

# REVUE DE STATISTIQUE APPLIQUÉE

S. MONCAUP

P. SARDA

P. VIEU

## **Une mise en œuvre d'estimateurs non paramétriques d'une fonction d'intensité sur des données météorologiques**

*Revue de statistique appliquée*, tome 43, n° 4 (1995), p. 77-88

[http://www.numdam.org/item?id=RSA\\_1995\\_\\_43\\_4\\_77\\_0](http://www.numdam.org/item?id=RSA_1995__43_4_77_0)

© Société française de statistique, 1995, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

# UNE MISE EN ŒUVRE D'ESTIMATEURS NON PARAMÉTRIQUES D'UNE FONCTION D'INTENSITÉ SUR DES DONNÉES MÉTÉOROLOGIQUES

S. Moncaup, P. Sarda, P. Vieu

*Laboratoire de Statistique et Probabilités  
U.A. C.N.R.S 745 Université Paul Sabatier  
118 route de Narbonne 31062 Toulouse cedex*

## RÉSUMÉ

Nous nous intéressons à l'estimation non paramétrique de la fonction d'intensité d'un processus de Poisson doublement stochastique. Une méthode de sélection du paramètre de lissage basée sur la minimisation d'un critère local de validation croisée est proposée, ainsi qu'une application à un problème météorologique.

**Mots-clés :** *Intensité, Processus de Cox, Estimation non paramétrique, Degré de lissage, Données météorologiques.*

## ABSTRACT

This paper deals with the nonparametric estimation of the intensity function of a doubly stochastic Poisson process. A data-driven smoothing parameter based on the minimization of a local cross-validation criterion is proposed, as well as an application to meteorological data.

**Keywords :** *Intensity, Cox processes, Nonparametric estimation, Bandwidth choice, Meteorological data.*

## 1. Introduction

Le travail présenté ici porte sur l'estimation non paramétrique de la fonction d'intensité d'un processus ponctuel. Le point de départ de cette étude est un problème météorologique concernant les phénomènes de formation des gouttes dans les nuages. Ainsi, on est amené à déterminer l'évolution de la concentration des particules d'eau dans un nuage afin d'analyser son effet sur les développements des précipitations et les propriétés radiatives de ce nuage. Pour cela, on analyse les mesures effectuées dans ce nuage avec des instruments embarqués sur avion. Plus précisément, on dispose, à l'intérieur d'un même nuage, d'une série de temps d'arrivée de gouttes d'eau dans un spectromètre permettant de compter et de dimensionner les particules le long de la trajectoire de l'avion. On veut donc décrire un phénomène aléatoire pour lequel

plusieurs hypothèses sur ses propriétés statistiques sont posées. Le comptage des gouttes est supposé être un processus de Poisson non homogène de fonction d'intensité  $\lambda$ . Bien que l'indépendance du nombre de gouttes dans deux régions disjointes du nuage semble plausible, une étude ultérieure serait souhaitable pour confirmer ou infirmer cette hypothèse. Généralement, lorsque la trajectoire de l'avion dans le nuage est suffisamment longue pour avoir un grand nombre de particules, on suppose que les variations observées de la concentration de ces particules ne dépendent pas de la trajectoire : la nature aléatoire du comptage est ainsi gommée. Cependant, depuis quelques années, on s'intéresse à des régions du nuage de plus en plus petites et donc on doit considérer que l'intensité  $\lambda$  est elle-même aléatoire. Les données dont on dispose sont ainsi supposées être la réalisation d'un processus de Poisson doublement stochastique (encore appelé processus de Cox : cf section 2).

Déterminer la concentration en eau du nuage revient à estimer la fonction  $\lambda$ , et on est en particulier intéressé à la détection d'éventuelles ruptures dans l'intensité de la concentration en eau. Toutefois, dans la méthode d'estimation choisie pour la fonction d'intensité on ne fait pas *a priori* d'hypothèse sur l'existence de telles ruptures. On ne fait donc, au départ, aucune hypothèse sur la forme de la fonction  $\lambda$ , et l'estimateur utilisé est un estimateur non paramétrique, au sens où il ne suppose pas l'appartenance de la fonction d'intensité à une classe de fonctions indexée par un nombre fini de paramètres réels. Nous utilisons ici l'estimateur à noyau défini dans la section suivante. Il présente l'intérêt d'une mise en œuvre très rapide par rapport à des méthodes alternatives de type paramétrique (reposant par exemple sur un estimateur de maximum de vraisemblance) qui nécessitent par ailleurs des hypothèses supplémentaires dans le modèle général.

