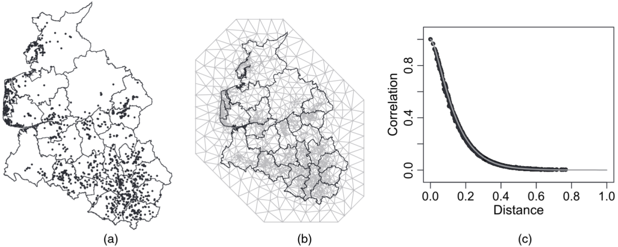


Fig. 2. (a) Locations of leukaemia survival observations, (b) triangulation using 3446 triangles and (c) a stationary correlation function (———) and the corresponding GMRF approximation (•) for $\nu = 1$ and approximate range 0.26

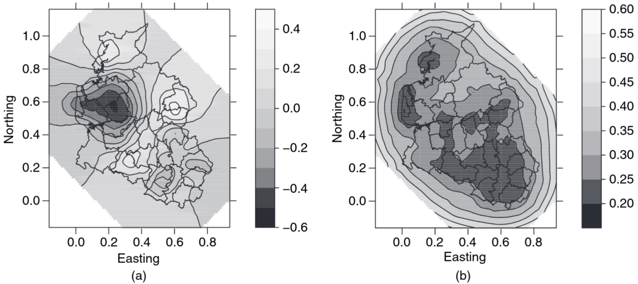
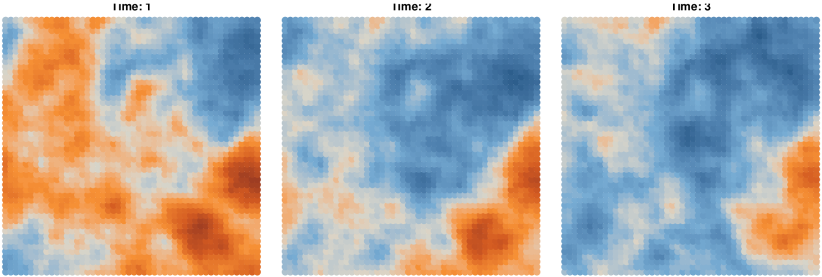


Fig. 3. (a) Posterior mean and (b) standard deviation of the spatial effect on survival by using the GMRF representation

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.

Moving to spatio-temporal models

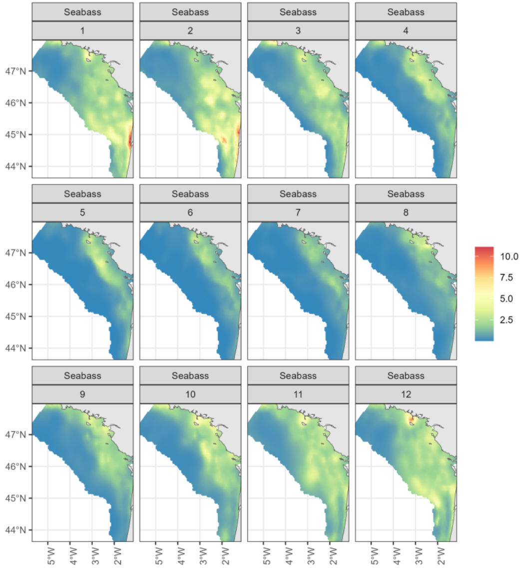
The ideas are similar, but we add temporal correlations in the expression of the random effect $\boldsymbol{\delta}$.



Let's introduce the model:

$$\delta(x, t) = \varphi \cdot \delta(x, t-1) + \omega(x, t) \text{ for } t = 2, \dots, T \text{ and } x \in \mathbb{R}^2$$

- $\varphi \in]-1; 1[$ is the autoregressive temporal term
- $\omega(x, t)$ is a purely spatial GRF



Average monthly distribution of sea bass in the Bay of Biscay from 2008 to 2018

Alglave Baptiste, Olmos et al. (2024). Investigating fish reproduction phenology and essential habitats by identifying the main spatio-temporal patterns of fish distribution.

