

Estimation de la distribution spatiale de la population de ragondins en Occitanie

Camille Mottier Iris Lang Justine Nicolas Olivier Gimenez

2 février 2026



Contexte et motivation de l'étude

Forte progression des
espèces exotiques envahissantes



Enjeu de suivi d'espèces

Contexte et motivation de l'étude

Forte progression des
espèces exotiques envahissantes



Enjeu de suivi d'espèces

Multiplication des données
de sciences participatives



Enjeu de traitement de données

Contexte et motivation de l'étude

Forte progression des
espèces exotiques envahissantes



Enjeu de suivi d'espèces



Multiplication des données
de sciences participatives



Enjeu de traitement de données



Objectif : Construction d'un modèle de distribution d'espèce prenant
en compte la spécificité des données de sciences participatives

Cas du ragondin en Occitanie : une espèce exotique envahissante



Spécificités des données de sciences participatives

- ▶ Données dites de présences seules : aucune information sur les absences
- ▶ Données d'observations et non de présences

Spécificités des données de sciences participatives

- ▶ Données dites de présences seules : aucune information sur les absences
- ▶ Données d'observations et non de présences

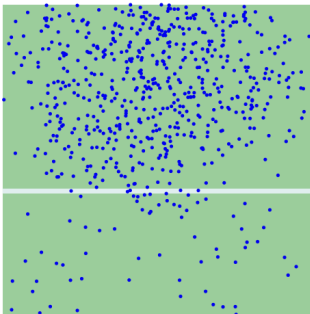
Zone d'étude



Spécificités des données de sciences participatives

- Données dites de présences seules : aucune information sur les absences
- Données d'observations et non de présences

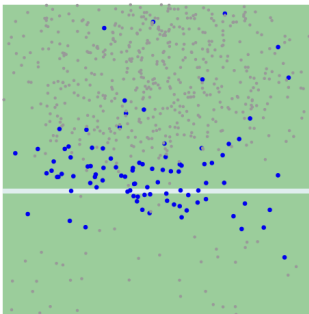
Présences de l'espèce



Spécificités des données de sciences participatives

- Données dites de présences seules : aucune information sur les absences
- Données d'observations et non de présences

Observations de l'espèce



Processus ponctuel

Un processus ponctuel \mathbf{X} sur une région $D \subset \mathbb{R}^2$, est un processus stochastique générant un ensemble fini de points $\mathbf{s} = \{s_1, \dots, s_n\}$ de D , dont le cardinal n n'est pas connu a priori.

- ▶ adapté aux données spatialisées
- ▶ à interpréter dans le contexte comme le processus régissant les présences réelles de l'espèce

Processus ponctuel de Poisson

Un processus ponctuel de Poisson, de fonction d'intensité λ , est un processus \mathbf{X} vérifiant pour tout $B \subset D$:

- ▶ $N(B) = \text{card}(\mathbf{X} \cap B) \hookrightarrow \mathcal{P} \left(\int_B \lambda(s) ds \right)$
- ▶ les points de $\mathbf{X} \cap B$, conditionnés par $N(B)$, sont indépendants et identiquement distribués avec pour densité :
$$f(s) = \frac{\lambda(s)}{\int_B \lambda(u) du}$$

Processus ponctuel de Poisson

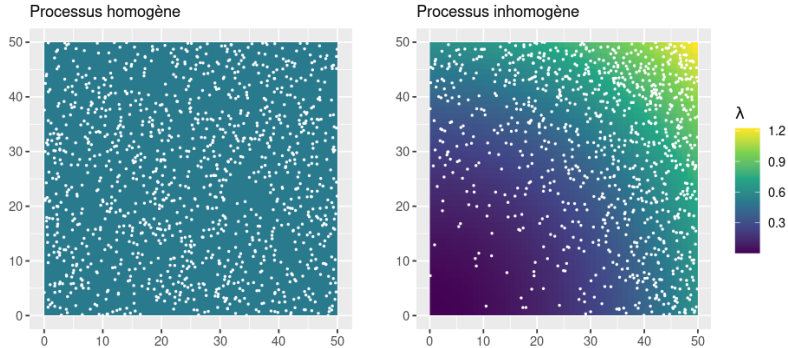
Un processus ponctuel de Poisson, de fonction d'intensité λ , est un processus \mathbf{X} vérifiant pour tout $B \subset D$:

- ▶ $N(B) = \text{card}(\mathbf{X} \cap B) \hookrightarrow \mathcal{P} \left(\int_B \lambda(s) ds \right)$
- ▶ les points de $\mathbf{X} \cap B$, conditionnés par $N(B)$, sont indépendants et identiquement distribués avec pour densité : $f(s) = \frac{\lambda(s)}{\int_B \lambda(u) du}$

- ▶ λ fonction positive caractéristique du processus de Poisson
- ▶ $\lambda(s)$: nombre moyen d'individus de l'espèce par unité d'aire à la position s
- ▶ Contrainte d'indépendance sur les présences, peu vraisemblable en pratique

Inhomogénéité

Les variations de λ permettent de traduire la disparité dans la répartition de la population étudiée.