Dans la pratique, l'utilisation d'un estimateur non paramétrique conduit, entre autres, au problème de choix d'un paramètre de lissage. Les premiers résultats établis par Diggle (1985) puis Diggle et Marron (1988) concernant des expressions d'erreurs quadratiques montrent que le problème du choix du paramètre de lissage est crucial dans la qualité de l'estimateur à noyau. Ces résultats sont rappelés dans la section 2. La bibliographie en matière de choix de paramètre de lissage sera présentée dans la section 3 dans laquelle une nouvelle méthode, basée sur les idées de validation croisée locale, sera présentée. La mise en œuvre sur les données météorologiques présentées ci-dessus fait l'objet de la section 4.

## 2. Le modèle de Cox et l'estimateur à noyau

On dispose de  $N$  points  $X_i$  sur un intervalle  $[0, T]$ . On veut déterminer «l'intensité locale»  $\lambda(x)$  de ces points au voisinage de tout point  $x$  de  $[0, T]$ . Le modèle introduit consiste à supposer que ces points constituent une réalisation partielle d'un processus de Cox stationnaire, de processus d'intensité  $\{\wedge(x), x \in \mathbb{R}\}$ . Ces processus sont également appelés «processus de Poisson doublement stochastiques» (Cox et Isham, 1980). Ainsi les deux conditions suivantes sont supposées être vérifiées :

- $\{\wedge(x), x \in \mathbb{R}\}$  est un processus aléatoire réel stationnaire, à valeurs positives, de moyenne  $\mu$  et on notera

$$\mathbb{E}[\wedge(x) \wedge(y)] = \nu(|x - y|), \quad \forall(x, y) \in \mathbb{R}^2;$$

- conditionnellement à la réalisation  $\lambda(x)$  de  $\wedge(x)$ , les points  $X_i$  sont générés selon un processus de Poisson non homogène, de fonction d'intensité  $\lambda(x)$ .

Le cadre choisi pour estimer l'intensité  $\lambda$  est celui de l'estimation non paramétrique : aucune hypothèse n'est faite concernant l'appartenance de la fonction  $\lambda$  à une classe paramétrée de fonctions. Un estimateur non paramétrique de l'intensité  $\lambda(x)$ , du type estimateur à noyau (Rosenblatt, 1956, Diggle, 1985), est alors donné par

$$\widehat{\lambda}_h(x) = \sum_{i=1}^N \delta_h(x - X_i), \quad x \in [0, T], \quad (2.1)$$

où  $h$  est un paramètre de lissage strictement positif, encore appelé largeur de fenêtre, et  $\delta_h(\cdot) = h^{-1}\delta(\cdot/h)$ ,  $\delta$  étant une densité de probabilité symétrique encore appelée noyau. Des résultats théoriques concernant l'estimateur à noyau d'une fonction d'intensité sont obtenus par Rammlau-Hansen (1983) et Diggle (1985). On se rapportera aussi à Crétois (1994) pour une bibliographie plus large sur ces estimateurs ainsi que pour des méthodes alternatives.

