



¡Escanéeme!

# INTRODUCCIÓN AL APRENDIZAJE BAYESIANO USANDO INLA

Joaquín Martínez Minaya

[jmarmin@eio.upv.es](mailto:jmarmin@eio.upv.es)

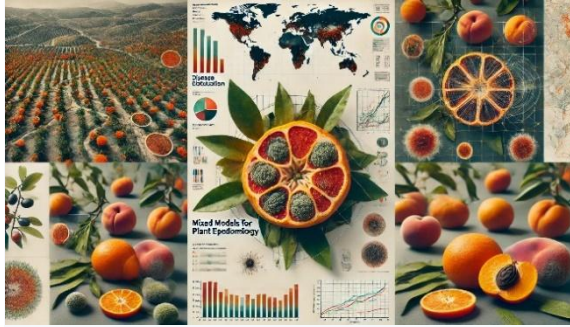
<https://jmartinez-minaya.github.io/>



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

10 de noviembre de 2025

# LÍNEAS DE INVESTIGACIÓN



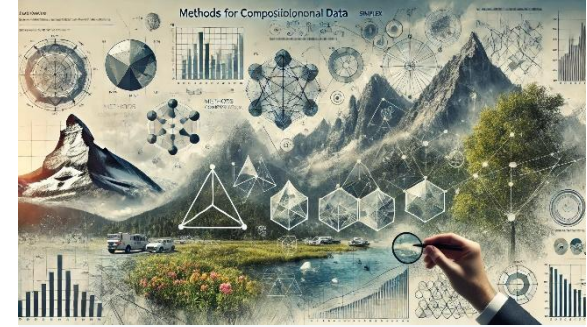
**ESTADÍSTICA ESPACIO-TEMPORAL  
EN LA MODELIZACIÓN DE LA  
DISTRIBUCIÓN DE ESPECIES**

**L1**



**APRENDIZAJE BAYESIANO EN  
SALUD, MEDIOAMBIENTE Y  
ECONOMÍA**

**L2**



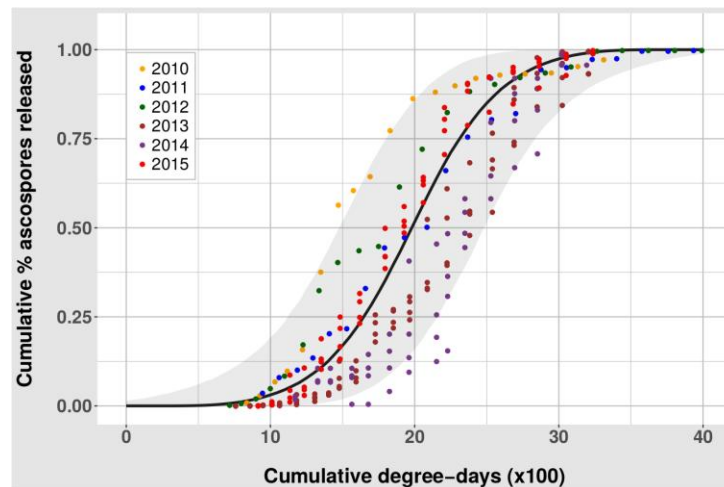
**MÉTODOS COMPUTACIONALES  
BAYESIANOS**

**L3**

- **Motivación. La mancha foliar del caqui**
- Modelos Latentes Gaussianos. INLA
- LGMs para la mancha negra de los cítricos
- L1: Estadística espacio-temporal en la modelización de la distribución de especies
- L2: Aprendizaje bayesiano en salud, medioambiente y economía
- L3: Métodos computacionales bayesianos
- Trabajos futuros

### Objetivos

- ▷ Modelizar la **producción de ascosporas** por el hongo *P. nawae* en la hojarasca del caqui
- ▷ Cultivos del caqui de la C. Valenciana afectados
- ▷ Lesiones necróticas y caída de frutos



### Variables

- ▷ Mediciones entre 2010 y 2015
- ▷ **Proporción acumulada de ascosporas**
- ▷ Covariables: Grados día acumulados (ADD)

### Regresión beta

- Modelizar el porcentaje de ascosporas liberado

#### Verosimilitud

$$y_i | \beta, u_i \sim \text{Beta}(\mu_i, \phi) ,$$

$$\text{logit}(\mu_i) = \mathbf{X}_i \beta + u_i ,$$

#### LGF

$$\beta \sim \mathcal{N}(0, \tau_b^{-1} \mathbf{I}) ,$$

$$\mathbf{u} \sim \mathcal{N}(0, \mathbf{Q}^{-1}) , \text{ donde } \mathbf{Q} = \tau_u \mathbf{I} ,$$

#### Hiperparámetros

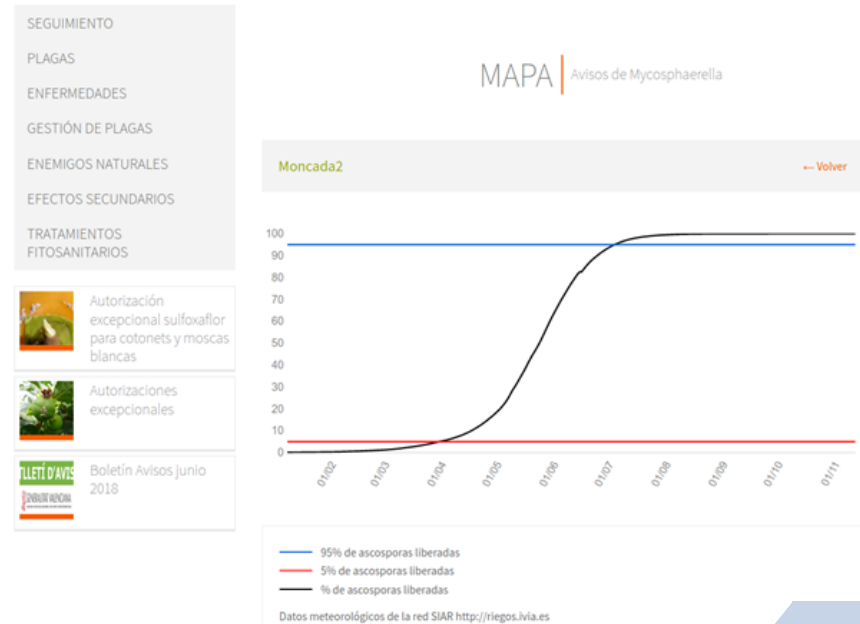
$$\sigma_u \sim \text{PC-prior}(\sigma_0, \alpha) ,$$

$$\phi \sim \text{Gamma}(a, b) ,$$

Phytopathology® • 2021 • 111:1184-1192 • <https://doi.org/10.1094/PHTO-08-20-0359-R>

