Model-based geostatistics for wildlife population monitoring : Northwestern Mediterranean fin whale population and other case studies

Pascal Monestiez

Biostatistique et Processus Spatiaux, INRA, Avignon, France

joint work with

Léa David, Nathalie Di-Méglio, Edwige Bellier, Grégoire Certain, Christophe Guinet, Vincent Bretagnole, ...

Chaire MMB, 25 mars 2010, MNHM, Paris

Context and aims

- Using data from distance sampling surveys to map animal spatial distributions (line and strip transects)
- Data pooled from multiple sources. Same visual line transect protocol, Only good quality records kept
- Low densities, rare sightings
- Capture-recapture methods based on photo-identification not efficient (recapture probability too low)

Context and aims

- Sighting data are summed on small spatial cells to get count data associated with effort
- Geostatistical methods are applied considering : count data, zero inflated distribution, known non-stationarity
- To propose improved form of Kriging giving maps of animal density and associated maps of standard error of prediction

Spatial hierarchical non-stationary model

For all site s (a small spatial cell), Z_s is the number of sightings

$$\begin{cases} Z_s | Y_s & \sim \mathcal{P}\left(Y_s\right) \\ Y_s & = m_s X_s \end{cases}$$

$\mathcal{P}$ independent (given Y) Poisson distributions
m_s a deterministic drift (habitat characteristics, historical data)
X_s a positive stationary random field with unit mean, variance σ_X^2, covariance function $C_{ss'} = \mathrm{Cov}[X_s, X_{s'}]$, and/or variogram $\gamma_{ss'}$

Drift definition and estimation

The drift m_s resumes spatial non-stationarities

- Explained non-stationarities (habitat characteristics)
  - environmental variables as proxies
  - knowledge on spatial potential habitat
  - cokriging of long range spatial components
- given non-stationarities (based on past data or a priori)
  - surveys from previous years
  - pooled data from independent sources
  - kernel smoother or filter kriging
- or else, an assumption of stationarity (constant mean)

Drift definition and estimation

The drift m_s resumes spatial non-stationarities

- Explained non-stationarities (habitat characteristics)
  - environmental variables as proxies
  - knowledge on spatial potential habitat
  - cokriging of long range spatial components
- given non-stationarities (based on past data or a priori)
  - surveys from previous years
  - pooled data from independent sources
  - kernel smoother or filter kriging
- or else, an assumption of stationarity (constant mean)

Variogram estimation

Let Z_α, $\alpha = 1, \ldots, n$ be the n measurements of $Z(s_\alpha)$.
An experimental variogram of the latent variable X is :

$$\gamma^*_X(h) = \frac{1}{2\,N(h)} \sum_{\substack{\alpha,\beta \\ d_{\alpha\beta} \sim h}} \left(\frac{m_\alpha\, m_\beta}{m_\alpha + m_\beta} \left(\frac{Z_\alpha}{m_\alpha} - \frac{Z_\beta}{m_\beta} \right)^2 - 1 \right)$$

Kriging in non-stationary hierarchical context

The Kriging of Y in site s_o is given by :

$$Y_o^* = \sum_{\alpha=1}^{n} \lambda_\alpha \frac{m_o Z_\alpha}{m_\alpha} \qquad \text{where } \lambda_\alpha \text{ are solutions of}$$

$$\begin{cases} \displaystyle\sum_{\beta=1}^{n} \lambda_\beta C_{\alpha\beta} + \frac{\lambda_\alpha}{m_\alpha} + \mu = C_{\alpha o} \quad \text{for} \quad \alpha = 1, \ldots, n \\ \displaystyle\sum_{\alpha=1}^{n} \lambda_\alpha = 1 \end{cases}$$

The error variance of prediction is :

$$\mathrm{Var}(Y_o^* - Y_o) = m_o^2 \left(\sigma_X^2 - \sum_{\alpha=1}^{n} \lambda_\alpha C_{\alpha o} - \mu \right)$$

