Who are we and What do we do?
Who are we?

- VaBaR is a *baby* team in the sense that it has just some moths old, but we have been collaborating together for some years.
- Different institutions:
  - University of Valencia (Spain): Anabel Forte, Antonio López, Blanca Sarzo, Carmen Armero, Danilo Alvares, David Conesa, Elena Lázaro and Joaquín Martínez-Minaya.
  - University of Castilla-La Mancha (Spain): Gonzalo García and Virgilio Gómez.
  - Institute National de la Recherche Agronomique (France): Facundo Muñoz.
  - University Rey Juan Carlos (Spain): Maria Eugenia Castellanos.
  - Foundation for the promotion of health and biomedical research of Valencia (Spain): Jordi Pérez and Rubén Amorós.
  - University Miguel Hernández de Elche (Spain): Xavier Barber.
  - University Carlos III de Madrid (Spain): Stefano Cabras.
Where are we from?
What do we do? Our lines of research:
Dedication of the group: **Bayesian inference**.

1 Theoretical:
   - *Solving Variable Selection in Linear Models*: selecting explanatory variables from a set of \( p \) potential ones.

2 Computational:
   - *Uncertainty quantification in Computer Models*: computer model aims at reproducing some real phenomenon.
   - *Approximate Bayesian Computation and Markov chain Monte Carlo algorithms*.

3 Biological and Medical studies:
   - *Detection of Influenza Outbreaks*.
   - *Species distribution modeling*.
   - *Joint Modelling of Longitudinal and Survival data*.
   - *Dynamic predition*.
   - *Study of virus prevalence in plants*.
   - *Sea-birds movements*.
S-T Detection of Influenza Outbreaks.

- Data in several regions
- Probability of being in epidemic → Detection of the onset
S-T Detection of Influenza Outbreak.

\[ Y_{ti} - Y_{t-1i} \sim N(R_{ti}Z_{ti}, \sigma^2_{Z_{ti}}) \]

\[ R_{ti0} = \mu_{t0} \]

\[ R_{ti1} = \mu_{t1} + AR + CAR \]

\[ \sigma_0 < \sigma_1 \]

\[ Z_{ti} \sim Cat(P_{Z_{t-1i}}) \]

- Differentiated rates \( \sim \) Norm. dist.
- Mean for non-epidemic state \( \rightarrow \) White noise
- Mean for epidemic state \( \rightarrow \) time and space autoregressive structure
- More variance in epidemic state
- Epidemic/Non-epidemic state \( \sim \) HMM(t-1,i)
Species distribution modeling

- **Species distribution modelling** links spatially referenced records of species occurrence with maps of environmental variables in order to create a statistical model of the relationship between a species and its environment.

- **Typical examples**: diseases, fish species, plants, animals, etc.
- **Typical covariates**: elevation, climate, vegetation, human disturbance, temperature, chlorophyll-a, etc.
- **Applications**: climate change, conservation of species, prevalence of diseases, etc.
- **Most of the available data come from the monitoring programmes and consists of** Presence/absence of the species (diseases) or Abundance.
Model for presence / absence of species:

1. Define a binary random variable as a response variable:
   
   \[ Z_i \sim Ber(\pi_i), \ 1 \leq i \leq n \]

2. \( \pi_i \) linked with linear predictor and spatial random effect:
   
   \[ \logit(\pi_i) = X_i \beta + W_i, \ 1 \leq i \leq n \]

\( W \) (Gaussian distributed) \( \rightarrow \) Latent Gaussian field.

3. A prior distribution for the hyperparameters of \( W \) is assigned. The resulting spatial model is a latent Gaussian model.

- We could compute approximations for the posterior distributions of the parameters with INLA but this is a continuously indexed Gaussian Field and INLA cannot be applied directly.
- GRMFs are discretely indexed \( \rightarrow \) the Markov property makes the precision matrix involved sparse (allowing the use of faster numerical algorithms).
In spite of its low commercial value, mackerel plays an important role in the observed transition zone between the Mediterranean and Atlantic sea.