### Modeling Inoculum Availability of *Plurivorosphaerella nawae* in Persimmon Leaf Litter with Bayesian Beta Regression

Joaquín Martínez-Minaya,<sup>1</sup> David Conesa,<sup>2</sup> Antonio López-Quílez,<sup>2</sup> José Luis Mira,<sup>3</sup> and Antonio Vicent<sup>3,†</sup>



- Motivación. La mancha foliar del caqui
- **Modelos Latentes Gaussianos. INLA**
- LGMs para la mancha negra de los cítricos
- L1: Estadística espacio-temporal en la modelización de la distribución de especies
- L2: Aprendizaje bayesiano en salud, medioambiente y economía
- L3: Métodos computacionales bayesianos
- Trabajos futuros

# MODELO JERÁRQUICO BAYESIANO (HBM)

## Nivel 1: Verosimilitud

$$y_i \mid \beta, \omega_i \sim p(y_i \mid \eta_i)$$

$$\eta_i = \mathbf{X}_i \beta + \omega_i$$

## Nivel 2: Parámetros y efectos aleatorios

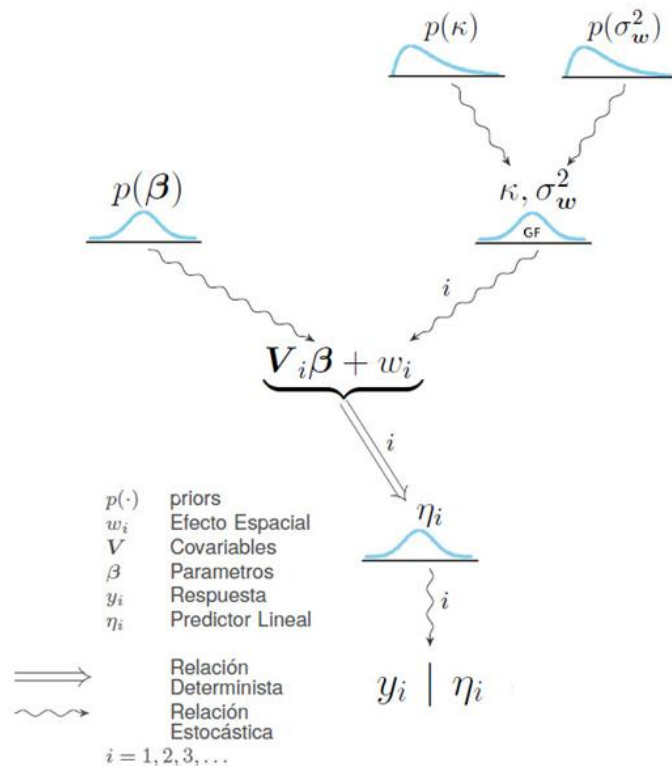
$$\beta \sim p(\beta)$$

$$\omega \sim \mathcal{N}(\mathbf{0}, \Sigma(\kappa, \sigma_w^2))$$

## Nivel 3: Hiperparámetros

$$\kappa \sim p(\kappa)$$

$$\sigma_w^2 \sim p(\sigma_w^2)$$



# COMPUTACIÓN EN HBMS

## Nivel 1: **Verosimilitud**

$$y_i \mid \beta, \omega_i \sim p(y_i \mid \eta_i)$$

$$\eta_i = \mathbf{X}_i \beta + \omega_i$$

## Nivel 2: **Parámetros y efectos aleatorios**

$$\beta \sim p(\beta)$$

$$\omega \sim \mathcal{N}(\mathbf{0}, \Sigma(\kappa, \sigma_\omega^2))$$

## Nivel 3: **Hiperparámetros**

$$\kappa \sim p(\kappa)$$

$$\sigma_\omega^2 \sim p(\sigma_\omega^2)$$

## MÉTODOS DE MONTE CARLO POR CADENAS DE MARKOV (MCMC)



# MODELOS LATENTES GAUSSIANOS, LGMs

## Nivel 1: Verosimilitud

$$y_i \mid \beta, \omega_i \sim p(y_i \mid \eta_i)$$

$$\eta_i = \mathbf{X}_i \beta + \omega_i$$

## Nivel 2: Campo latente Gaussiano (LGF)

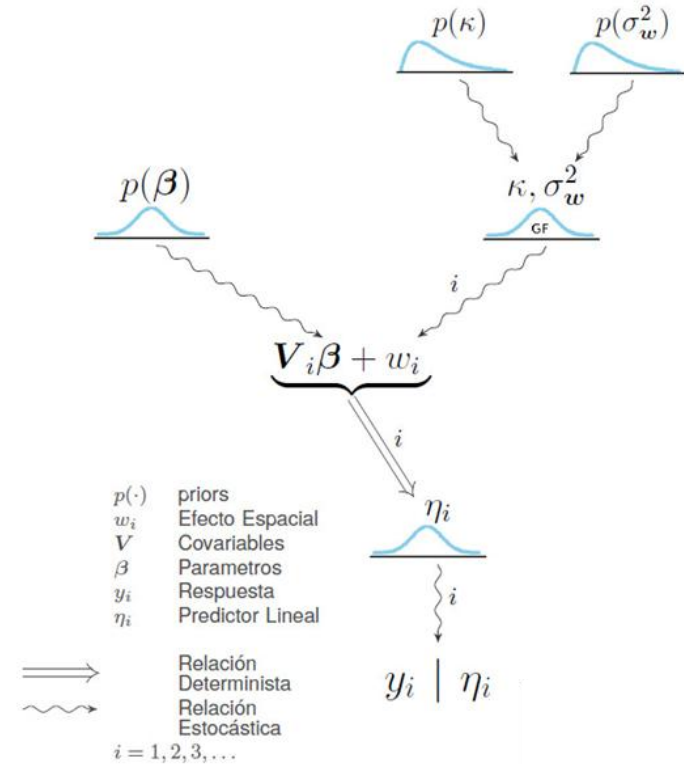
$$\beta \sim p(\beta)$$

$$\omega \sim \mathcal{N}(\mathbf{0}, \Sigma(\kappa, \sigma_\omega^2))$$

## Nivel 3: Hiperparámetros

$$\kappa \sim p(\kappa)$$

$$\sigma_\omega^2 \sim p(\sigma_\omega^2)$$



# MODELOS LATENTES GAUSSIANOS, LGMs

## Nivel 1: Verosimilitud

$$y_i \mid \beta, \omega_i \sim p(y_i \mid \eta_i)$$

$$\eta_i = \mathbf{X}_i \beta + \omega_i$$

## Nivel 2: Campo latente Gaussiano (LGF)

$$\beta \sim p(\beta)$$

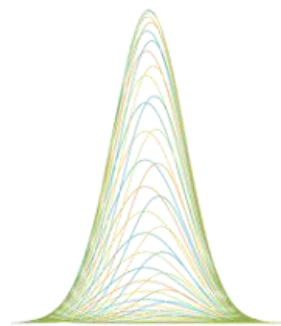
$$\omega \sim \mathcal{N}(\mathbf{0}, \Sigma(\kappa, \sigma_\omega^2))$$