Amincissement : la prise en compte de l'effort d'observation

L'**amincissement** de \mathbf{X} par $p : D \rightarrow [0, 1]$ est le processus ponctuel obtenu en décidant pour chaque point s du processus \mathbf{X} , et de manière indépendante

- ▶ de le garder avec probabilité $p(s)$,
- ▶ ou de le supprimer avec probabilité $1 - p(s)$.

Ceci définit un nouveau processus ponctuel de Poisson, de fonction d'intensité λp .

Amincissement : la prise en compte de l'effort d'observation

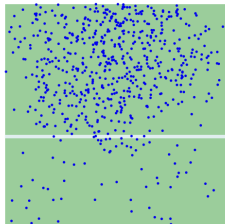
L'**amincissement** de \mathbf{X} par $p : D \rightarrow [0, 1]$ est le processus ponctuel obtenu en décidant pour chaque point s du processus \mathbf{X} , et de manière indépendante

- ▶ de le garder avec probabilité $p(s)$,
- ▶ ou de le supprimer avec probabilité $1 - p(s)$.

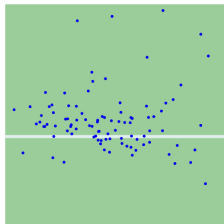
Ceci définit un nouveau processus ponctuel de Poisson, de fonction d'intensité λp .

- ▶ p est la probabilité qu'un individu soit observé

Processus ponctuel de Poisson inhomogène aminci : retour à l'exemple

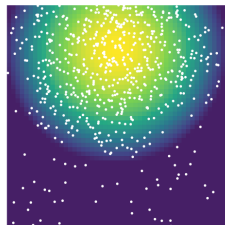


Présences $\{S_1, \dots, S_\ell\}$

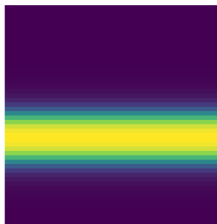


Observations $\{s_1, \dots, s_n\}$

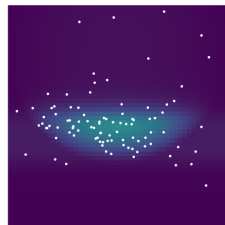
Processus ponctuel de Poisson inhomogène aminci : retour à l'exemple



Présences $\{S_1, \dots, S_\ell\}$
 $\mathbf{X} \sim \text{IPPP}(\lambda)$

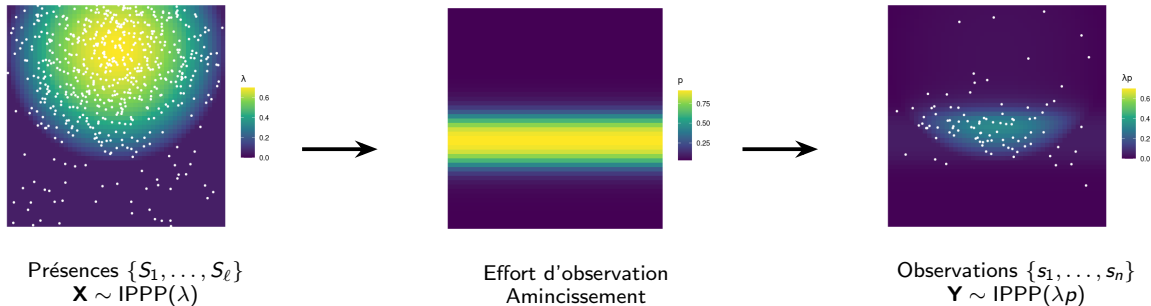


Effort d'observation
Amincissement



Observations $\{s_1, \dots, s_n\}$
 $\mathbf{Y} \sim \text{IPPP}(\lambda p)$

Processus ponctuel de Poisson inhomogène aminci : retour à l'exemple



- λ et p exprimées en fonction de variables d'environnements et d'effort :