Le comportement de l'estimateur présenté en (2.1) est en particulier lié au choix du paramètre de lissage  $h$ . Plus on prend  $h$  grand plus le biais de l'estimateur augmente avec pour conséquence un surlissage, tandis que si l'on utilise  $h$  trop petit c'est la variance qui se détériore. Cet effet de largeur de fenêtre est bien connu dans tous les problèmes d'estimation fonctionnelle. Ici il apparaît à la fois dans les expressions asymptotiques d'erreurs d'estimation (voir par exemple (2.5) ci-dessous) et dans l'étude réelle qui nous intéresse (voir figures 1,3 et 4). Si l'on considère l'erreur quadratique moyenne (EQM) comme une mesure de la qualité de l'estimateur

$$\text{EQM}(h) = \mathbb{E}[(\widehat{\lambda}_h(x) - \wedge(x))^2], \quad (2.2)$$

alors la meilleure largeur de fenêtre  $h_0$  est définie de la manière suivante

$$h_0 = \underset{h}{\text{Argmin}} \text{EQM}(h). \quad (2.3)$$

Le problème du choix de la largeur de fenêtre revient donc à estimer  $h_0$ . Diggle (1985) établit sous des conditions générales que

$$\text{EQM}(h) = \mu h^{-1} \int \delta^2(y) dy + \mu^2 h^{-1} K_{\delta * \delta}(h) - 2\mu^2 h^{-1} K_{\delta}(h) + \nu(0), \quad (2.4)$$

avec

$$K_{\delta}(h) = \int_0^{+\infty} 2\mu^{-2} \delta(uh^{-1}) \nu(u) du,$$

et

$$\delta * \delta(u) = \int \delta(u-x) \delta(x) dx.$$

Dans le cas particulier d'une fonction  $\nu$  de la forme  $\nu(x) = \mu^2 \nu_0(x)$  et sous des hypothèses de dérivabilité de  $\nu$ , Diggle et Marron (1988) précisent l'expression (2.4) et obtiennent

$$\text{EQM}(h) = \mu h^{-1} \int \delta^2(y) dy + \mu^2 h^4 \nu_0^{(4)}(0) \left[ \frac{1}{2} \int u^2 \delta(u) du \right]^2 + o(h^4). \quad (2.5)$$

On obtient ainsi la décomposition classique de  $\text{EQM}(h)$  en un terme lié à la variance et un terme de biais au carré. Des expressions du type (2.4) et (2.5) permettent d'avoir une expression simple de la fenêtre optimale  $h_0$  définie en (2.3). Malheureusement  $h_0$  dépend de quantités inconnues. La section 3 est centrée sur la question de choix optimaux et automatiques de fenêtre. Malgré l'importance de la question, la bibliographie sur ce sujet est peu abondante, bien moins en tout cas que dans d'autres problèmes d'estimation fonctionnelle. Dans la section 3, nous nous attarderons plus particulièrement sur la méthode de validation croisée (Brooks et Marron, 1990), et nous en proposerons une version nouvelle qui permet de faire évoluer ce paramètre de lissage avec le temps.

### 3. Méthode de validation croisée

Les techniques de validation croisée sont couramment utilisées pour les questions de choix d'estimateurs non paramétriques (voir Marron, 1988, Vieu, 1993, et Broniatowsky, 1993, pour des revues bibliographiques). Ces techniques sont surtout développées en estimation de régression ou de densité, et elles ont été adaptées récemment par Brooks et Marron (1990) à notre problème. En se fixant comme objectif la minimisation de l'erreur quadratique intégrée,

$$\begin{aligned} \text{EQI}(h) &= \int_0^T [\hat{\lambda}_h(x) - \lambda(x)]^2 dx \\ &= \int_0^T \hat{\lambda}_h(x)^2 dx - 2 \int_0^T \hat{\lambda}_h(x) \lambda(x) dx + \int_0^T \lambda(x)^2 dx, \end{aligned} \quad (3.1)$$

ces auteurs proposent d'estimer la fenêtre optimale

$$h_1 = \underset{h}{\text{Argmin}} \text{EQI}(h), \quad (3.2)$$

en utilisant le critère suivant,

$$\text{CV}(h) = \int_0^T \hat{\lambda}_h(x)^2 dx - 2 \sum_{i=1}^n \sum_{j \neq i} h^{-1} \delta((X_i - X_j)/h). \quad (3.3)$$