## Nivel 3: Hiperparámetros

$$\kappa \sim p(\kappa)$$

$$\sigma_\omega^2 \sim p(\sigma_\omega^2)$$

## INTEGRATED NESTED LAPLACE APPROXIMATION



INLA

### Diseñadores:

- ▷ Prof. Haavard Rue, KAUST
- ▷ Prof. Finn Lindgren, Universidad de Edimburgo

- Motivación. La mancha foliar del caqui
- Modelos Latentes Gaussianos. INLA
- **LGMs para la mancha negra de los cítricos**
- L1: Estadística espacio-temporal en la modelización de la distribución de especies
- L2: Aprendizaje bayesiano en salud, medioambiente y economía
- L3: Métodos computacionales bayesianos
- Trabajos futuros

Eur J Plant Pathol (2015) 143:69–83  
DOI 10.1007/s10658-015-0666-z

## Climatic distribution of citrus black spot caused by *Phyllosticta citricarpa*. A historical analysis of disease spread in South Africa

Joaquín Martínez-Minaya • David Conesa •  
Antonio López-Quílez • Antonio Vicent

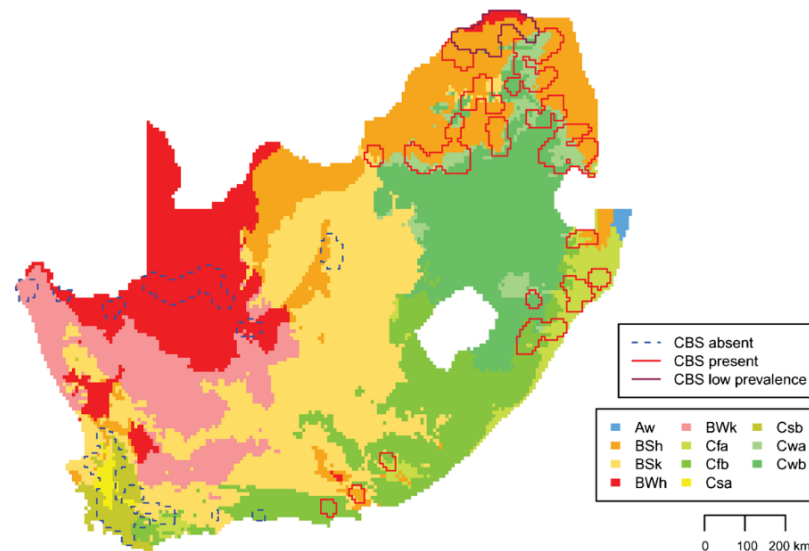
Eur J Plant Pathol (2017) 148:497–502  
DOI 10.1007/s10658-016-1056-x

## Scientific critique of the paper “Climatic distribution of citrus black spot caused by *Phyllosticta citricarpa*. A historical analysis of disease spread in South Africa” by Martínez-Minaya et al. (2015)

Eur J Plant Pathol (2017) 148:503–508  
DOI 10.1007/s10658-017-1163-3

## Response to the letter on “Climatic distribution of citrus black spot caused by *Phyllosticta citricarpa*. A historical analysis of disease spread in South Africa” by Fourie et al. (2017)

Joaquín Martínez-Minaya • David Conesa •  
Antonio López-Quílez • Antonio Vicent



Este trabajo fue seleccionado por EFSA para evaluar nueva información científica con respecto al hongo *Phyllosticta citricarpa*.

### Objetivos y variables

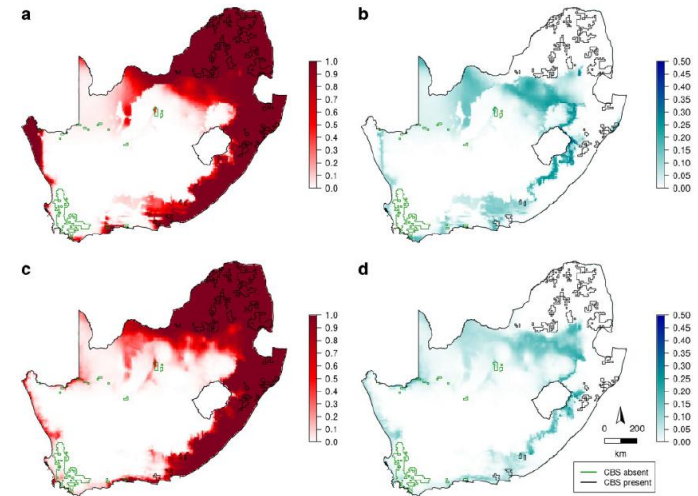
- Estudiar factores espaciales y climáticos asociados con la distribución del CBS en Sudáfrica
- **Respuesta: presencia/ausencia**
- Covariables: 21 variables climáticas



Eur J Plant Pathol (2018) 151:991–1007  
<https://doi.org/10.1007/s10658-018-1435-6>

### Spatial and climatic factors associated with the geographical distribution of citrus black spot disease in South Africa. A Bayesian latent Gaussian model approach

Joaquín Martínez-Minaya • David Conesa •  
Antonio López-Quílez • Antonio Vicent



### Objetivos y variables

- Estudiar factores espaciales y climáticos asociados con la distribución del CBS en Sudáfrica
- Respuesta: presencia/ausencia
- Covariables: 21 variables climáticas

### Verosimilitud

$$y_i | \beta, \omega_i \sim \text{Ber}(\pi_i) ,$$

$$\text{logit}(\pi_i) = X_i \beta + \omega_i ,$$

### LGF

$$\beta \sim \mathcal{N}(0, \tau_b^{-1} I) ,$$

$$\omega \sim \mathcal{N}(0, Q^{-1}(\phi, \sigma_\omega)) ,$$

### Hiperparámetros

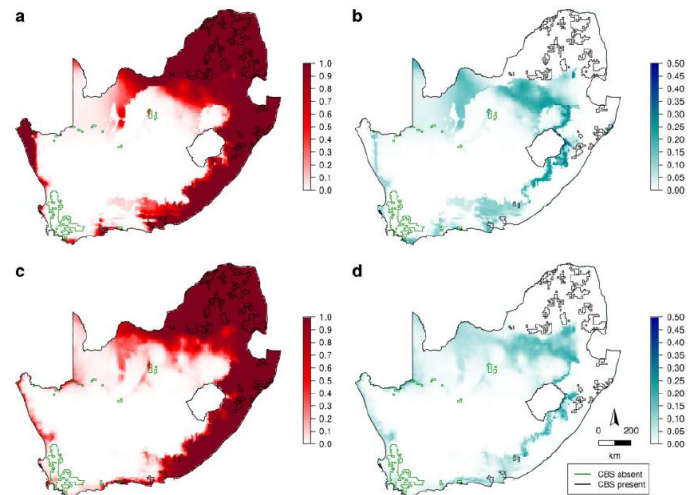
$$\sigma_\omega \sim \text{PC-prior}(\sigma_0, \alpha_1) ,$$

$$\phi \sim \text{PC-prior}(\phi_0, \alpha_2) ,$$

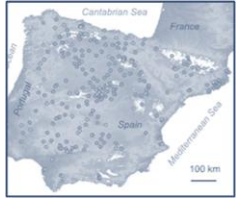
Eur J Plant Pathol (2018) 151:991–1007  
<https://doi.org/10.1007/s10658-018-1435-6>

