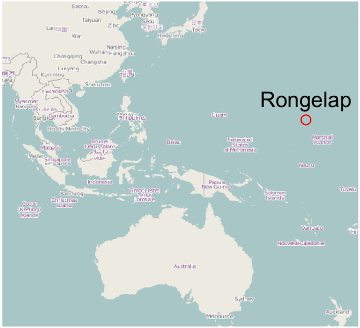
Modèles Hiérarchique Bayésien (2/2): Géostatistique

Aurélien Latouche

Conservatoire national des arts et métiers

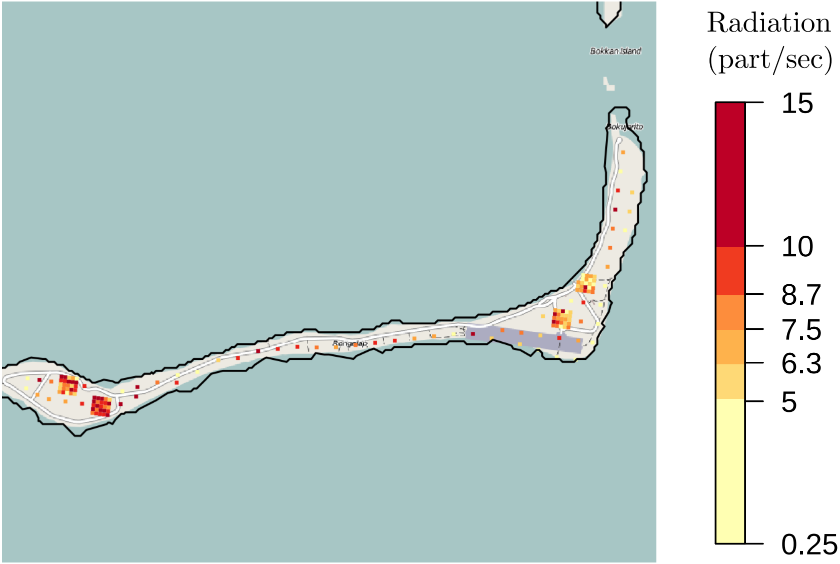
Etude de cas : Iles Marshall



- Essai nucléaire en 1954 : contamination
- Habitants évacuées après les essais (bcp de décès)
- En 1957, les autorités autorisent le retour des habitants

- Nombreux cas de Leucémies et Tumeurs de la thyroïde. Greenpeace évacue les habitants en 1985
- Des mesures de radiation sont faites et cartographiées
- Après nettoyage du sol (couche superficielle) les niveaux de radiation baissent
- Retour de 400 habitants en 1999

Iles Marshall : Atoll de Rongelap

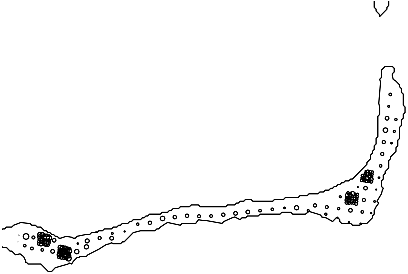


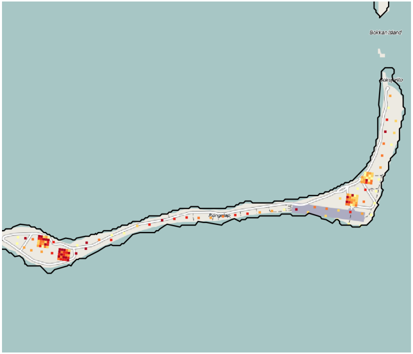
Description du problème

- Enquête de terrain des mesures de Césium 137 : ^{137}Cs
- Estimation de la variabilité spatiale de la radioactivité du césium ^{137}Cs
- Comparaison avec les seuils tolérés
- Utiliser les données de comptage Y_i sur un intervalle de temps t_i aux sites s_i $(i = 1, ..., n)$
- Prédire l'intensité sur l'île
- Déterminer les sites où l'intensité est au dessus des niveaux tolérés

$\Rightarrow$ Cartographier toutes ces estimations

Cartographie





Iles Marshalls : un modèle

- Y_i nombre particules en s_i, observées en T_i secondes.
- $\lambda(s)$ l'intensité de radiation qui depend de $U(s)$.

- $Y_i \sim$

Iles Marshalls : un modèle

- Y_i nombre particules en s_i, observées en T_i secondes.
- $\lambda(s)$ l'intensité de radiation qui depend de $U(s)$.