$$\log(\lambda(s)) = \beta_0 + \beta_1 x_1(s) + \dots \beta_m x_m(s)$$

$$\text{logit}(p(s)) = \alpha_0 + \alpha_1 w_1(s) + \dots \alpha_q w_q(s)$$

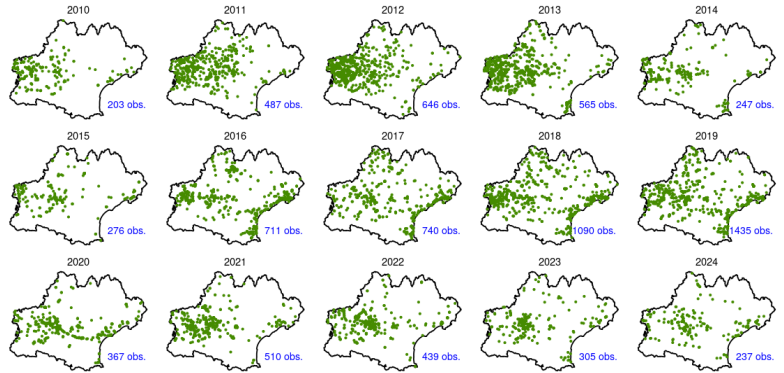
- Inférence bayésienne avec Nimble pour estimer les coefficients α et β et toutes les quantités d'intérêt

Le ragondin en Occitanie

Le ragondin, une espèce
exotique envahissante



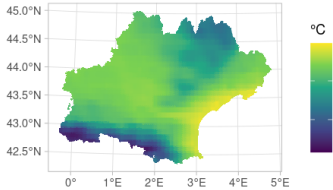
Données provenant du Système d'information de l'inventaire du
patrimoine naturel (SINP)



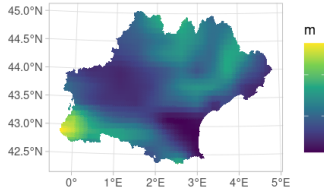
Les variables d'intensité

Variables d'intensité pour l'année 2016

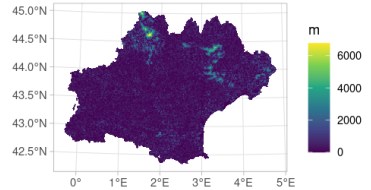
Température moy. mensuelle min.



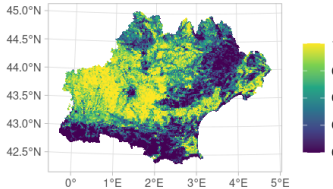
Précipitation annuelle cumulée



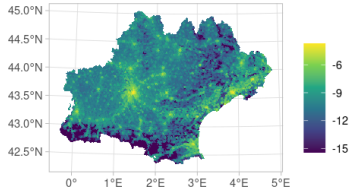
Distance à l'eau la plus proche



Proportion de surface agricole

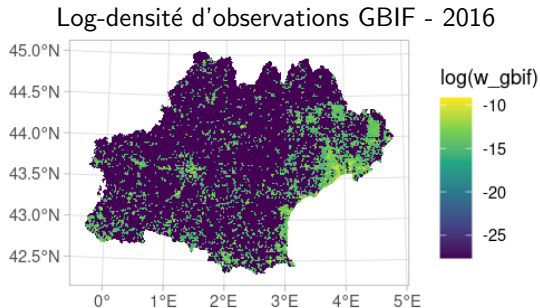


Log-densité de population

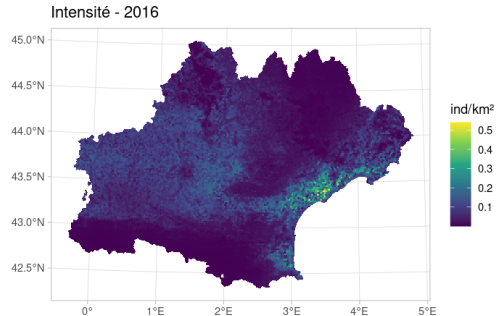
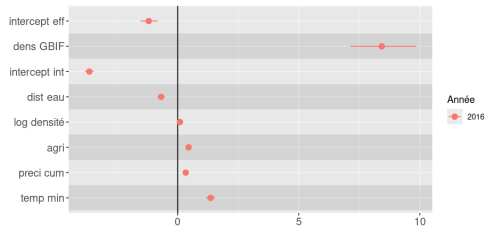


La variable d'effort

Une approche par densité d'observations multi-espèces (GBIF - Global Biodiversity Information Facility).



Exemple de l'année 2016

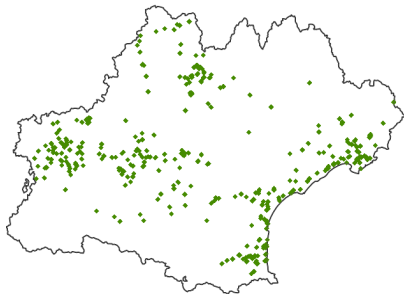


Validation du modèle

Le modèle ajusté explique-t-il le clustering observé ?