### Spatial and climatic factors associated with the geographical distribution of citrus black spot disease in South Africa. A Bayesian latent Gaussian model approach

Joaquín Martínez-Minaya · David Conesa ·  
Antonio López-Quílez · Antonio Vicent



- Motivación. La mancha foliar del caqui
- Modelos Latentes Gaussianos. INLA
- LGMs para la mancha negra de los cítricos
- **L1: Estadística espacio-temporal en la modelización de la distribución de especies**
- L2: Aprendizaje bayesiano en salud, medioambiente y economía
- L3: Métodos computacionales bayesianos
- Trabajos futuros



## BESAG YORK Y MOLLIE

### Verosimilitud

$$y_i | \beta, u_i, v_i \sim \text{Binomial}(n_i, \pi_i),$$

$$\text{logit}(\pi_i) = X_i \beta + \frac{1}{\tau} \left( \sqrt{1 - \phi} u_i + \sqrt{\phi} v_i \right),$$

### LGF

$$\beta \sim \mathcal{N}(0, \tau_b^{-1} I),$$

$$u \sim \mathcal{N}(0, Q^{-1}), \text{ donde } Q = \tau_u I,$$

$$v_i | v_{-i} \sim \mathcal{N} \left( \frac{1}{k_i} \sum_{j \sim i} v_j, \frac{1}{\tau_v k_i} \right), \quad i \neq j,$$

### Hiperparámetros

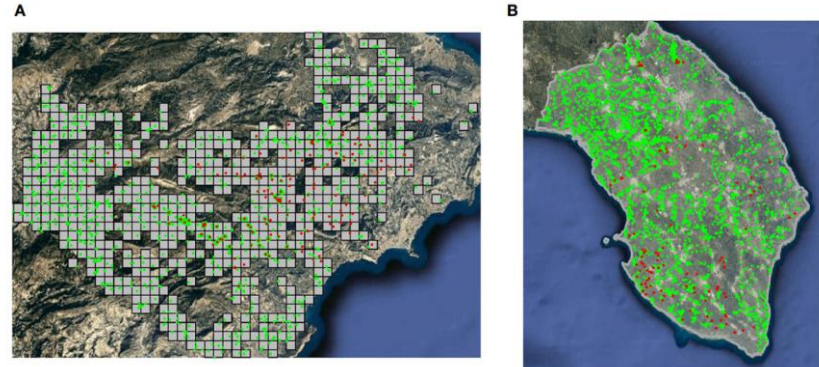
$$\sigma_u = \sqrt{1/\tau_u} \sim \text{PC-prior}(\sigma_{u0}, \alpha_1),$$

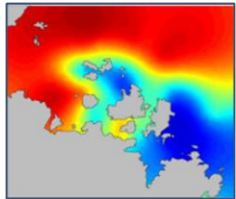
$$\sigma_v = \sqrt{1/\tau_v} \sim \text{PC-prior}(\sigma_{v0}, \alpha_2),$$

$$\phi \sim \text{PC-prior}(\phi_0, \alpha_3),$$

## Spatial Bayesian Modeling Applied to the Surveys of *Xylella fastidiosa* in Alicante (Spain) and Apulia (Italy)

Martina Cendoya<sup>1\*</sup>, Joaquín Martínez-Minaya<sup>2</sup>, Vicente Dalmau<sup>3</sup>, Amparo Ferrer<sup>3</sup>, Maria Saponari<sup>4</sup>, David Conesa<sup>5</sup>, Antonio López-Quílez<sup>5</sup> and Antonio Vicent<sup>1\*</sup>





## MODELO CON SISTEMA DE SPDES

### Verosimilitud

$$y_i | \beta, \omega_i \sim \text{Ber}(\pi_i),$$

$$\text{logit}(\pi_i) = \mathbf{X}_i \beta + \omega_i,$$

### LGF

$$\beta \sim \mathcal{N}(0, \tau_b^{-1} I),$$

$$\omega \sim \mathcal{N}(0, Q^{-1}(\phi, \sigma_\omega)),$$

### Hiperparámetros

$$\sigma_\omega \sim \text{PC-prior}(\sigma_0, \alpha_1),$$

$$\phi \sim \text{PC-prior}(\phi_0, \alpha_2),$$

$$\omega(s) - \nabla \cdot \frac{r^2}{8} \nabla \omega(s) = r \sqrt{\frac{\pi}{2}} \sigma_\omega \mathcal{W}(s), \quad \text{for } s \in \Omega_n,$$

$$\omega(s) - \nabla \cdot \frac{r_b^2}{8} \nabla \omega(s) = r_b \sqrt{\frac{\pi}{2}} \sigma_\omega \mathcal{W}(s), \quad \text{for } s \in \Omega_b.$$



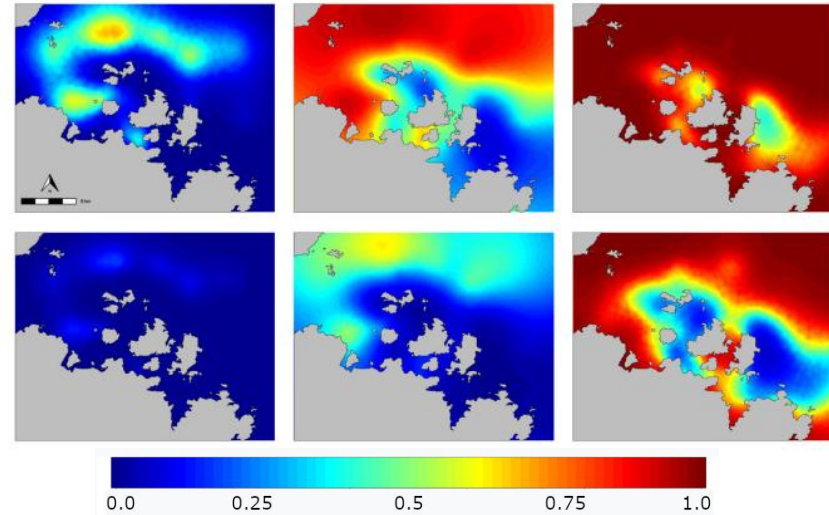
Contents lists available at ScienceDirect