Covariates in final model: \( \log(\text{Depth}) \) (neg. effect) and chlorophyll-a (pos. effect).

There seems to be a east-west effect.

This was associated a posteriori by local experts with the fact that the western area of the bay is a protected coastline with favourable conditions for the species.

The model provided (unexpectedly) a quantification of the impact of this protective action on the Mediterranean horse mackerel.
In spite of its low commercial value, mackerel plays an important role in the observed transition zone between the Mediterranean and Atlantic sea.

Covariates in final model: \( \log(\text{Depth}) \) (neg. effect) and chlorophyll-a (pos. effect).

There seems to be an east-west effect.

This was associated a posteriori by local experts with the fact that the western area of the bay is a protected coastline with favourable conditions for the species.

The model provided (unexpectedly) a quantification of the impact of this protective action on the Mediterranean horse mackerel.
In spite of its low commercial value, mackerel plays an important role in the observed transition zone between the Mediterranean and Atlantic sea.

- Covariates in final model: $\log(\text{Depth})$ (neg. effect) and chlorophyll-a (pos. effect).

There seems to be an east-west effect.

This was associated a posteriori by local experts with the fact that the western area of the bay is a protected coastline with favourable conditions for the species.

The model provided (unexpectedly) a quantification of the impact of this protective action on the Mediterranean horse mackerel.
In spite of its low commercial value, mackerel plays an important role in the observed transition zone between the Mediterranean and Atlantic sea. Covariates in final model: log(Depth) (neg. effect) and chlorophyll-a (pos. effect).

There seems to be a east-west effect.

This was associated a posteriori by local experts with the fact that the western area of the bay is a protected coastline with favourable conditions for the species.

The model provided (unexpectedly) a quantification of the impact of this protective action on the Mediterranean horse mackerel.
In spite of its low commercial value, mackerel plays an important role in the observed transition zone between the Mediterranean and Atlantic sea.

Covariates in final model: $\log(\text{Depth})$ (neg. effect) and chlorophyll-a (pos. effect).

There seems to be an east-west effect.

This was associated a posteriori by local experts with the fact that the western area of the bay is a protected coastline with favourable conditions for the species.

The model provided (unexpectedly) a quantification of the impact of this protective action on the Mediterranean horse mackerel.
Distribution of Mediterranean horse mackerel in Gulf of Almería (Muñoz et al., 2013)

- In spite of its low commercial value, mackerel plays an important role in the observed transition zone between the Mediterranean and Atlantic sea.
- Covariates in final model: \( \log(\text{Depth}) \) (neg. effect) and chlorophyll-a (pos. effect).

**Spatial effect**

![Spatial effect](image)

Posterior mean and standard deviation of the spatial effect

- There seems to be a east-west effect.
- This was associated a posteriori by local experts with the fact that the western area of the bay is a protected coastline with favourable conditions for the species.
- The model provided (unexpectedly) a quantification of the impact of this protective action on the Mediterranean horse mackerel.
Distribution of three elasmobranch species (Pennino et al., 2013).

- There is an increasing concern over elasmobranch species because they are highly vulnerable to fishing pressure.

- Main predictors of elasmobranch habitats are depth, slope of seabed and type of substrate, followed by temperature and chlorophyll-a.

- Species show different optimum depths: could indicate a sort of fine-tuned bathymetric segregation, though they coexist on shelf and slope bottoms.

- These maps can be used to identify sensitive habitats, with a final aim to improve management and conservation of these vulnerable species.
Distribution of three elasmobranch species (Pennino et al., 2013).

- There is an increasing concern over elasmobranch species because they are highly vulnerable to fishing pressure.
- Main predictors of elasmobranch habitats are depth, slope of seabed and type of substrate, followed by temperature and chlorophyll-a.

- Species show different optimum depths: could indicate a sort of fine-tuned bathymetric segregation, though they coexist on shelf and slope bottoms.
- These maps can be used to identify sensitive habitats, with a final aim to improve management and conservation of these vulnerable species.
There is an increasing concern over elasmobranch species because they are highly vulnerable to fishing pressure.