La fenêtre sélectionnée par validation croisée est celle qui minimise ce critère :

$$\hat{h}_{\text{CV}} = \underset{h}{\text{Argmin}} \text{CV}(h). \quad (3.4)$$

De manière intuitive, ce critère se justifie par le fait que chacun de ses termes est un estimateur des deux premiers termes que l'on obtient en développant l'expression de  $\text{EQI}(h)$  (le troisième terme de  $\text{EQI}(h)$  ne dépendant pas de  $h$ ). Plus précisément, des résultats d'optimalité asymptotique du paramètre sont obtenus par Brooks et Marron (1990) pour un processus de Poisson non homogène à intensité multiplicative  $\lambda(x) = c \cdot \alpha(x)$  où  $c$  est une constante positive et  $\alpha$  une fonction déterministe. Dans ce cadre, Brooks et Marron ont en effet démontré que :

$$\frac{\text{EQI}(\hat{h}_{\text{CV}})}{\text{EQI}(h_1)} \longrightarrow 1, p.s., \text{ quand } c \longrightarrow \infty. \quad (3.5)$$

Des résultats concernant la vitesse de convergence dans (3.5) sont également obtenus par Grégoire (1993) dans un cadre légèrement différent. Concernant le modèle de Cox qui nous intéresse ici, Diggle et Marron (1988) ont montré (dans le cas particulier d'un noyau uniforme) l'équivalence entre ce critère et un estimateur direct de l'EQM construit à partir de l'expression (2.4) et d'estimateurs de  $\mu$  et  $K_\delta$ . Nous renvoyons à Diggle (1985) pour des détails sur cet estimateur de l'EQM.

Ce paramètre  $\hat{h}_{\text{CV}}$  est un paramètre global. Il minimise (du moins asymptotiquement) l'erreur globale  $\text{EQI}(h)$  et il est le même en tout point de  $[0, T]$ . On peut cependant envisager de choisir une largeur de fenêtre locale, c'est-à-dire dépendant du point  $x$  où la courbe  $\lambda$  est estimée. En effet, en un point  $x$  où la courbe présente une forte pente (par exemple près d'une rupture), un choix de fenêtre «petite» devrait conduire à une meilleure estimation alors qu'une fenêtre plus large serait préférable dans une zone où la courbe est plus «plate». On cherche ainsi à minimiser en chaque point  $x$  de  $[0, T]$  une version locale de  $\text{EQI}(h)$

$$\text{EQI}_x(h) = \int_0^T (\hat{\lambda}_h(u) - \lambda(u))^2 I_{[x-\varepsilon; x+\varepsilon]}(u) du, \quad (3.6)$$

où  $\varepsilon$  (dépendant de  $N$ ) est un réel positif fixé. Ici aussi,  $\text{EQI}_x(h)$  dépendant de l'intensité réelle  $\lambda$ , ces paramètres locaux (c'est-à-dire  $h_1(x)$  minimisant  $\text{EQI}_x(h)$ ) sont impossibles à calculer. L'idée est donc d'utiliser dans l'estimateur à noyau des paramètres locaux choisis comme minimisant un estimateur des termes de  $\text{EQI}_x(h)$  dépendant de  $h$ . La fonction de validation croisée locale est alors construite sur le même principe que le critère (3.3)

$$\begin{aligned} \text{CV}_x(h) = & \int_0^T \hat{\lambda}_h(u)^2 I_{[x-\varepsilon; x+\varepsilon]}(u) du \\ & - 2 \sum_{i=1}^N \sum_{j \neq i} h^{-1} \delta((X_j - X_i)/h) I_{[x-\varepsilon; x+\varepsilon]}(X_i). \end{aligned} \quad (3.7)$$

En chaque point  $x$ , nous proposons de choisir un paramètre local défini par

$$\hat{h}(x) = \underset{h}{\text{Argmin}} \text{CV}_x(h). \quad (3.8)$$

Le critère (3.7) est analogue à celui introduit par Mielniczuk *et al.* (1989) en estimation de densité. Dans la pratique, l'utilisation de  $CV_x(h)$  impose le choix de  $\varepsilon$  (qui fixe un intervalle autour du point  $x$ , dans lequel le critère de validation croisée est calculé). Hall et Schucany (1989) discutent, également pour l'estimateur à noyau de la densité (avec un noyau uniforme), le choix de  $\varepsilon$ . Ces auteurs montrent à l'aide de simulations que le choix de  $\varepsilon$  n'est pas un problème critique (en regard du problème du choix de la largeur de fenêtre) et suggèrent qu'un choix pertinent serait celui d'une largeur de fenêtre «pilote».