- Validation sur intensité amincie uniquement
- on compare les données observées avec des simulations faites à partir du modèle

Observations - 2016

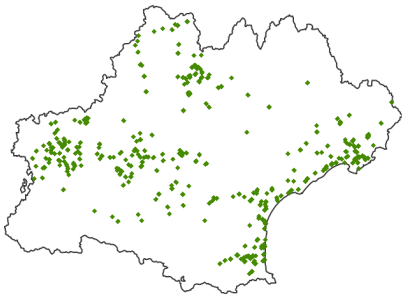


Validation du modèle

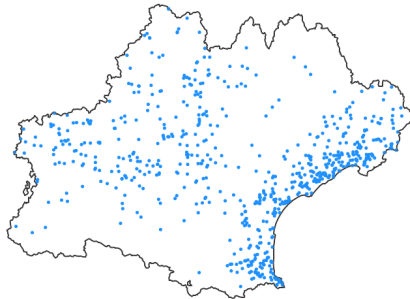
Le modèle ajusté explique-t-il le clustering observé ?

- Validation sur intensité amincie uniquement
- on compare les données observées avec des simulations faites à partir du modèle

Observations - 2016



Simulations IPPP ajusté - 2016



Fonction K inhomogène

Le modèle ajusté explique-t-il le clustering observé ?

- une mesure du clustering : la fonction K de Ripley inhomogène

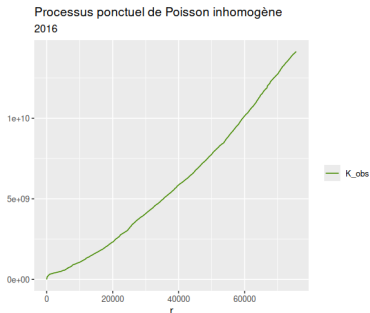
$$K_{\text{inhom}}(r) = \frac{1}{|D|} \mathbb{E} \left[\sum_{s_i \in \mathbf{X} \cap D} \sum_{s_j \in \mathbf{X} \setminus \{s_i\}} \frac{\mathbb{1}(\|s_i - s_j\| \leq r)}{\lambda(s_i)\lambda(s_j)} \right], \quad r \geq 0$$

Fonction K inhomogène

Le modèle ajusté explique-t-il le clustering observé ?

- une mesure du clustering : la fonction K de Ripley inhomogène

$$K_{\text{inhom}}(r) = \frac{1}{|D|} \mathbb{E} \left[\sum_{s_i \in \mathbf{X} \cap D} \sum_{s_j \in \mathbf{X} \setminus \{s_i\}} \frac{\mathbb{1}(\|s_i - s_j\| \leq r)}{\lambda(s_i)\lambda(s_j)} \right], \quad r \geq 0$$

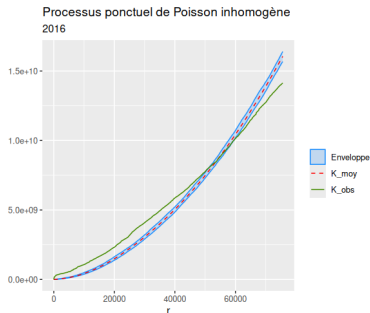


Fonction K inhomogène

Le modèle ajusté explique-t-il le clustering observé ?

- une mesure du clustering : la fonction K de Ripley inhomogène

$$K_{\text{inhom}}(r) = \frac{1}{|D|} \mathbb{E} \left[\sum_{s_i \in \mathbf{X} \cap D} \sum_{s_j \in \mathbf{X} \setminus \{s_i\}} \frac{\mathbb{1}(\|s_i - s_j\| \leq r)}{\lambda(s_i)\lambda(s_j)} \right], \quad r \geq 0$$

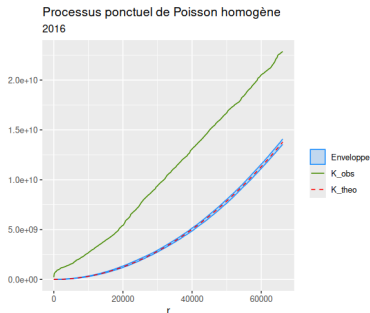


Fonction K inhomogène

Le modèle ajusté explique-t-il le clustering observé ?