Main predictors of elasmobranch habitats are depth, slope of seabed and type of substrate, followed by temperature and chlorophyll-a.

Species show different optimum depths: could indicate a sort of fine-tuned bathymetric segregation, though they coexist on shelf and slope bottoms.

These maps can be used to identify sensitive habitats, with a final aim to improve management and conservation of these vulnerable species.
Distribution of three elasmobranch species (Pennino et al., 2013).

- There is an increasing concern over elasmobranch species because they are highly vulnerable to fishing pressure.
- Main predictors of elasmobranch habitats are depth, slope of seabed and type of substrate, followed by temperature and chlorophyll-a.

Median of the posterior probability of the presence of elasmobranch species

- Species show different optimum depths: could indicate a sort of fine-tuned bathymetric segregation, though they coexist on shelf and slope bottoms.
- These maps can be used to identify sensitive habitats, with a final aim to improve management and conservation of these vulnerable species.
Distribution of three elasmobranch species (Pennino et al., 2013).

- There is an increasing concern over elasmobranch species because they are highly vulnerable to fishing pressure.
- Main predictors of elasmobranch habitats are depth, slope of seabed and type of substrate, followed by temperature and chlorophyll-a.

Species show different optimum depths: could indicate a sort of fine-tuned bathymetric segregation, though they coexist on shelf and slope bottoms.

These maps can be used to identify sensitive habitats, with a final aim to improve management and conservation of these vulnerable species.
Big concern in Galicia about paramphistomosis in cows and the possibility of infecting humans.

Interest of detecting the zones of higher prevalence.

Gain knowledge about the different covariates: most relevant are Temperature and log(slope).

Spatial component of the fitted model in dairy cows throughout Galicia: Posterior mean (A) and standard deviation (B).
Distribution of prevalence of Bovine paramphistomosis in Galicia (González-Warleta et al., 2013).

- Big concern in Galicia about paramphistomosis in cows and the possibility of infecting humans.

- Interest of detecting the zones of higher prevalence.

- Gain knowledge about the different covariates: most relevant are Temperature and log(slope).

- Spatial component of the fitted model in dairy cows throughout Galicia: Posterior mean (A) and standard deviation (B).
Distribution of prevalence of Bovine paramphistomosis in Galicia (González-Warleta et al., 2013).

- Big concern in Galicia about paramphistomosis in cows and the possibility of infecting humans.

- Interest of detecting the zones of higher prevalence.

- Gain knowledge about the different covariates: most relevant are Temperature and log(slope).

- Spatial component of the fitted model in dairy cows throughout Galicia: Posterior mean (A) and standard deviation (B).
Distribution of prevalence of Bovine paramphistomosis in Galicia (González-Warleta et al., 2013).

- Big concern in Galicia about paramphistomosis in cows and the possibility of infecting humans.
- Interest of detecting the zones of higher prevalence.
- Gain knowledge about the different covariates: most relevant are Temperature and log(slope).
- Spatial component of the fitted model in dairy cows throughout Galicia: Posterior mean (A) and standard deviation (B).
Preferential sampling. The previous model usually assume that sampling locations and the process being modelled are stochastically independent. But, sampling locations are deliberately concentrated in areas where the abundance of species is known or expected to be high: 

\[ \text{fishermen fish in areas where they are likely to find fish.} \]

In the first situation hot spots of high-CPUE density can be appreciated. In the other case the spatial component surface is smoother: no particular hot spot is marked.
Preferential sampling. The previous model usually assume that sampling locations and the process being modelled are stochastically independent. But, sampling locations are deliberately concentrated in areas where the abundance of species is known or expected to be high:

→ fishermen fish in areas where they are likely to find fish.