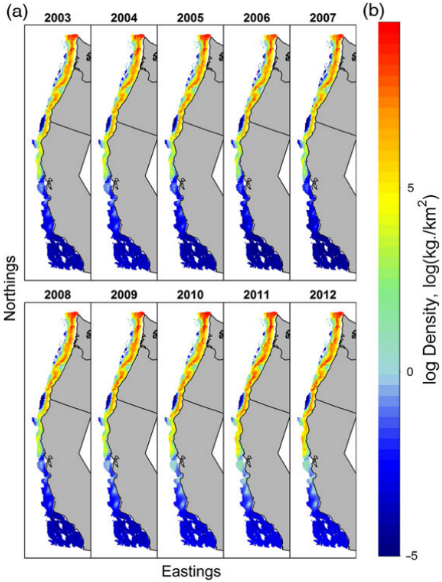


Figure 1. Density for arrowtooth flounder 2003 – 2012, estimated by the geostatistical delta-generalized linear mixed model (note that the white space in southern California represents the cowcod conservation area, which prohibits trawl gears including the survey design and hence is excluded when estimating spatial densities and abundance indices).

Thorson, J. T., Shelton, A. O., Ward, E. J., and Skaug, H. J. (2015). Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES Journal of Marine Science, 72(5), 1297-1310.

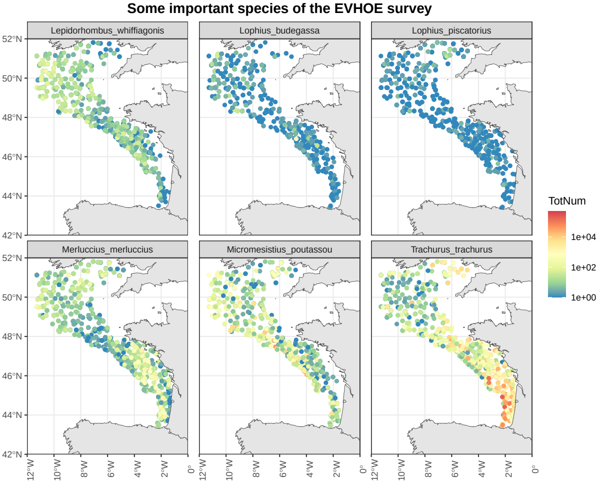
Multivariate spatio-temporal models

Multivariate modelling

Often, spatial data are multivariate

e.g. in survey data, several species are recorded in a single haul.

➠ **How to extract the common patterns between these species?**



How to model jointly these species?

Let's now denote $\boldsymbol{S}_{(n \times l)}$ the matrix of observations of species in n locations for l species.

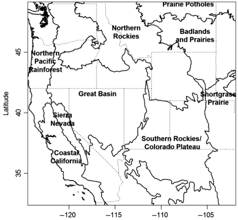
$$\log(\boldsymbol{S}_{i,\cdot}) = \boldsymbol{X} \cdot \beta + \boldsymbol{L\Delta}_{i,\cdot}^T$$

$\rightarrow \log(\boldsymbol{S}_{i,\cdot})$ is the log-expected densities for all species l at locations $i \in \{1, \cdots, n\}$.

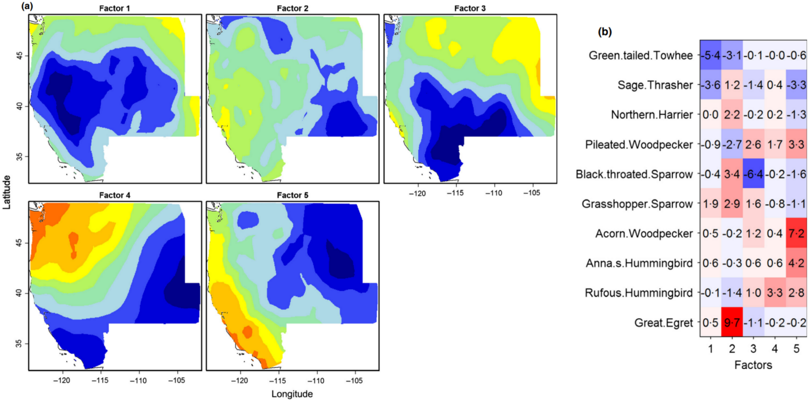
$\rightarrow \boldsymbol{X} \cdot \beta$ is the trend part.

$\rightarrow \boldsymbol{L\Delta}_{i,\cdot}^T$ is the part of the model that controls species (or variables) interaction.

➠ How is it parameterized?



Results for the spatial factor analysis model applied to breeding bird survey data for 10 species in 2013:
(a) the estimated factors after varimax rotation,
(b) the varimax-rotated loadings matrix.