- une mesure du clustering : la fonction K de Ripley inhomogène

$$K_{\text{inhom}}(r) = \frac{1}{|D|} \mathbb{E} \left[\sum_{s_i \in \mathbf{X} \cap D} \sum_{s_j \in \mathbf{X} \setminus \{s_i\}} \frac{\mathbb{1}(\|s_i - s_j\| \leq r)}{\lambda(s_i)\lambda(s_j)} \right], \quad r \geq 0$$

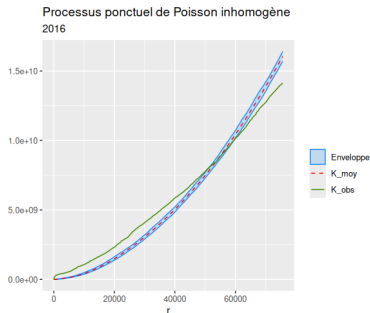


Fonction K inhomogène

Le modèle ajusté explique-t-il le clustering observé ?

- une mesure du clustering : la fonction K de Ripley inhomogène

$$K_{\text{inhom}}(r) = \frac{1}{|D|} \mathbb{E} \left[\sum_{s_i \in \mathbf{X} \cap D} \sum_{s_j \in \mathbf{X} \setminus \{s_i\}} \frac{\mathbb{1}(\|s_i - s_j\| \leq r)}{\lambda(s_i)\lambda(s_j)} \right], \quad r \geq 0$$



Question soulevée

Comment prendre en compte la dépendance spatiale ?

Dépendance spatiale

- Processus de Poisson : dépendance spatiale portée uniquement par les covariables du modèle

Dépendance spatiale

- ▶ Processus de Poisson : dépendance spatiale portée uniquement par les covariables du modèle
- ▶ Processus de Cox log-gaussien (LGCP) : dépendance spatiale introduite via un champ gaussien

Un processus ponctuel \mathbf{X} est un processus de Cox, s'il est régi par une intensité aléatoire Λ telle que :

$$\mathbf{X} | \Lambda \sim PPP(\Lambda)$$

Il est dit log-gaussien si :

$$\Lambda = \exp(Y) \text{ où } Y \text{ est un champ gaussien}$$

Dépendance spatiale

- ▶ Processus de Poisson : dépendance spatiale portée uniquement par les covariables du modèle
- ▶ Processus de Cox log-gaussien (LGCP) : dépendance spatiale introduite via un champ gaussien

Un processus ponctuel \mathbf{X} est un **processus de Cox**, s'il est régi par une intensité aléatoire Λ telle que :

$$\mathbf{X} | \Lambda \sim PPP(\Lambda)$$

Il est dit **log-gaussien** si :

$$\Lambda = \exp(Y) \text{ où } Y \text{ est un champ gaussien}$$

- ▶ Avec les covariables et l'amincissement :

$$Y(s) = \log(p(s)) + \beta x^t(s) + \underbrace{Z(s)}_{CG}$$

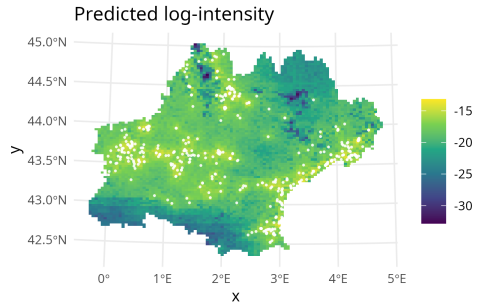
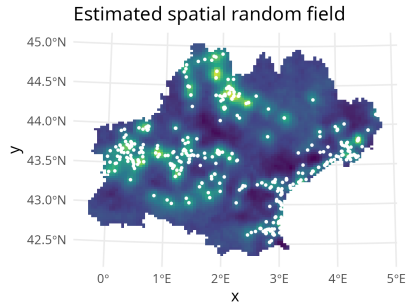
Ajustement du LGCP

- Ajustement à l'aide du package `inlabru`, basé sur l'approximation de Laplace et la discrétisation du champ gaussien

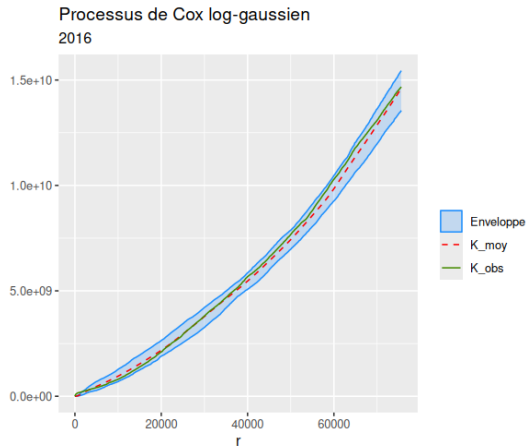
Ajustement du LGCP

- Ajustement à l'aide du package `inlabru`, basé sur l'approximation de Laplace et la discrétisation du champ gaussien

2016



Validation



Question soulevée

Comment prendre en compte la dépendance spatiale ?