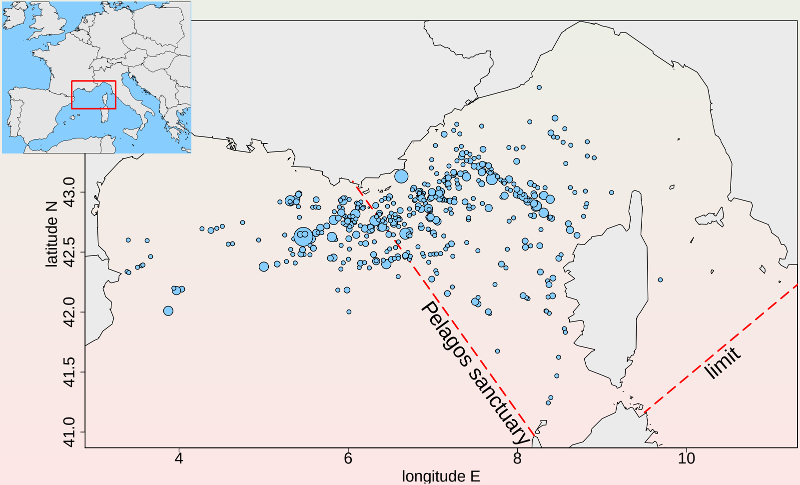
Mapping spatial distributions of Fin whale in Mediterranean Sea

- Presence of a resident population in the western Mediterranean Sea (estimations range from 700 to 3500 individuals, *Forcada 1996, Gannier 2006*)
- Classified as endangered by the IUCN Red List.

- population concentrates in summer in the Ligurian basin (100.000 km^2)
- An International Sanctuary (France, Italy, Monaco) established in 1999 and effective since 2001 (PELAGOS)