## 4. Application

### 4.1. Présentation de l'étude

L'application pratique concerne le traitement des données décrites dans l'introduction et qui nous ont été fournies par le centre météorologique de Toulouse (voir Brenguier, 1993, pour une description détaillée). On s'intéresse ici à une portion de la trajectoire (rectiligne) de l'avion : on dispose ainsi dans l'intervalle [46940, 47100] d'une série  $(X_i)_{i=1, \dots, 10229}$  représentant la position de 10229 gouttes (l'échelle étant exprimée en mm). D'après ce qui a été vu en introduction,  $(X_i)_{i=1, \dots, 10229}$  est considérée comme étant la réalisation partielle d'un processus de Cox. La concentration en eau est mesurée à l'aide de l'intensité  $\lambda(x)$  de ce processus.

Nous avons utilisé pour estimer  $\lambda(x)$ , l'estimateur à noyau présenté à la section 2 en sélectionnant le paramètre de lissage par la méthode de validation croisée introduite dans la section précédente. L'idée d'utiliser l'estimateur  $\hat{\lambda}_h$  se justifie par la simplicité de son écriture ainsi que par le grand nombre de données dont on dispose (puisque les méthodes non paramétriques à cause de leur caractère local nécessitent généralement des échantillons importants). Par ailleurs, n'ayant pas d'information *a priori* sur la forme de l'intensité  $\lambda$ , il est naturel de l'estimer en ne posant pas d'hypothèses restrictives sur celle-ci.

Bien que les résultats de type (2.5) montrent que la forme du noyau influe moins que le paramètre  $h$  sur la qualité de l'estimateur (du moins asymptotiquement), il serait intéressant dans un futur travail d'en étudier les effets sur des échantillons finis. Ici nous avons utilisé pour des commodités de calcul le noyau uniforme largement utilisé dans la littérature

$$\delta(\cdot) = 2^{-1} I_{[-1,1]}(\cdot).$$

### 4.2. Construction d'un estimateur globalement optimal

Nous avons tout d'abord utilisé un paramètre de lissage global, obtenu par la méthode de validation croisée définie par (3.3) et (3.4). La minimisation dans (3.4) est effectuée en considérant 80 valeurs de  $h$  comprises entre 0 et 5. La fonction de validation croisée  $CV(h)$  représente un estimateur de l'erreur quadratique intégrée  $EQI(h)$ . On remarque bien sur son tracé (figure 1) que  $CV(h)$  augmente quand  $h$  prend

des valeurs trop petites, et que  $CV(h)$  croît également quand  $h$  prend des valeurs trop grandes.

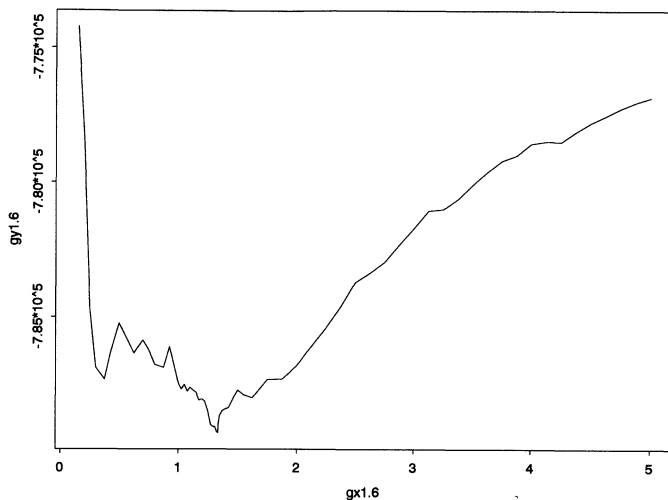


FIGURE 1 :

*Fonction de validation croisée globale (pour  $h$  variant de 0 à 5.)*

Ceci est conforme aux expressions asymptotiques précédentes (voir (2.5) par exemple). On sélectionne donc comme valeur de  $h$  celle pour laquelle la fonction CV atteint son minimum : 1,335.