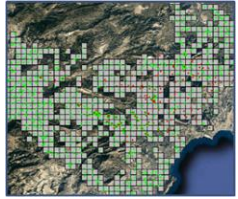
Ecological Modelling

journal homepage: [www.elsevier.com/locate/ecolmodel](http://www.elsevier.com/locate/ecolmodel)

Dealing with physical barriers in bottlenose dolphin (*Tursiops truncatus*) distribution

Joaquín Martínez-Minaya<sup>a,b,\*</sup>, David Conesa<sup>a,b</sup>, Haakon Bakka<sup>c</sup>, Maria Grazia Pennino<sup>b,d</sup>





## MODELOS HURDLE. EFECTOS ESPACIALES COMPARTIDOS

### Verosimilitud

$$Z_i | p_i \sim \text{Ber}(p_i),$$

$$\text{logit}(p_i) = \mathbf{X}_i^{(0)} \boldsymbol{\beta}^{(0)} + \omega_i,$$

$$HS_i | (Z_i = 1), \mu_i, \phi_{HS} \sim \text{Beta}(\mu_i, \phi_{HS}),$$

$$\text{logit}(\mu_i) = \mathbf{X}_i \boldsymbol{\beta} + \omega_i,$$

### LGF

$$\boldsymbol{\beta} \sim \mathcal{N}(\mathbf{0}, \tau_b^{-1} \mathbf{I}),$$

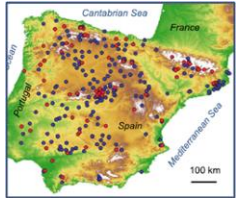
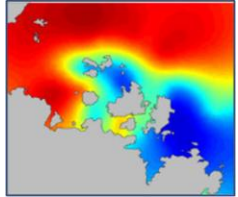
$$\boldsymbol{\omega} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}^{-1}(\phi, \sigma_\omega)),$$

### Hiperparámetros

$$\sigma_\omega \sim \text{PC-prior}(\sigma_0, \alpha_1),$$

$$\phi \sim \text{PC-prior}(\phi_0, \alpha_2)$$

$$\phi_{HS} \sim \text{Gamma}(a_\phi, b_\phi).$$

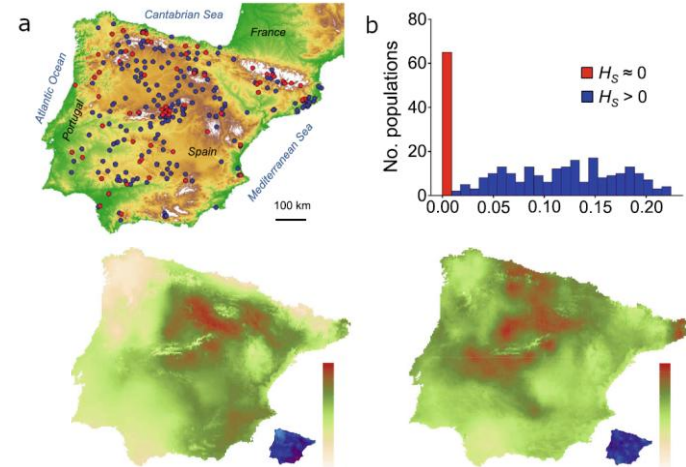


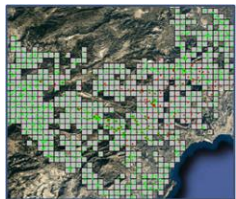
Castilla et al. *BMC Evolutionary Biology* (2020) 20:71  
<https://doi.org/10.1186/s12862-020-01635-2>

BMC Evolutionary Biology

## Ecological, genetic and evolutionary drivers of regional genetic differentiation in *Arabidopsis thaliana*

Antonio R. Castilla<sup>1,2</sup>, Belén Méndez-Vigo<sup>3</sup>, Arnald Marcer<sup>4,5</sup>, Joaquín Martínez-Minaya<sup>6</sup>, David Conesa<sup>7</sup>, F. Xavier Picó<sup>2\*</sup> and Carlos Alonso-Blanco<sup>3</sup>





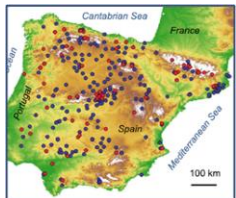
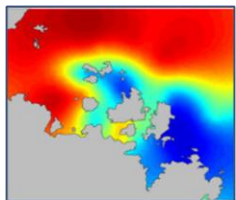
**MODELOS CON VEROSIMILITUDES UNIVARIANTES:**  
BINOMIAL, BETA, ETC.

**EXTENSIÓN AL CASO MULTIVARIANTE: DATOS COMPOSICIONALES**

- ▷ Metodología INLA no lidia directamente con distribuciones multivariantes

**LITERATURA: DOS VEROSIMILITUDES**

- ▷ Dirichlet
- ▷ Logistic-Normal

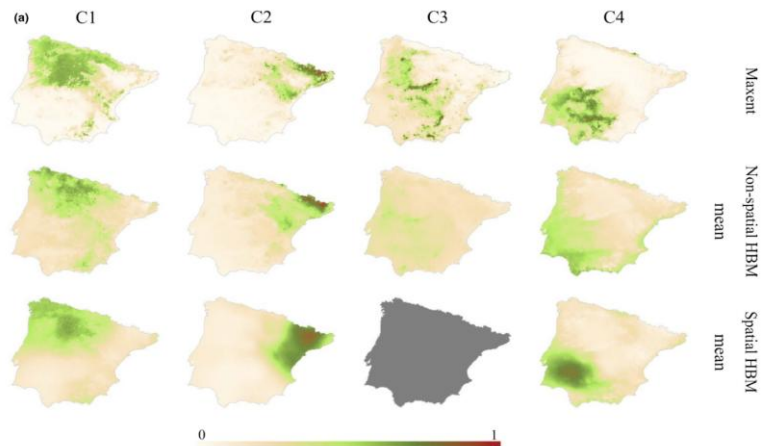


### RESOURCE ARTICLE

MOLECULAR ECOLOGY  
RESOURCES

**A hierarchical Bayesian Beta regression approach to study the effects of geographical genetic structure and spatial autocorrelation on species distribution range shifts**

Joaquín Martínez-Minaya<sup>1</sup> | David Conesa<sup>1</sup> | Marie-Josée Fortin<sup>2</sup> | Carlos Alonso-Blanco<sup>3</sup> | F. Xavier Pico<sup>4</sup> | Arnald Marcer<sup>5,6</sup>



Journal of Computational and Graphical Statistics

### The Integrated Nested Laplace Approximation for Fitting Dirichlet Regression Models

Joaquín Martínez-Minaya, Finn Lindgren, Antonio López-Quílez, Daniel Simpson & David Conesa

Statistics and Computing (2024) 34:116  
<https://doi.org/10.1007/s11222-024-10427-3>

### A flexible Bayesian tool for CoDa mixed models: logistic-normal distribution with Dirichlet covariance

Joaquín Martínez-Minaya<sup>1</sup> · Haavard Rue<sup>2</sup>

dirinla (con **Finn Lindgren** en **UE**)

- Ajuste de regresión de Dirichlet
  - ▷ IDEA: convertir el vector de **observaciones Dirichlet en pseudo-observaciones Gaussianas condicionalmente independientes**

- Paquete de R
  - ▷ Programar algunas funciones del R-INLA fuera del R-INLA para poder utilizar los LGFs
  - ▷ Disponible: CoDa con efectos aleatorios i.i.d.

