

Processus de Hawkes non stationnaire

Paul Pierrat

Septembre 2023

On présente dans un premier temps le modèle de Hawkes non stationnaire ainsi que la fonction de vraisemblance associée. On ajuste ensuite le processus aux données épidémiques de libération de spores en considérant la covariable humidité des feuilles.

1 La fonction d'intensité

On considère une intensité conditionnelle de la forme :

$$\lambda^*(t) = \lambda(t) + \sum_{t_i < t} \mu(t - t_i) \quad (1)$$

avec t_i les instants d'arrivées.

La fonction $\mu(\cdot)$ est la fonction d'excitation exponentielle :

$$\mu(t) = \alpha \exp(-\beta t) \quad (2)$$

avec $\alpha, \beta > 0$.

La fonction $\lambda(\cdot)$ est **l'intensité d'arrière plan**. Elle correspond à une intensité log-linéaire d'un processus de Poisson non homogène :

$$\log(\lambda(t)) = \theta^\top Z(t) \quad (3)$$

avec $\theta = (\theta_1, \dots, \theta_p)$ le vecteur des paramètres et $Z(t) = (Z_1(t), \dots, Z_p(t))$ le vecteur des covariables. Par la suite on considérera une fonction d'intensité log-linéaire dépendant d'une seule covariable :

$$\log(\lambda(t)) = \theta_1 + \theta_2 f(t) \quad (4)$$

avec $f(t)$ la valeur de la covariable à l'instant t .

Remarque 1.1. Pour le processus de Hawkes stationnaire l'intensité d'arrière plan s'écrit :

$$\lambda(t) = \lambda \quad (5)$$

où $\lambda > 0$.

Année	θ_1	θ_2	α	β
2011	-5.0	3.7	0.21	0.35
2012	-5.1	3.5	0.30	0.47

TABLE 1 – Estimations des paramètres par MLE.

2 La fonction de vraisemblance

En se restreignant à $T = t_k$, la fonction de log-vraisemblance de $\theta = (\theta_1, \theta_2, \alpha, \beta)$ s'écrit :

$$l(\theta|t_1, \dots, t_k) = \sum_{i=1}^k \ln(\exp(\theta_1 + \theta_2 f(t_i)) + \alpha A(i)) - \lambda t_k + \frac{\alpha}{\beta} \sum_{i=1}^k (\exp(-\beta(t_k - t_i)) - 1) \quad (6)$$

avec $A(i) = \sum_{j=1}^{i-1} \exp(-\beta(t_i - t_j))$ pour $i > 1$, et $A(1) = 0$ sinon.

On définit l'estimateur du maximum de vraisemblance par :

$$\hat{\theta} = \operatorname{argmax}_{\theta} l(\theta|t_1, \dots, t_k). \quad (7)$$

"The MLE $\hat{\theta}$ can be obtained by maximizing $l(\theta)$ over θ , or by solving the score equation."
[Chen and Hall, 2013].

3 Mise en oeuvre numérique

On ajuste le processus de Hawkes non stationnaire aux données de libérations de spores en considérant une covariable : l'humidité des feuilles. Les paramètres sont estimés avec l'estimateur du maximum de vraisemblance (7). Les résultats des ajustements sont présentés Table 1.

Le calcul du gradient par différences finies montre que le maximum est atteint aux points trouvés (maximum 10^{-3}). L'intensité conditionnelle et l'intensité d'arrière plan sont représentées Figure 1. On observe que les instants d'arrivées sont localisés sur les intervalles où l'intensité d'arrière plan est élevée.

Le critère AIC pour les différents modèles est présenté Table 2. Celui du modèle de Hawkes non stationnaire est le meilleur.

La Figure 2 donne le graphique quantiles-quantiles des résidus du processus de Hawkes stationnaire avec une intensité d'arrière plan constante et une fonction d'excitation exponentielle (à gauche) ; et le QQ-plot des résidus du processus de Hawkes non stationnaire dépendant de l'humidité des feuilles.

Références

[Chen and Hall, 2013] Chen, F. and Hall, P. (2013). Inference for a nonstationary self-exciting point process with an application in ultra-high frequency financial data modeling. *Journal of Applied Probability*, 50(4) :1006–1024.