L'estimateur optimal de l'intensité sera donc  $\hat{\lambda}_{1.335}$ . La figure 2 en est le tracé.

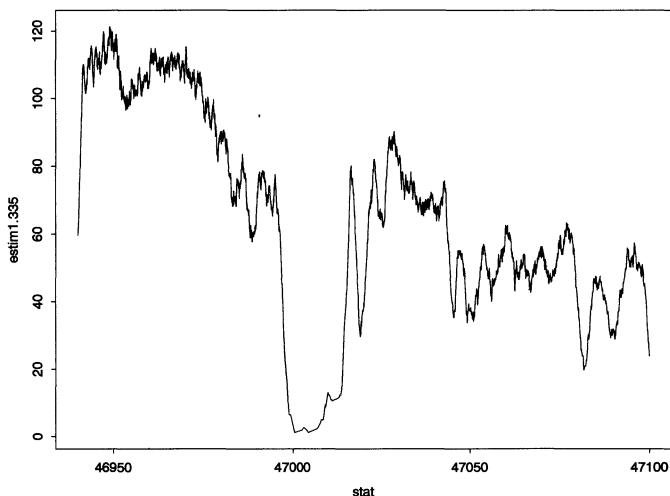


FIGURE 2 :

*Fonction d'intensité estimée ( $h = 1.335$ )*



L'importance du choix de  $h$ , qui apparaissait nettement sur l'expression de l'erreur quadratique moyenne en (2.5) ainsi que sur la courbe de la fonction de validation croisée, peut aussi être directement observée sur les figures 3 et 4 qui représentent l'estimateur  $\hat{\lambda}_h$  pour deux autres valeurs du paramètre  $h$ . On remarque en effet de grands changements dans l'allure de l'estimateur pour des valeurs de  $h$  très différentes. Comme on le voit sur la figure 3, pour une grande valeur de  $h$  (ici  $h = 20$ ) la fonction  $\hat{\lambda}_h$  est lisse et les variations importantes de l'intensité observées sur la figure 2 ont été gommées. A l'inverse, pour une petite valeur de  $h$  ( $h = 0, 1$ ), la courbe 4 oscille beaucoup trop. Le bon choix du paramètre  $h$  doit permettre de trouver un «compromis» entre ces deux phénomènes : la courbe de l'estimateur doit être suffisamment lisse aux endroits où l'intensité réelle varie peu, sans pour autant faire disparaître les changements d'intensité plus importants qui peuvent être les ruptures qui nous intéressent.