INLAcomp (con **Haavard Rue** en **KAUST**)

- Uso de log-ratios para lidiar con CoDa
  - ▷ IDEA: se propone una distribución, la **logistic-normal con covarianza Dirichlet**, que encaja perfectamente con R-INLA
- Paquete de R: incorporación uso de log-ratios para lidiar con CoDa
  - ▷ Es fácil incorporar cualquier LGF
  - ▷ Se pueden utilizar todas las herramientas para lidiar con varias verosimilitudes: copy, etc.

## LOGISTIC-NORMAL CON COVARIANZA DIRICHLET

## ■ Equivalencia entre Dirichlet y la Distribución Normal Logística

- ▶ El operador KL se minimizaba cuando:

$$\mu_d = \phi(\alpha_d) - \phi(\alpha_D), \quad d = 1, \dots, D-1,$$

$$\Sigma_{dd} = \phi'(\alpha_d) + \phi'(\alpha_D), \quad d = 1, \dots, D-1,$$

$$\Sigma_{dk} = \phi'(\alpha_D), \quad d \neq k.$$

■  $\mathbf{y} \in \mathbb{S}^D$  sigue una **distribución normal-logística con covarianza Dirichlet** si y solo si:

- ▶  $alr(\mathbf{y}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$   $\mathbf{y}$ ,
- ▶  $\Sigma_{dd} = \sigma_d^2 + \gamma, \quad d = 1, \dots, D-1,$
- ▶  $\Sigma_{dk} = \gamma, \quad d \neq k,$

Statistics and Computing (2024) 34:116  
<https://doi.org/10.1007/s11222-024-10427-3>

## A flexible Bayesian tool for CoDa mixed models: logistic-normal distribution with Dirichlet covariance

Joaquín Martínez-Minaya<sup>1</sup> · Haavard Rue<sup>2</sup>

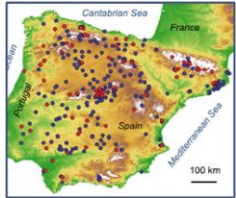
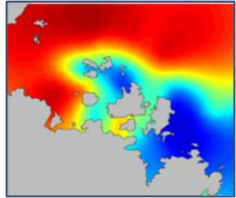
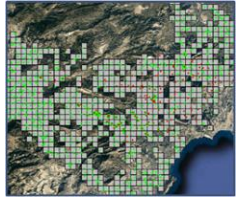
## INLAcomp (con Haavard Rue en KAUST)

## ■ Uso de log-ratios para lidiar con CoDa

- ▶ IDEA: se propone una distribución, la **logistic-normal con covarianza Dirichlet**, que encaja perfectamente con R-INLA.

## ■ Paquete de R: incorporación uso de log-ratios para lidiar con CoDa

- ▶ Es fácil incorporar cualquier LGF
- ▶ Se pueden utilizar todas las herramientas para lidiar con varias verosimilitudes: copy, etc.



## MODELO NORMAL-LOGÍSTICA CON COVARIANZA DIRICHLET (LNDM)

### Verosimilitud

$$alr(\mathbf{Y}) \sim \mathcal{ND}((\boldsymbol{\mu}^{(1)}, \dots, \boldsymbol{\mu}^{(d)}), \boldsymbol{\Sigma}),$$

$$\boldsymbol{\mu}^{(D-1)} = \mathbf{X}\boldsymbol{\beta}^{(d)} + \boldsymbol{\omega}^{(d)}, d = 1, \dots, D-1,$$

### LGF

$$\{\boldsymbol{\beta}^{(d)}, \boldsymbol{\omega}^{(d)} : d = 1, \dots, (D-1)\}$$

### Hiperparámetros

$$\boldsymbol{\theta}_1 = \{\sigma_d^2, \gamma : d = 1, \dots, (D-1)\}$$

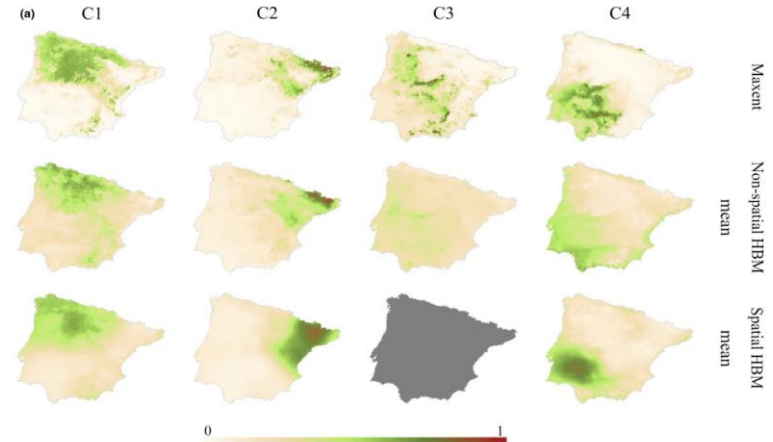
$$\boldsymbol{\theta}_2 = \{\sigma_{\omega}, \phi\}$$

## RESOURCE ARTICLE

## MOLECULAR ECOLOGY RESOURCES

### A hierarchical Bayesian Beta regression approach to study the effects of geographical genetic structure and spatial autocorrelation on species distribution range shifts

Joaquín Martínez-Minaya<sup>1</sup> | David Conesa<sup>1</sup> | Marie-Josée Fortin<sup>2</sup> |  
Carlos Alonso-Blanco<sup>3</sup> | F. Xavier Pico<sup>4</sup> | Arnald Marcer<sup>5,6</sup>



- Motivación. La mancha foliar del caqui
- Modelos Latentes Gaussianos. INLA
- LGMs para la mancha negra de los cítricos
- L1: Estadística espacio-temporal en la modelización de la distribución de especies
- **L2: Aprendizaje bayesiano en salud, medioambiente y economía**
- L3: Métodos computacionales bayesianos
- Trabajos futuros



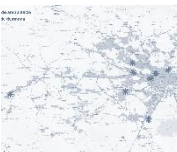
## How good is this medical device? Bayesian mixed models for agreement measures

IPs: Joaquín Martínez-Minaya y Vanda Inacio (UE)  
Período: 01 Sep 2023–31 Aug 2026.  
Financiado por The Royal Society.