Question soulevée

Comment prendre en compte la dépendance spatiale ?

Comment prendre en compte la dépendance temporelle ?

Modèle spatio-temporel

Considérer la période entière d'étude :

- ▶ pour gagner en précision
- ▶ pour tenir compte d'une éventuelle dépendance temporelle

Modèle spatio-temporel

Considérer la période entière d'étude :

- ▶ pour gagner en précision
- ▶ pour tenir compte d'une éventuelle dépendance temporelle

Quelle structure temporelle ?

$$Y(s, t) = \log(p(s, t)) + \beta x^t(s, t) + \underbrace{Z(s, t)}_{CG}$$

Modèle spatio-temporel

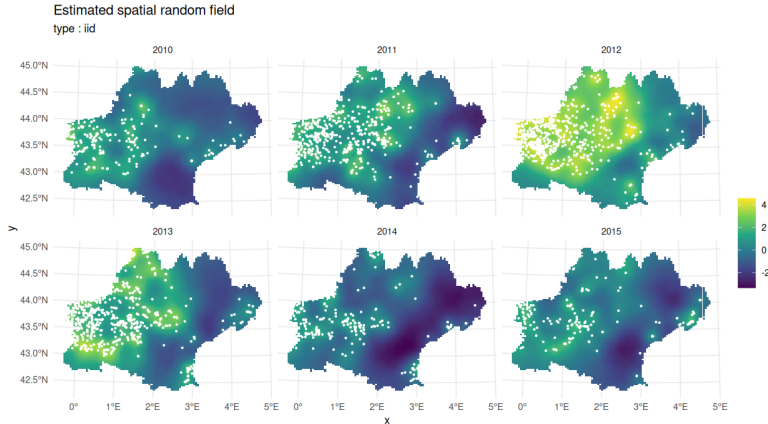
Quelle dépendance temporelle entre les champs gaussiens ?

- Des essais en considérant des paramètres constants pour la covariance

Modèle spatio-temporel

Quelle dépendance temporelle entre les champs gaussiens ?

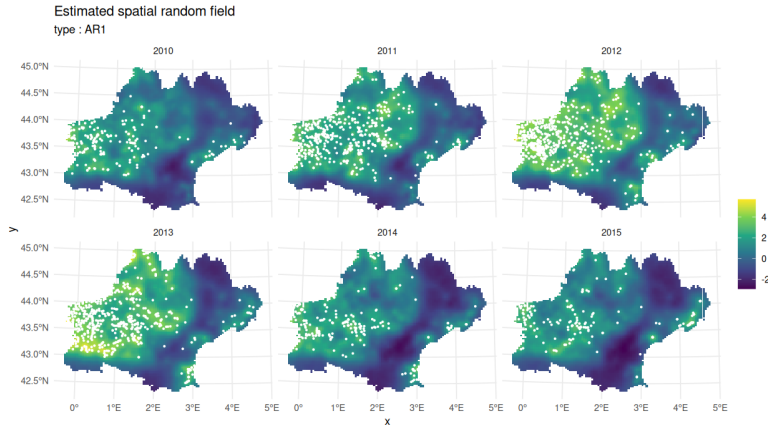
- Des essais en considérant des paramètres constants pour la covariance



Modèle spatio-temporel

Quelle dépendance temporelle entre les champs gaussiens ?

- Des essais en considérant des paramètres constants pour la covariance



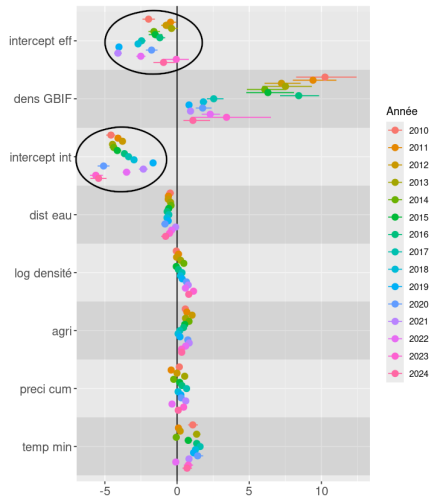
Conclusion et perspectives

- ▶ Des résultats encourageants
- ▶ Des problèmes d'identifiabilité
- ▶ Mieux comprendre le modèle à l'aide de simulations
- ▶ Construire et ajuster le modèle temporel