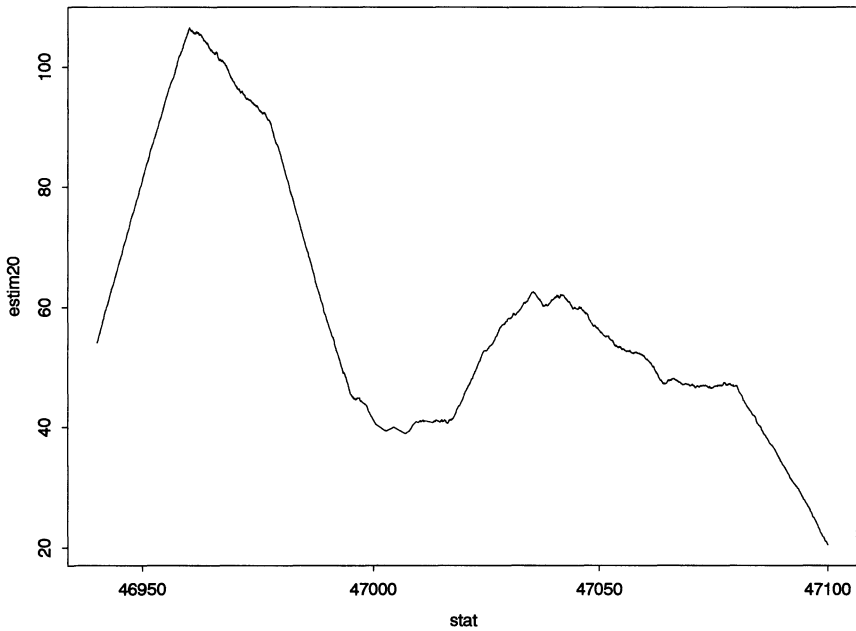


FIGURE 3 :  
*Fonction d'intensité estimée ( $h = 20$ )*

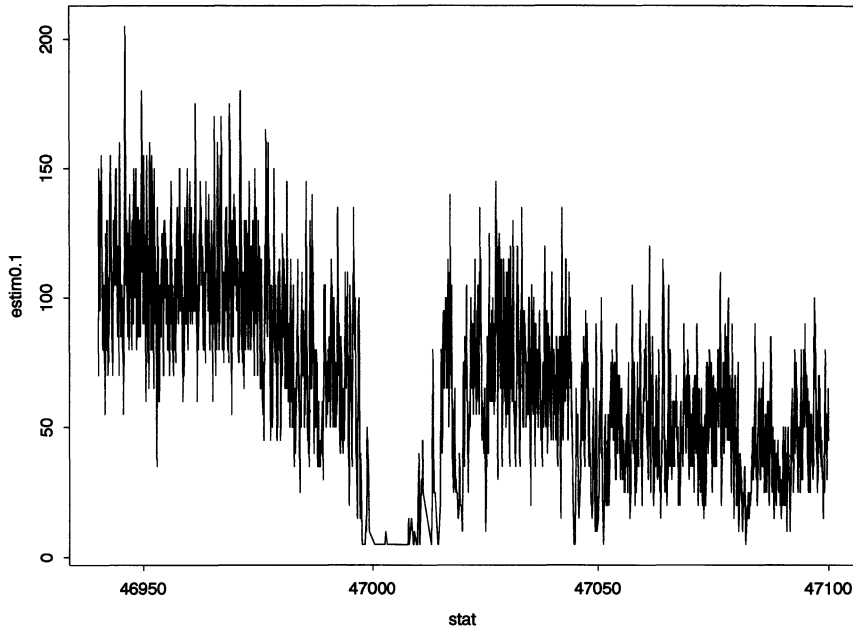


FIGURE 4 :

*Fonction d'intensité estimée ( $h = 0.1$ )*

Sur la figure 2, représentant le tracé de l'estimateur optimal, on observe, au niveau des points d'abscisses 46995 à 47015, de brusques changements dans la valeur de l'intensité estimée qui semblent révéler une rupture dans l'intensité réelle. On passerait d'un endroit du nuage fortement concentré en eau, à un endroit où la concentration est très faible, avant de revenir à une concentration plus importante. Ailleurs, les valeurs sont beaucoup moins changeantes, les variations ont moins d'amplitude, ce qui laisse supposer une plus grande constance de l'intensité de la concentration en eau dans les autres endroits du nuage. L'étude de cette courbe soulève tout de même des questions : si une rupture assez nette dans l'intensité est détectée au niveau des points d'abscisses 46995 à 47015, qu'en est-il au niveau des points d'abscisses 47045 et 47082? Ces deux «pics» sont-ils dûs seulement à du bruit, ou bien représentent-ils de réelles variations de l'intensité? On peut alors se demander si le choix de fenêtres locales permettrait une meilleure précision de l'estimateur à ces endroits-là.

### ***4.3. Construction d'un estimateur localement optimal***

Le tracé de l'estimateur  $\hat{\lambda}_h$  utilisant un paramètre  $h$  variant localement a été réalisé sur la figure 5. Les différentes valeurs du paramètre ont été sélectionnées par la méthode de validation croisée locale définie en (3.7). Comme nous l'avons vu plus haut, la valeur de  $\varepsilon$  influe peu sur la qualité de l'estimateur  $\hat{\lambda}_h$  (voir Hall et Schucany,

1989). Nous avons pris pour  $\varepsilon$  une valeur de l'ordre de la fenêtre globale obtenue par validation croisée :  $\varepsilon = 1,5$ . L'allure de cette courbe est très similaire à celle obtenue pour le paramètre  $h$  global (figure 2). La présence d'une rupture au niveau des points d'abscisses 46995 à 47015 est confirmée. Le peu de différence entre les courbes 2 et 5 tient ici essentiellement au fait que le nombre de données est très grand.