- $Y_i \sim \text{Poisson}(\lambda(s_i)\,T_i)$

Iles Marshalls : un modèle

- Y_i nombre particules en s_i, observées en T_i secondes.
- $\lambda(s)$ l'intensité de radiation qui depend de $U(s)$.

- $Y_i \sim \text{Poisson}(\lambda(s_i)\, T_i)$
- $\log[\lambda(s)] =$

Iles Marshalls : un modèle

- Y_i nombre particules en s_i, observées en T_i secondes.
- $\lambda(s)$ l'intensité de radiation qui depend de $U(s)$.

- $Y_i \sim \text{Poisson}(\lambda(s_i)\,T_i)$
- $\log[\lambda(s)] = \mu +$

Iles Marshalls : un modèle

- Y_i nombre particules en s_i, observées en T_i secondes.
- $\lambda(s)$ l'intensité de radiation qui depend de $U(s)$.

- $Y_i \sim \text{Poisson}(\lambda(s_i)\, T_i)$
- $\log[\lambda(s)] = \mu + U(s)$

Iles Marshalls : un modèle

- Y_i nombre particules en s_i, observées en T_i secondes.
- $\lambda(s)$ l'intensité de radiation qui depend de $U(s)$.

- $Y_i \sim \text{Poisson}(\lambda(s_i)\, T_i)$
- $\log[\lambda(s)] = \mu + U(s) + (\beta X(s))$

Iles Marshalls : un modèle

- Y_i nombre particules en s_i, observées en T_i secondes.
- $\lambda(s)$ l'intensité de radiation qui depend de $U(s)$.

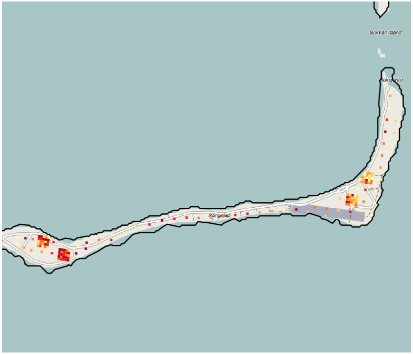
- $Y_i \sim \text{Poisson}(\lambda(s_i)\, T_i)$
- $\log[\lambda(s)] = \mu + U(s) + (\beta X(s))$
- $\text{cov}[U(s+h), U(s)] =$

Iles Marshalls : un modèle

- Y_i nombre particules en s_i, observées en T_i secondes.
- $\lambda(s)$ l'intensité de radiation qui depend de $U(s)$.

- $Y_i \sim \text{Poisson}(\lambda(s_i)\, T_i)$
- $\log[\lambda(s)] = \mu + U(s) + (\beta X(s))$
- $\text{cov}[U(s+h), U(s)] = \sigma^2 \exp(-h/\theta)$

- **Que valent μ, β, θ (range) , σ^2, ?**
- Et $\lambda(s)$?



Distribution a posteriori

$$\begin{aligned} \text{Jointe} &: [U, \sigma, \theta, \beta | Y] \\ \text{Marginales} &: [U|Y], [\sigma|Y], [\theta|Y], [\beta|Y] \end{aligned}$$

β correspond aux effets de covariables dans le modèle pour $\lambda()$

Problème

- Estimer β, σ, θ et $U(s)$?

Distribution aposteriori

$$pr(\beta,\sigma,\theta|Y) = \frac{pr(Y;\beta,\sigma,\theta)pr(\beta,\sigma,\theta)}{\int\int\int pr(Y;\beta,\sigma,\theta)d\beta d\sigma d\theta}$$

Integrales incalculable explicitement $\rightarrow$ MCMC

Iles Marshalls : Estimation du modèle

$$
\begin{aligned}
Y_i \sim & \mathrm{Poisson}(\lambda_i) \\
\log(\lambda_i) = & \underbrace{\log(T_i)}_{\texttt{Offset}} + \mu + U(s_i) \\
\mathrm{cov}[U(s+h), U(s)] = & \sigma^2 \exp(-h/\theta)
\end{aligned}
$$

1. Spécifier un modèle
2. Echantillon de l'aposteriori $[\mu, \sigma, \theta, U(s_i) | Y]$
3. Convergence des chaines
4. Pour une grille de points $g_1 \ldots g_M$ échantillonner $[U(g_m) | Y]$
5. Faire une carte