Merci

Problème d'identifiabilité des intercepts

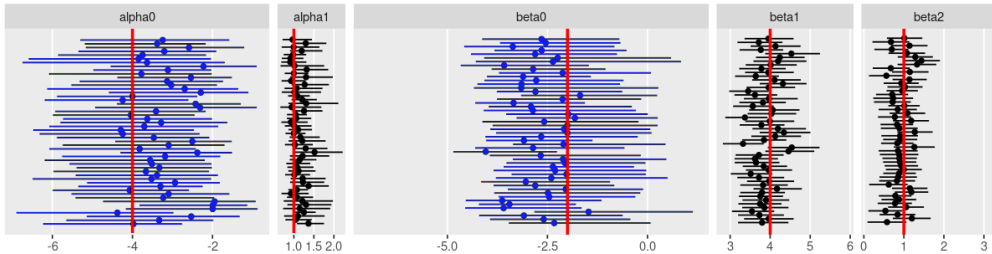


Problème d'identifiabilité des intercepts

Problème d'identifiabilité des intercepts confirmé par simulations, en particulier lorsque la probabilité de détection est faible partout :

$$\lambda(s)p(s) \simeq \exp(\beta x^t(s) + \alpha w^t(s))$$

Estimation des coefficients, cas d'un effort faible partout

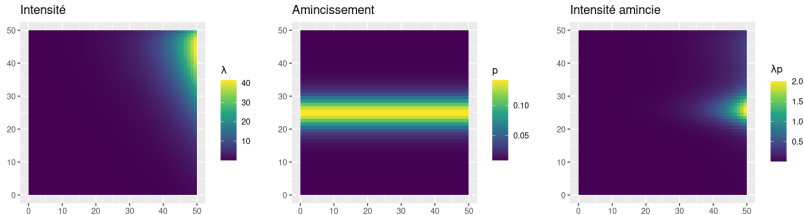


$$\log(\lambda(s)) = \beta_0 + \beta_1 x_1(s) + \beta_2 x_2(s), \quad \text{logit}(p(s)) = \alpha_0 + \alpha_1 w(s), \quad \text{avec } p(s) \leq 0.15$$

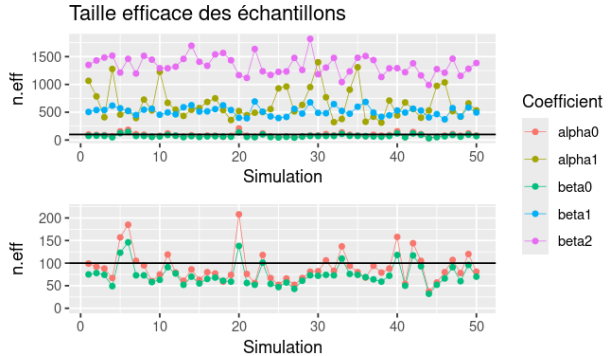
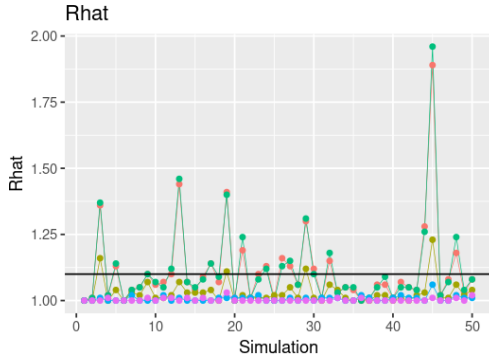
Cas simulé

Illustration sur des simulations d'un processus ponctuel de Poisson aminci

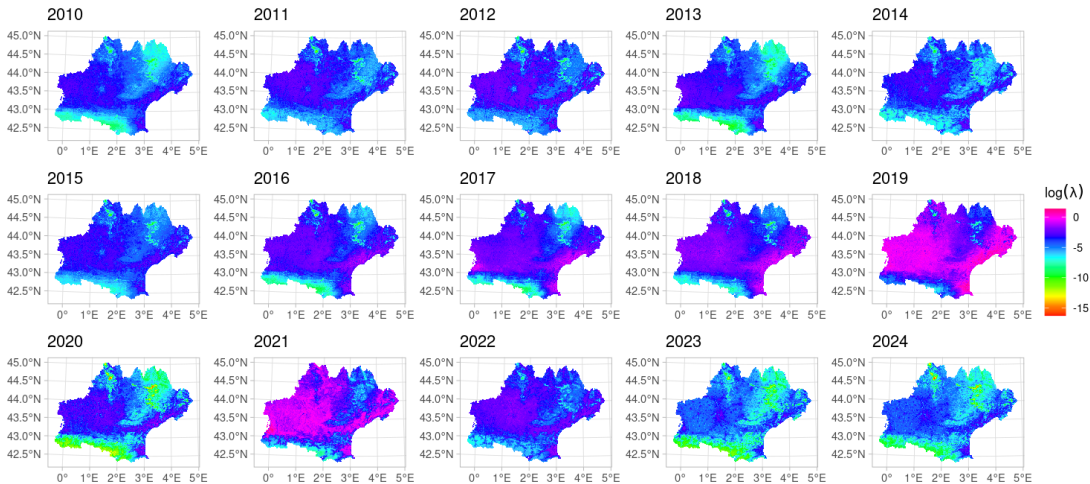
$$\log(\lambda(s)) = -4 + 4x_1(s) + x_2(s), \quad \text{logit}(p(s)) = -2 + 2w(s)$$