## Novel methodologies for multi-omics integrative analysis with cutting-edge sequencing technologies

IPs: Sonia Tarazona (UPV) y Ana Conesa (CSIC)  
Período: 01 Sep 2024–31 Aug 2027.  
Financiado por Gobierno de España.



### Verosimilitud

$$y_{ijlt} = \mu + \alpha_i + \beta_j + \gamma_l + \epsilon_{ijlt}$$

### LGF

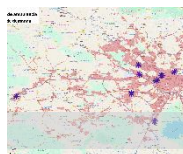
$$\begin{aligned} \beta_j &\sim \mathcal{N}(0, \sigma_\beta^2), \sigma_\beta^2 \text{fixed} . \\ \alpha_i &\sim \mathcal{N}(0, \sigma_\alpha^2), \gamma_l \sim \mathcal{N}(0, \sigma_\gamma^2), \\ \epsilon_{ijlt} &\sim \mathcal{N}(0, \sigma_\epsilon^2) \end{aligned}$$

### Hiperparámetros

$$\sigma_\alpha, \sigma_\gamma, \sigma_\epsilon \sim \text{PC-prior}(\sigma_0, \lambda) ,$$

*Retos metodológicos:*  
*Alta dimensión de los datos ómicos. Alto coste computacional*

Fenotipo	Donante	Célula	Tipo celular	Gen 1	Gen 2	Gen 3	...
Grupo A	Muestra 1	Célula 1	Tipo celular X				
		Célula 2	Tipo celular Y				
		...	...				
	Muestra 2	Célula 1	Tipo celular X				
		Célula 2	Tipo celular Y				
		...	...				
Grupo B	Muestra 1	Célula 1	Tipo celular X				
		Célula 2	Tipo celular Y				
		...	...				
	Muestra 2	Célula 1	Tipo celular X				
		Célula 2	Tipo celular Y				
		...	...				



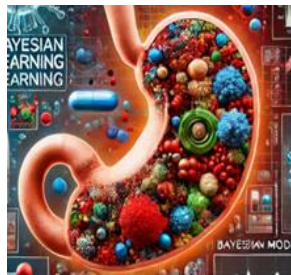
## Aprendizaje Bayesiano para Datos Composicionales de Alta Dimensión: Análisis del Impacto Socioeconómico en la Microbiota Vaginal (COMSOVA)

IP: Joaquín Martínez-Minaya.

Investigadores: Rubén Amorós (UV) y Blanca Sarzo (Fisabio).

Período: 01 Sep 2025–31 Aug 2026.

Financiado por Generalitat Valenciana.



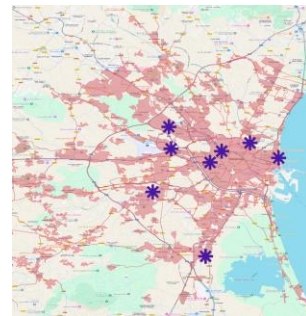
## Hireves: Herramienta Interactiva de Relocalización de Vehículos de Emergencias Sanitarias

IP: Fulgencia Villa Juliá.

Investigadores: Yulia Karpova, Esther Ródenas, Víctor Sánchez, Eva Vallada (UPV)

Período: 01 Ene 2024–31 Dic 2026.

Financiado por la Agencia Valenciana de Innovación.



### Verosimilitud

$$\underbrace{\mathbf{y}_i \mid \mathbf{p}_i}_{\text{Nivel 1: conteos por mujer}} \sim \text{Multinomial}(S_i, \mathbf{p}_i), \quad i = 1, \dots, N,$$

$$\underbrace{\mathbf{p}_i \mid z_i, \boldsymbol{\alpha}_{z_i}}_{\text{Nivel 2: composición condicionada al clúster}} \sim \text{Dirichlet}(\boldsymbol{\alpha}_{z_i}),$$

$$\underbrace{z_i \mid \boldsymbol{\pi}_i}_{\text{Nivel 3: asignación de clúster}} \sim \text{Categorical}(\boldsymbol{\pi}_i),$$

$$\log \left( \frac{\pi_i^{(k)}}{\pi_i^{(1)}} \right) = \mathbf{X}_i \boldsymbol{\beta}_k, \quad k = 2, \dots, K,$$

### Verosimilitud

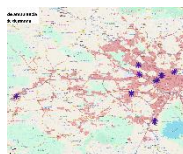
$$N(W) \mid \lambda(\cdot) \sim \text{Poisson} \left( \int_W \lambda(s) ds \right),$$

$$\lambda(s) = \exp\{\eta(s)\}, \quad s \in W,$$

$$\eta(s) = \boldsymbol{\beta}_0 + \mathbf{X}(s)^\top \boldsymbol{\beta} + \omega(s),$$

*Retos metodológicos:*

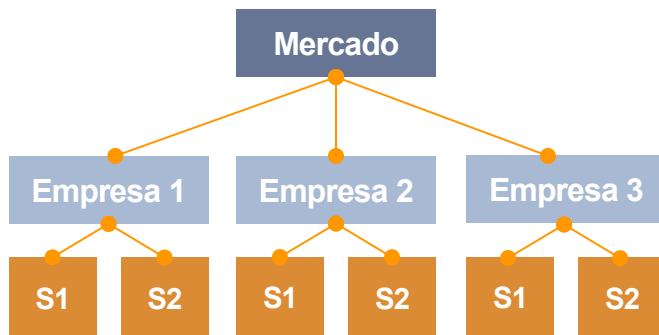
*Alta dimensión composicional, clustering, alto coste computacional*



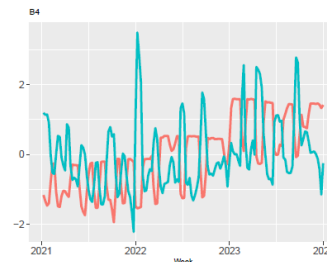
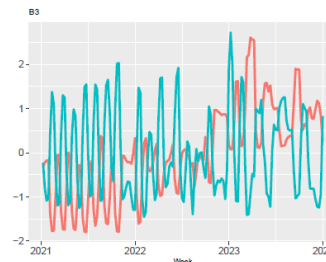
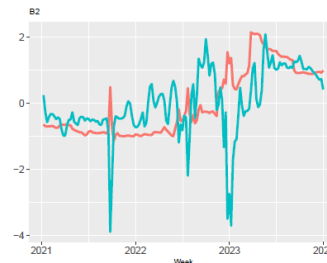
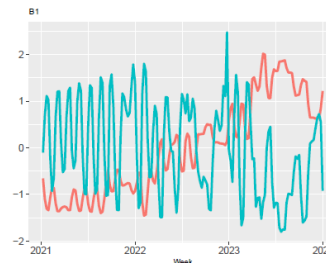
¿QUÉ PASA SI B1 INCREMENTA SU PRECIO EN UN 5%?