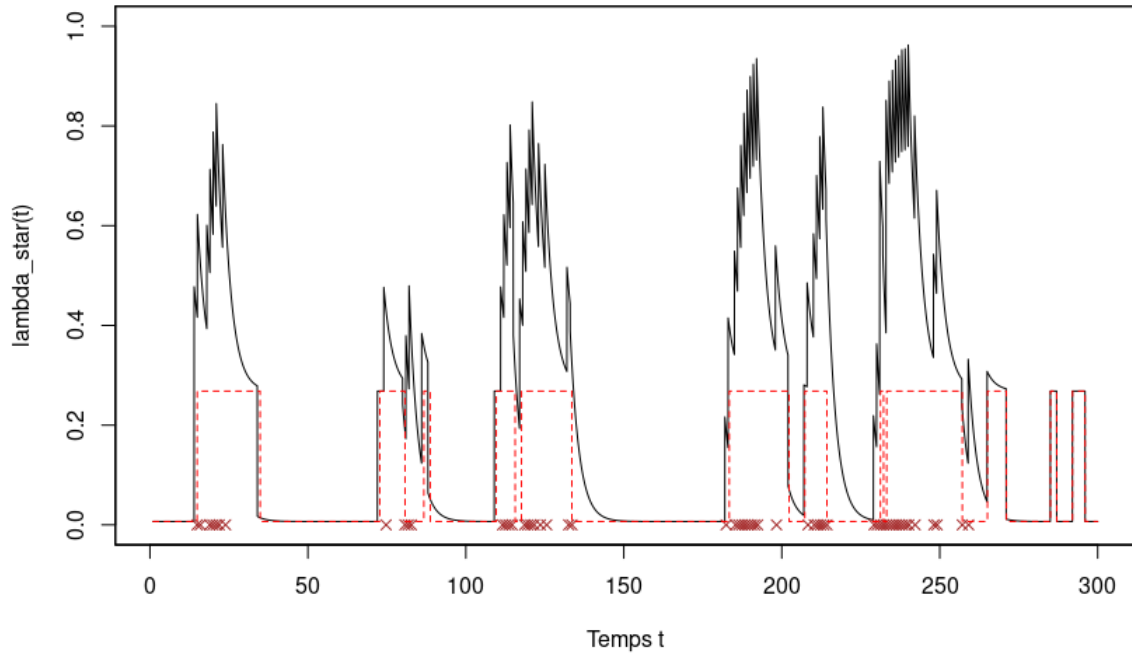


FIGURE 1 – Intensité conditionnelle ajustée au processus de Hawkes non stationnaire pour les données de libérations de spores en 2011 (en noir). Intensité d'arrière plan dépendant de l'humidité des feuilles (en rouge pointillé), et instants de libérations de spores (croix marron).

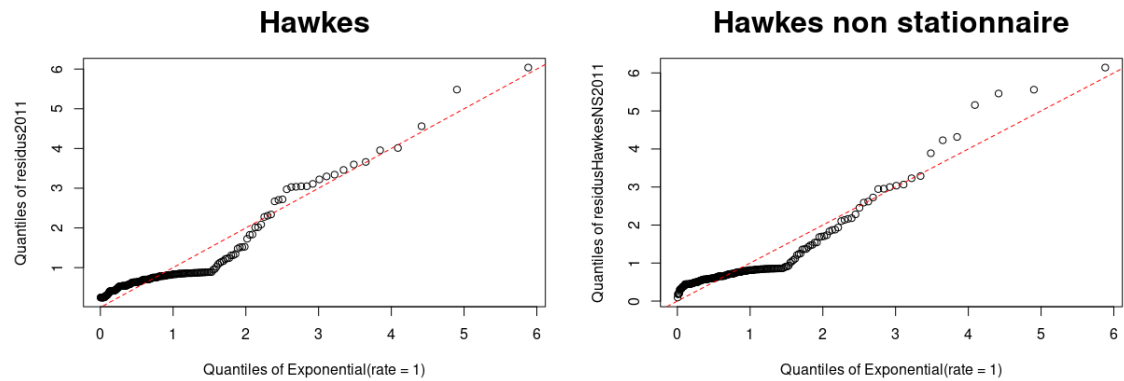


FIGURE 2 – QQ-plot des résidus du processus de Hawkes stationnaire et non stationnaire.

AIC	Année 2011	Année 2012
Poisson non homogène	983	751
Hawkes stationnaire	917	726
Hawkes non stationnaire	833	674

TABLE 2 – Comparaison des AIC pour le modèle de Poisson non homogène avec une intensité log-linéaire dépendant de l’humidité des feuilles, de l’humidité de l’air et de la luminosité, le modèle de Hawkes stationnaire avec une fonction d’excitation exponentielle et le modèle de Hawkes non stationnaire avec une intensité log-linéaire dépendant de l’humidité des feuilles et une fonction d’excitation exponentielle.