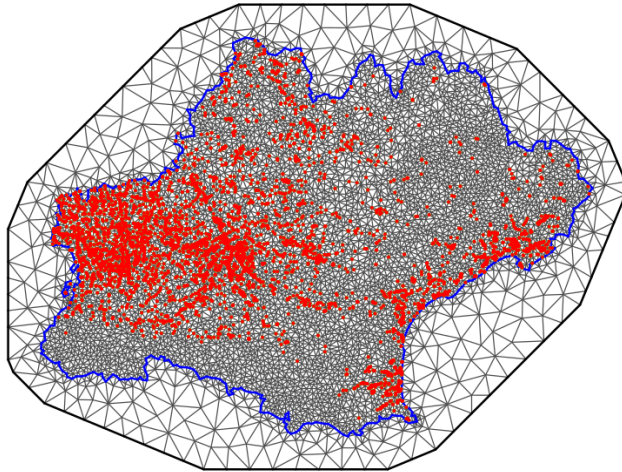
Diagnostics des simulations



Intensités (échelle log) - Modèles annuels

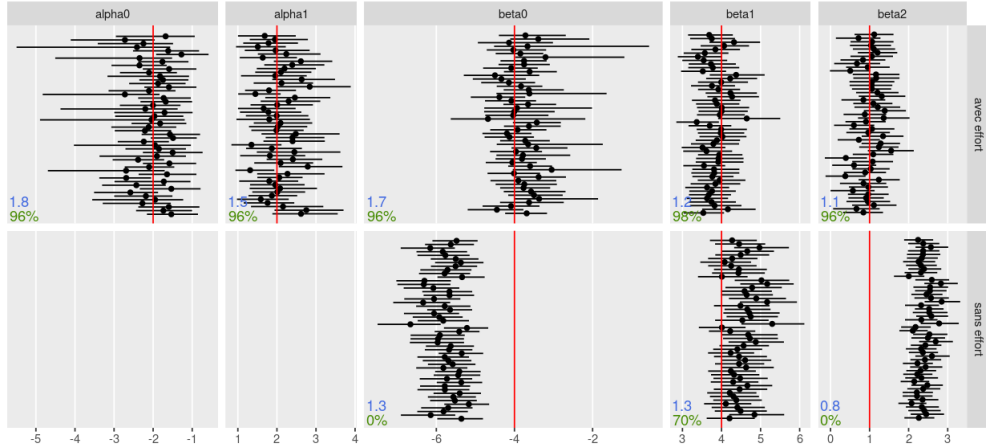


Mesh



Comparaison avec et sans effort

Estimation des coefficients, comparaison des modèles avec et sans effort



Vraisemblance

Vraisemblance d'un processus ponctuel de Poisson :

$$\begin{aligned} L(n, (s_1, \dots, s_n) | \alpha, \beta) &= \frac{1}{n!} \exp \left(- \int_D \lambda(s) p(s) ds \right) \left(\int_D \lambda(s) p(s) ds \right)^n \prod_{i=1}^n \frac{\lambda(s_i) p(s_i)}{\int_D \lambda(s) p(s) ds} \\ &= \frac{1}{n!} \exp \left(\int_D \frac{\exp(\beta x^t(s) + \alpha w^t(s))}{1 + \exp(\alpha w^t(s))} ds \right) \prod_{i=1}^n \frac{\exp(\beta x^t(s_i) + \alpha w^t(s_i))}{1 + \exp(\alpha w^t(s_i))} \end{aligned}$$

► Approximation de l'intégrale après découpage de la région d'étude en une grille de cellules :

$$\int_D \lambda(s) p(s) ds \simeq \sum_{i=1}^m a_i \lambda_i p_i$$