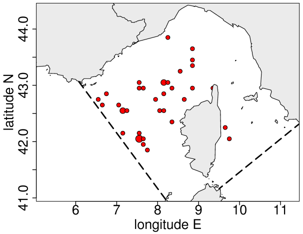
Fin whale summer spatial distribution

Fin whale sightings : raw data 1994 to 2008

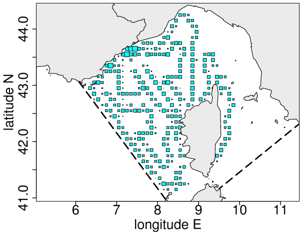


Poisson kriging and density map: Fin whales year 2001

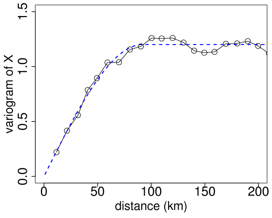
37 fin whale sightings

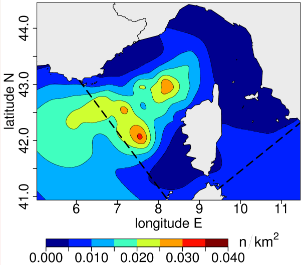


effort 220 hours or 2440 km

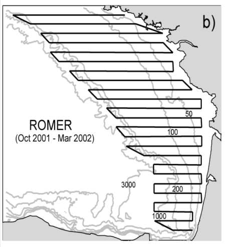


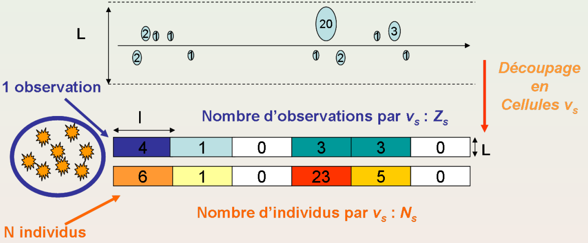
variogram of density X

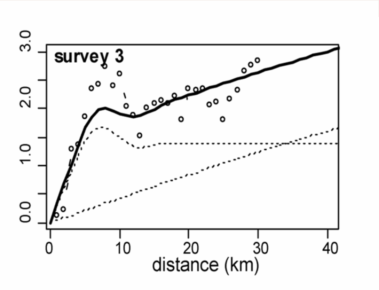


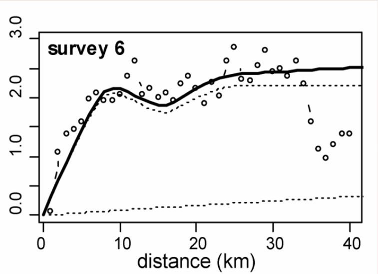


Variograms : Awks in the Bay of Biscay

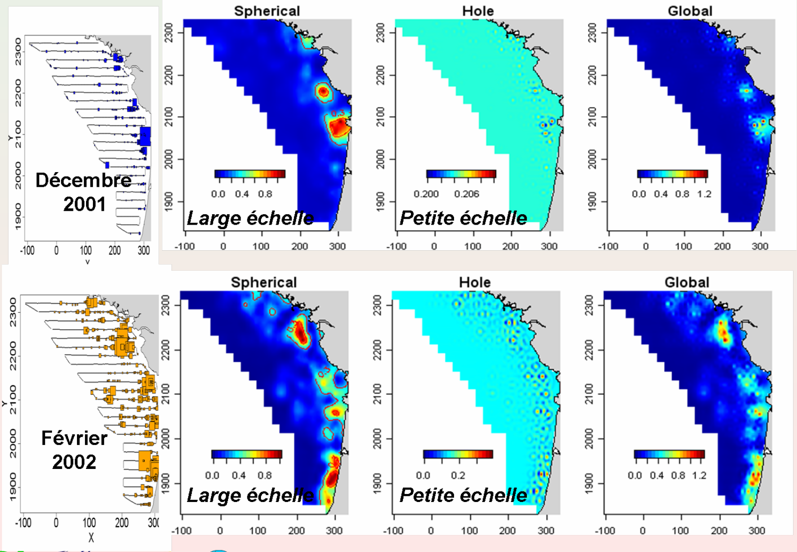




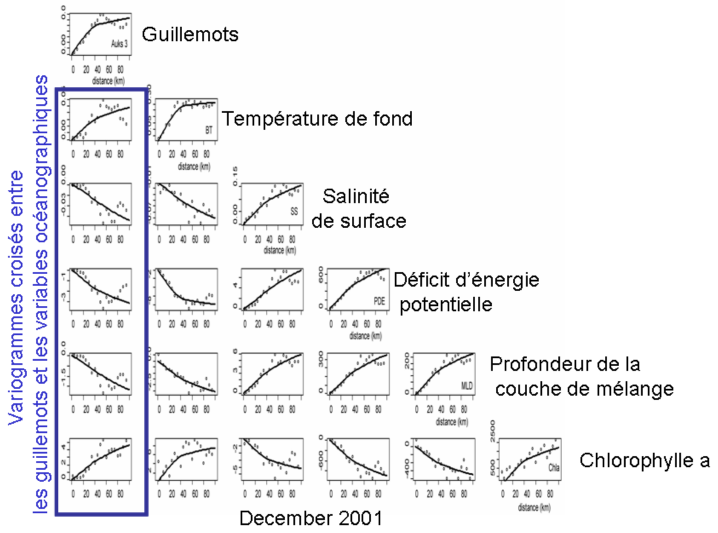




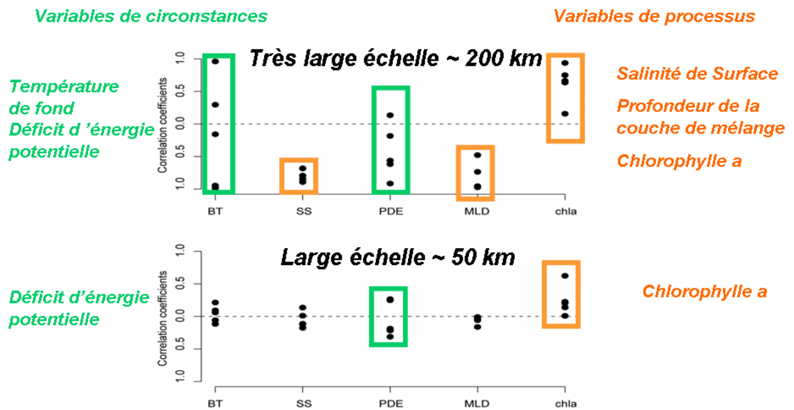
Filter kriging : Awks in the Bay of Biscay



