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
1 Context and aims

2 Spatial hierarchical non-stationary model
   - Model definition
   - Inference: spatial drift, variogram
   - Prediction: non stationary Poisson Kriging

3 Mapping spatial distributions
   - Fin whale case study
   - Awks case study in the Bay of Biscay
   - Nested spatial scales
   - Spatial multivariate model

4 Total abundance estimations
   - Line transect method and Distance Sampling
   - From whale density to abundance by Block Poisson Kriging

5 Conclusions
   - On methods and models
   - On monitoring
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

\[
\begin{aligned}
Z_s | Y_s & \sim \mathcal{P}(Y_s) \\
Y_s & = m_s X_s
\end{aligned}
\]

\( \mathcal{P} \) 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 \( \sigma_X^2 \),

covariance function \( C_{ss'} = \text{Cov}[X_s, X_{s'}] \), and/or variogram \( \gamma_{ss'} \)
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)
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)
Let $Z_\alpha, \alpha = 1, \ldots, n$ be the $n$ measurements of $Z(s_\alpha)$. An experimental variogram of the latent variable $X$ is:

$$
\gamma^*_X(h) = \frac{1}{2 \, N(h)} \sum_{\alpha, \beta} \left( \frac{m_\alpha m_\beta}{m_\alpha + m_\beta} \left( \frac{Z_\alpha}{m_\alpha} - \frac{Z_\beta}{m_\beta} \right)^2 - 1 \right)
$$
The Kriging of $Y$ in site $s_o$ is given by:

$$
Y_o^* = \sum_{\alpha=1}^{n} \lambda_\alpha \frac{m_\alpha Z_\alpha}{m_\alpha}
$$

where $\lambda_\alpha$ are solutions of

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

The error variance of prediction is:

$$
\text{Var}(Y_o^* - Y_o) = m_o^2 \left( \sigma_X^2 - \sum_{\alpha=1}^{n} \lambda_\alpha C_{\alpha o} - \mu \right)
$$
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 km²).
- 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

variogram of density $X$

effort 220 hours or 2440 km
Variograms: Awks in the Bay of Biscay

ROMER (Oct 2001 - Mar 2002)

1 observation

N individus

Nombre d'observations par \( v_s \) : \( Z_s \)

Nombre d'individus par \( v_s \) : \( N_s \)

Découpage en Cellules \( v_s \)

survey 3

distance (km)

survey 6

distance (km)

P. Monestiez

Model-based geostatistics for population monitoring
Filter kriging: Awks in the Bay of Biscay

Décembre 2001

Février 2002

Large échelle

Petite échelle

Spherical

Hole

Global

INRA

P. Monestiez
Model-based geostatistics for population monitoring
Multivariate model: Linear Model of Coregionalisation with oceanic variables

Guillemots

Température de fond

Salinité de surface

Déficit d'énergie potentielle

Profondeur de la couche de mélangé

Chlorophylle a

December 2001
Multivariate model: time variation and/or stability

Variables de circonstances

Température de fond
Déficit d’énergie potentielle

Variables de processus

Salinité de Surface
Profondeur de la couche de mélange
Chlorophylle a

Très large échelle ~ 200 km

Large échelle ~ 50 km

P. Monestiez
Model-based geostatistics for population monitoring
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 Pelagros Sanctuary after 2001?
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

\[
\begin{align*}
\sum_{\beta=1}^{n} \lambda_{\beta} C_{\alpha\beta} + \frac{\lambda_{\alpha}}{m_{\alpha}} + \mu &= \frac{1}{V} \int_{V} m_{s} C_{\alpha s} \, ds \quad \text{for} \quad \alpha = 1, \ldots, n \\
\sum_{\alpha=1}^{n} \lambda_{\alpha} &= \frac{1}{V} \int_{V} m_{s} \, ds = m_{V}
\end{align*}
\]

Block kriging variance

\[
\text{Var}(Y_{V}^{*} - Y_{V}) = \frac{1}{V^{2}} \int \int_{V \times V} m_{s} m_{s'} C_{ss'} \, ds \, ds' - \sum_{\alpha=1}^{n} \frac{\lambda_{\alpha}}{V} \int_{V} m_{s} C_{\alpha s} \, ds \\
- m_{V} \mu
\]
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 $Y_V^* \times \text{domain } V \text{ area})$

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?

P. Monestiez

Model-based geostatistics for population monitoring
References


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*

Context Model Mapping Abundance Conclusion Ref.

P. Monestiez

Model-based geostatistics for population monitoring
Model-based geostatistics for population monitoring