Syntaxe du modèle hiérarchique `model.bug`

```
model{

for(Dsite in 1:Nsite) {

  meansite[Dsite] <- intercept

  for(Dobservations in Ssite[Dsite]:(Ssite[Dsite+1]-1)){

    count[Dobservations] ~ dpois(meanobservations[Dobservations])
    log(meanobservations[Dobservations]) <- Rsite[Dsite]+ logOffset[Dobservations]

  }#observations
}#site


Rsite[1:NsiteSpatial] ~ spatial.exp(meansite[1:NsiteSpatial], xSpatialsite[1:NsiteSpatial],
ySpatialsite[1:NsiteSpatial], Tsite, phisite, 1)
phisite ~ dgamma(10,.1)


# priors

intercept ~ dflat()

Tsite <- pow(SDsite, -2)
SDsite ~ dunif(0, 100)

} # model
```

```
> library(glmmBUGS)
> load("rongelapUTM.RData")
> names(rongelapUTM)

[1] "x"     "y"     "count" "time"

> rongelapUTM$logOffset = log(rongelapUTM$time)
> rongelapUTM$site = seq(1, dim(rongelapUTM@data)[1])
```

```
library(geoRglm)
#data(rongelap)

rongelap<-read.table(rongelap.txt",header=TRUE)

library(spdep)
coordinates(rongelap) = ~cX+cY


rongelap$logOffset = log(rongelap$time)
rongelap$site = seq(1, length(rongelap$time))
```

Implémentation :MCMC

```
> forBugs = glmmBUGS(
+count + logOffset ~ 1, family="poisson",
+  data=rongelapUTM@data, effects="site",
+  spatial=rongelapUTM,
+  priors=list(phisite="dgamma(10,.1)"))
```

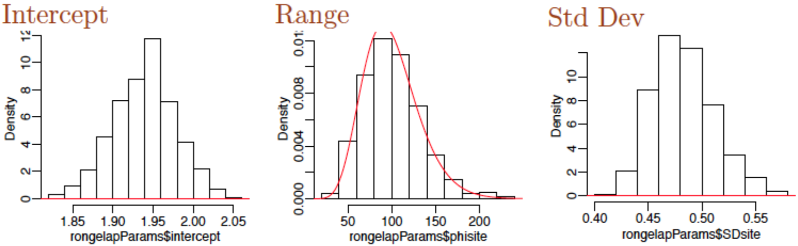
- `count` variable réponse ,
- `logOffset` offset, modèle avec seulement l'intercept, on ajoute `~1`
- un effet aléatoire pour chaque site
- L'effet aléatoire est spatialité avec les sites fournis par `rongelapUTM`

```
> startingValues = forBugs$startingValues
> startingValues$phi$site = 100
> source("getInits.R")
> library(R2WinBUGS)
>
```

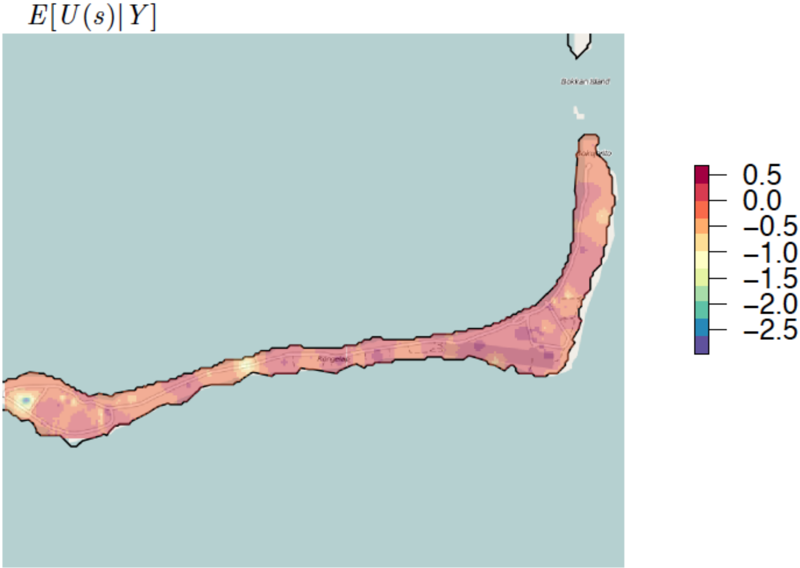
- `getInits.R` contient les initalisations des hyper paramètres `getInits`, qui utilise `startingValues`

Résultats

Apriori-Aposteriori



Résultats: U



Résultats : Dépassement de seuil pour U

$pr[U(s) > 0 | Y]$

Résultats: Prédiction du Risque

$E[\lambda(s)|Y]$

Résultats : Dépassement de seuil pour λ

$pr[\lambda(s) > 7 | Y]$