Multivariate model : Linear Model of Coregionalisation with oceanic variables



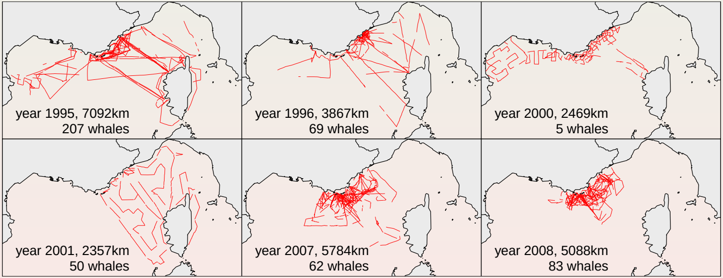
Multivariate model : time variation and/or stability



Distance Sampling : line transect approach

Data collection

- Line transect sampling with quantified efforts and standard protocol to record whale sightings and school size

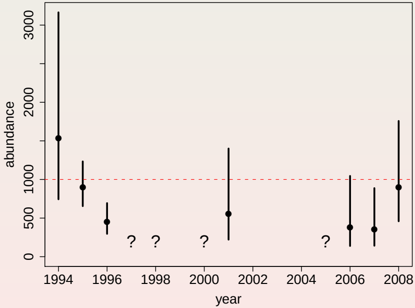


- Changes in sample designs from year to year : logistic, funding, targeted area (examples for 6 different years)

Distance Sampling : line transect approach

Data processing

- fitting the detection function $\Rightarrow$ effective strip width
- variance and confidence intervals by block bootstrap

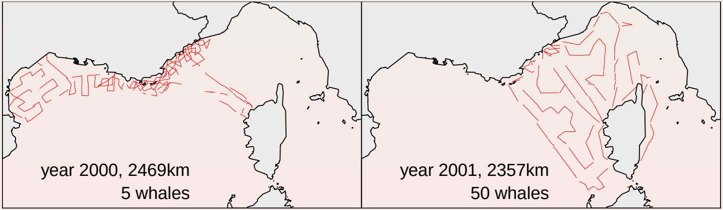


- U shape curve ? Effect of the Pelagos Sanctuary after 2001 ?

Distance Sampling : discussion

- Violation of one major assumption of Distance Sampling

"Independence between sampling scheme and whale spatial distribution"



From whale density to abundance : Averaged spatial mean by Block Poisson Kriging

block kriging system : (n+1) equations

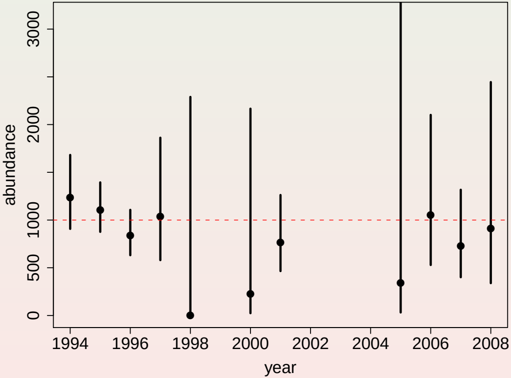
$$\begin{cases} \displaystyle\sum_{\beta=1}^{n} \lambda_\beta C_{\alpha\beta} + \frac{\lambda_\alpha}{m_\alpha} + \mu = \frac{1}{V}\int_V m_s C_{\alpha s}\, ds \quad \text{for} \quad \alpha = 1,\dots,n \\ \displaystyle\sum_{\alpha=1}^{n} \lambda_\alpha = \frac{1}{V}\int_V m_s\, ds = m_V \end{cases}$$

block kriging variance

$$\mathrm{Var}(Y_V^* - Y_V) = \frac{1}{V^2}\iint_{V\times V} m_s m_{s'} C_{ss'}\, ds\, ds' - \sum_{\alpha=1}^{n} \frac{\lambda_\alpha}{V}\int_V m_s C_{\alpha s}\, ds - m_V\, \mu$$