Posterior mean of the spatial effect of the preferential sampling model with fisherman error (left) and (right) the non-preferential sampling model

- In the first situation hot spots of high-CPUE density can be appreciated.
- In the other case the spatial component surface is smoother: no particular hot spot is marked.
Preferential sampling. The previous model usually assume that sampling locations and the process being modelled are stochastically independent. But, sampling locations are deliberately concentrated in areas where the abundance of species is known or expected to be high: fishermen fish in areas where they are likely to find fish.

Posterior mean of the spatial effect of the preferential sampling model with fisherman error (left) and (right) the non-preferential sampling model

In the first situation hot spots of high-CPUE density can be appreciated. In the other case the spatial component surface is smoother: no particular hot spot is marked.
Preferential sampling. The previous model usually assume that sampling locations and the process being modelled are stochastically independent. But, sampling locations are deliberately concentrated in areas where the abundance of species is known or expected to be high: fishermen fish in areas where they are likely to find fish.

In the first situation hot spots of high-CPUE density can be appreciated. In the other case the spatial component surface is smoother: no particular hot spot is marked.
- **Misalignment.** It appears when **measurement values of the covariates are not known** at the observed locations nor at those locations where we are going to make predictions.
- **Misalignment.** It appears when *measurement values of the covariates are not known* at the observed locations nor at those locations where we are going to make predictions.

**Example of misalignment:** the 67 official weather stations in Galicia do not coincide with the farms where data were observed.
Fasciolosis in Galicia solved not taking into account misalignment

Posterior mean of the probability of occurrence (left) and the first (center) and third (right) quantiles. Red points mean Presence and black points mean Absence.
Fasciolosis in Galicia solved taking into account misalignment

- Modelling under uncertainty in the covariates can still be performed using the SPDE approach.
- Barber et al. (2016) have used this approach to analyze the presence of fasciolosis in Galicia.
Fasciolosis in Galicia solved taking into account misalignment

- Modelling under uncertainty in the covariates can still be performed using the SPDE approach.
- Barber et al. (2016) have used this approach to analyze the presence of fasciolosis in Galicia.
Fasciolosis in Galicia solved taking into account misalignment

- Modelling under uncertainty in the covariates can still be performed using the SPDE approach
- Barber et al. (2016) have used this approach to analyze the presence *fasciolosis* in Galicia.

**Posterior mean of the probability of occurrence (left) and the first (center) and third (right) quantiles**
Spatio-temporal analysis

- It is also possible to incorporate temporal components (autoregressive, random walks, etc.) to the previous model.
- Example: the persistence over time of abundance hot-spots is key in order to identify nursery areas in fisheries.
- Paradinas et al. (2015) have analyzed persistence by comparing different spatio-temporal models.

Other extensions

- Shared components, splines, and many others.
Spatio-temporal analysis

- It is also possible to incorporate temporal components (autoregressive, random walks, etc.) to the previous model.
- Example: the persistence over time of abundance hot-spots is key in order to identify nursery areas in fisheries.
- Paradinas et al. (2015) have analyzed persistance by comparing different spatio-temporal models.

Other extensions

- Shared components, splines, and many others.
**Spatio-temporal analysis**

- It is also possible to incorporate temporal components (autoregressive, random walks, etc.) to the previous model.
- Example: the persistence over time of abundance hot-spots is key in order to identify nursery areas in fisheries.
- Paradinas et al. (2015) have analyzed persistence by comparing different spatio-temporal models.

**Other extensions**

- Shared components, splines, and many others.
Spatio-temporal analysis

- It is also possible to incorporate temporal components (autorregressive, random walks, etc.) to the previous model.
- Example: the persistence over time of abundance hot-spots is key in order to identify nursery areas in fisheries.
- Paradinas et al. (2015) have analyzed persistence by comparing different spatio-temporal models.

Other extensions

- Shared components, splines, and many others.
**Spatio-temporal analysis**

- It is also possible to incorporate temporal components (autoregressive, random walks, etc.) to the previous model.
- Example: the persistence over time of abundance hot-spots is key in order to identify nursery areas in fisheries.
- Paradinas et al. (2015) have analyzed persistence by comparing different spatio-temporal models.

**Other extensions**

- Shared components, splines, and many others.