Brand	Current MS(%)	New MS(%)	Change in MS (PP)
B1	42.51	38.78	-3.74
B2	26.10	27.77	1.67
B3	26.86	28.81	1.95
B4	4.53	4.64	0.11

Tanto B2 como B3 aumentan sus cuotas de mercado



Retos metodológicos: CoDa temporal y jerárquico



Este trabajo forma parte de una colaboración con el Dr. **Balaji Raman**, de una empresa India.

TFG de Àngel Lova desarrollando CoDa temporales

- Motivación. La mancha foliar del caqui
- Modelos Latentes Gaussianos. INLA
- LGMs para la mancha negra de los cítricos
- L1: Estadística espacio-temporal en la modelización de la distribución de especies
- L2: Aprendizaje bayesiano en salud, medioambiente y economía
- **L3: Métodos computacionales bayesianos**
- Trabajos futuros

### CoDA TEMPORAL Y ESPACIO-TEMPORAL

- ▷ Ninguno de los paquetes ha sido testeado en el contexto temporal y espacio-temporal

#### OBJETIVO

- ▷ Incluir modelización **temporal y espacio temporal en CoDa**
- ▷ Jorge Mestre, estudiante de doctorado, lo está desarrollando y aplicando en un contexto pesquero

### CoDA JERÁRQUICO

- ▷ En el caso de las cuotas de mercado, existe un problema de **composiciones dentro de composiciones**

#### OBJETIVO

- ▷ Construir un método en que se puedan modelizar los diferentes niveles a nivel composicional

### REDUCCIÓN DIMENSIONALIDAD EN CoDA

- ▷ No es trivial realizar un análisis cluster en CoDa

#### OBJETIVO

- ▷ Usar árboles generadores mínimos + INLA para detectar clústeres en CoDa
- ▷ Jorge Mestre para datos funcionales con Paula Moraga (KAUST)
- ▷ Aplicación en Microbiota

### INLAcomp

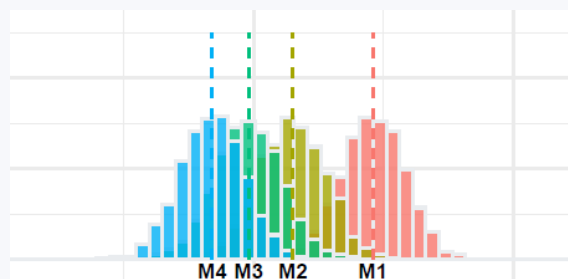
- ▷ INLAcomp, dirinla están basado en R-INLA
- ▷ CoDabrms está basado en brms
- ▷ Regresión Dirichlet, alr, clr e ilr

#### OBJETIVO

- ▷ Unificar usando **inlabru**.
- ▷ Incorporación de funciones no lineales

### $R^2$ BAYESIANO EN CoDA




- ▶ Obtención de una distribución a posteriori para el  $R^2$
- ▶ Elaborar un  $R^2$  para CoDa basado en la varianza total de una composición



- ▶ Extender la idea del **R2D2 al caso multivariante**. Colaboración con Javier Aguilar (TU Dortmund, desarrollador de brms) y Gabriel Calvo (UV)
- ▶ Medidas de validación de **IA bayesianas** (Adrián Colomer, UPV)
- ▶ Desarrollo de **medidas predictivas**: LGOCV. Aritz Adín (UPN), Haavard Rue (KAUST), Amanda Lenzi (UE)

Enviado a *SERRA*

SOIL TEXTURE PREDICTION WITH BAYESIAN GENERALIZED ADDITIVE MODELS FOR SPATIAL COMPOSITIONAL DATA

BY JOAQUÍN MARTÍNEZ-MINAYA <sup>\*1,a</sup>,  
LORE ZUMETA-OLASKOAGA <sup>2,3,b</sup> AND DAE-JIN LEE <sup>4,c</sup>

Spatial Statistics

Automatic cross-validation in structured models: Is it time to leave out leave-one-out?

Aritz Adin <sup>\*1</sup>, Elias Teixeira Krainski<sup>2</sup>, Amanda Lenzi<sup>3</sup>, Zhedong Liu<sup>4</sup>,  
Joaquín Martínez-Minaya<sup>5</sup>, Håvard Rue<sup>2</sup>

### META-LEARNING BAYESIANO

- ▶ Utilizar Deep learning para aprender modelos estadísticos
- ▶ Aplicación en casos donde el coste computacional es alto: datos ómicos, emergencias, microbiota vaginal

- Motivación. La mancha foliar del caqui
- Modelos Latentes Gaussianos. INLA
- LGMs para la mancha negra de los cítricos
- L1: Estadística espacio-temporal en la modelización de la distribución de especies
- L2: Aprendizaje bayesiano en salud, medioambiente y economía
- L3: Métodos computacionales bayesianos
- **Trabajos futuros**

# PRÓXIMOS PROYECTOS

## CLIMATE-SMART PUBLIC HEALTH

L2

FRANCESCA DOMINICI (U. HARVARD)

CHRISTOPHER GOLDEN (U. HARVARD)

## CONTAMINACIÓN EN SALUD EN LA COMUNIDAD VALENCIANA

L2

ANTONIO MARTÍNEZ MILLANA (UPV)

## TERAPIA HORMONAL Y SALUD CARDIOVASCULAR FEMENINA

L2

MARÍA EUGENIA MÉNDEZ ORTUONDO  
(ESTUDIANTE DEL GCDD, UPV)

## MODELOS DE SUPERVIVENCIA EN MELANOMA

L2

VERA TOMAZELLA, KATY CRUZ (U. SAO  
CARLOS), DANILO ALVARES (U. CAMBRIDGE)

## SUPERVIVENCIA PARA MENARQUIA Y MENOPAUSIA

L2

BLANCA SARZO, OIHANE ÁLVAREZ (FISABIO).  
M. JOSÉ LÓPEZ (UV), ALEX SCHMIDT (MCGILL)

## RIESGO/INCIDENCIA DE INCENDIOS FORESTALES

L2

ÁLVARO PÉREZ HERNANDO (ESTUDIANTE DEL  
DOBLE GRADO ADE-CTA, UPV)

## TUTORIALES

L1

BRMS: ELENA LÁZARO (UV)

INLA EN SDMS: XABIER BARBER (UMH)

## TEORÍA BAYESIANA PARA LOS NÚMEROS COMPLEJOS

L3

RENZO CABALLERO (U. TOKIO)

## DEEP LEARNING BAYESIANO EN IMÁGENES MÉDICAS

L2

ADRIÁN COLOMER (UPV)



¡Escanéeme!

# INTRODUCCIÓN AL APRENDIZAJE BAYESIANO USANDO INLA

Joaquín Martínez Minaya

[jmarmin@eio.upv.es](mailto:jmarmin@eio.upv.es)

<https://jmartinez-minaya.github.io/>



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

10 de noviembre de 2025