Abundance estimations by Block Poisson Kriging

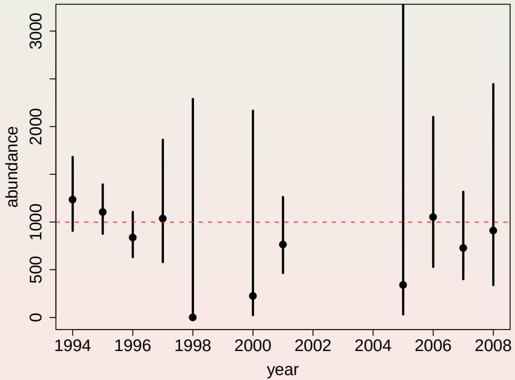
Total abundances versus years
(spatial averaged mean $Y_V^* \times$ domain V area)



No significant time trend : population size remains constant ?

Abundance estimations by Block Poisson Kriging

Total abundances versus years
(spatial averaged mean $Y_V^* \times$ domain V area)



No significant time trend : population size remains constant ?

Conclusions on methods and models

- Spatial modelling and Poisson Kriging do not totally replace Distance Sampling (some parameters of DS as the detection function, or the school sizes always needed)
- Rigorous protocols remain necessary to get total abundance
- Spatial modelling generally does not reduce confidence intervals, the main purpose is to correct bias from inhomogeneous sampling schemes.

Conclusions on monitoring

- In long term monitoring, survey designs and/or sampling schemes never remain constant (empirical law)
- This heterogeneity is accentuated when getting data from multiple sources
- Need to prevent bias or error underestimation due to sample variation, need to prevent bias due to population spatial shifts (changing environment)
- Then, it is crucial to model the animal spatial distribution in density (trend and stochastic part), before concluding on long term variation : decline, stability or recovery ?

Conclusions on monitoring

- In long term monitoring, survey designs and/or sampling schemes never remain constant (empirical law)
- This heterogeneity is accentuated when getting data from multiple sources
- Need to prevent bias or error underestimation due to sample variation, need to prevent bias due to population spatial shifts (changing environment)
- Then, it is crucial to model the animal spatial distribution in density (trend and stochastic part), before concluding on long term variation : decline, stability or recovery ?

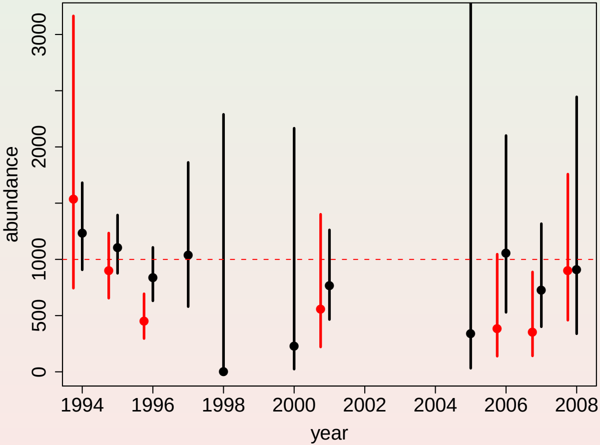
References

Monestiez P. Dubroca L. Bonnin E. Durbec J-P. Guinet C. (2006)
Geostatistical modelling of spatial distribution of *Balaenoptera physalus* in the Northwestern Mediterranean Sea from sparse count data
Ecological Modelling **193**, 615–628

Bellier E. (2007)
Spatial distribution within hierachical patchs dynamics systems : identifying species-environment spatial relationships at multiscales
PhD Thesis Université de la Méditérranée (Aix-Marseille II)

Bellier, E., Monestiez, P. (2009).
Model based block kriging to estimate spatial abundance of wildlife populations from count data.
submitted to *Environmental and Ecological Statistics*)

Bellier, E., Certain, G., Planque, B., Monestiez, P, Bretagnolle, V. (2010)
Modelling habitat selection at multiple scales with multivariate Geostatistics: an application to seabirds at sea
Oikos (in press)


abundance
0
500
1000
2000
3000
1994
1996
1998
2000
2002
2004
2006
2008
year