Thorson, J. T., Scheuerell, M. D., Shelton, A. O., See, K. E., Skaug, H. J., & Kristensen, K. (2015). Spatial factor analysis: a new tool for estimating joint species distributions and correlations in species range. Methods in Ecology and Evolution, 6(6), 627-637.

Model inference

How to infer such model?

Standard methods are not efficient.

The keystone is the likelihood:

$$L_M(\boldsymbol{\theta}) = P(\boldsymbol{Y}|\boldsymbol{\theta}) = \int_{\mathbb{R}^q} P(\boldsymbol{Y}, \delta|\boldsymbol{\theta}) d\delta$$

$\boldsymbol{Y}$ are the observations, $\boldsymbol{\theta}$ are the parameters, δ are the latent random variables.

In a spatio-temporal context q can be very high

➠ require efficient numerical methods (1) to bypass the integration step, (2) to reduce the dimensionality of δ and (3) to derive efficiently the likelihood.

Three methods to enhance computing:

- Laplace approximation (= approximation of the likelihood)
- SPDE approach (= approximation of the Gaussian random effect)
- Automatic differentiation (= efficient derivation technics)

R package: `R-INLA` for bayesian inference (Rue et al., 2017) and `TMB` for maximum likelihood inference (Kristensen et al., 2015).

How to infer such model?

Standard methods are not efficient.

The keystone is the likelihood:

$$L_M(\theta) = P(Y|\theta) = \int_{\mathbb{R}^q} P(Y, \delta|\theta) d\delta$$

Y are the observations, θ are the parameters, δ are the latent random variables.

In a spatio-temporal context q can be very high

➟ require efficient numerical methods (1) to bypass the integration step, (2) to reduce the dimensionality of δ and (3) to derive efficiently the likelihood.

Three methods to enhance computing:

- Laplace approximation (= approximation of the likelihood)
- SPDE approach (= approximation of the Gaussian random effect)
- Automatic differentiation (= efficient derivation technics)

R package: `R-INLA` for bayesian inference (Rue et al., 2017) and `TMB` for maximum likelihood inference (Kristensen et al., 2015).

How to infer such model?

Standard methods are not efficient.

The keystone is the likelihood:

$$L_M(\boldsymbol{\theta}) = P(\boldsymbol{Y}|\boldsymbol{\theta}) = \int_{\mathbb{R}^q} P(\boldsymbol{Y}, \delta|\boldsymbol{\theta}) d\delta$$

$\boldsymbol{Y}$ are the observations, $\boldsymbol{\theta}$ are the parameters, δ are the latent random variables.

In a spatio-temporal context q can be very high

➠ require efficient numerical methods (1) to bypass the integration step, (2) to reduce the dimensionality of δ and (3) to derive efficiently the likelihood.

Three methods to enhance computing:

- Laplace approximation (= approximation of the likelihood)
- SPDE approach (= approximation of the Gaussian random effect)
- Automatic differentiation (= efficient derivation technics)

R package: `R-INLA` for bayesian inference (Rue et al., 2017) and `TMB` for maximum likelihood inference (Kristensen et al., 2015).

How to infer such model?

Standard methods are not efficient.

The keystone is the likelihood:

$$L_M(\boldsymbol{\theta}) = P(\boldsymbol{Y}|\boldsymbol{\theta}) = \int_{\mathbb{R}^q} P(\boldsymbol{Y}, \delta|\boldsymbol{\theta}) d\delta$$

$\boldsymbol{Y}$ are the observations, $\boldsymbol{\theta}$ are the parameters, δ are the latent random variables.

In a spatio-temporal context q can be very high

➟ require efficient numerical methods (1) to bypass the integration step, (2) to reduce the dimensionality of δ and (3) to derive efficiently the likelihood.

Three methods to enhance computing:

- Laplace approximation (= approximation of the likelihood)
- SPDE approach (= approximation of the Gaussian random effect)
- Automatic differentiation (= efficient derivation technics)

R package: `R-INLA` for bayesian inference (Rue et al., 2017) and `TMB` for maximum likelihood inference (Kristensen et al., 2015).

Towards more complex methods

Modeling population dynamics

Modeling movement

Modeling causality (structural equation models)

Take home message

Main ideas:

hierarchical modelling

spatial/temporal correlation

multivariate analysis

➠ extract structuring/shared patterns, and their variation in time

Many applications in ecology

Possible statistical developments too