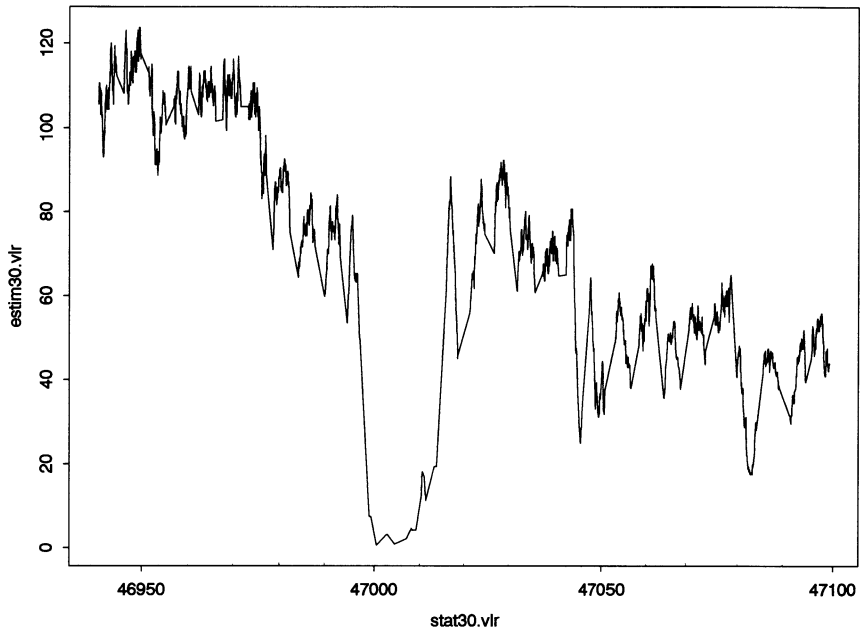


FIGURE 5 :

*Fonction d'intensité estimée (fenêtres locales) ( $\varepsilon = 1,5$ )*

## 5. Conclusion

Cette étude a permis de mettre en évidence le bon comportement des techniques de validation croisée pour le choix d'estimateurs non paramétriques de fonctions d'intensité. Bien sûr, ce bon comportement est lié au nombre important de données dont nous disposons. Par ailleurs, la nature même du problème physique qui nous intéresse dans cette étude a mis en relief la nécessité de développer des méthodes non paramétriques de détection de rupture. Dans le cadre d'estimation d'intensité nous ne connaissons aucun travail sur ce sujet, mais les méthodes récentes développées dans d'autres contextes (nous pensons en particulier au travail de Wu et Chu 1993, en régression) devraient pouvoir inspirer de futurs développements dans cette direction. Il est clair que les méthodes locales développées dans ce papier devraient constituer un outil intéressant pour l'étude de ruptures.

Par ailleurs cela nous a permis de faire le point sur les divers résultats mathématiques existant dans le domaine, et de mettre en évidence la nécessité d'apporter des réponses théoriques à certains problèmes importants non encore résolus. En particulier, dans un premier temps il serait intéressant de s'attacher à démontrer des résultats de type (3.5) pour des mesures de pertes locales dans le cadre du processus de Cox.

### Remerciements

Les discussions que nous avons eues avec Jean-Louis BRENGUIER et Anna PAWLOVSKA du Centre National de la Recherche Météorologique de Toulouse, nous ont permis de nous familiariser avec les données traitées dans ce papier. Leurs connaissances, tant sur les phénomènes physiques liés à ce jeu de données que sur les méthodes statistiques paramétriques habituellement utilisées dans ces problèmes, ont été précieuses pour nous tout au long de ce travail et nous tenons à les en remercier.

### Bibliographie

- BRENGUIER, J.L. (1993), Observations of cloud microstructure at the centimeter scale, *J. of Applied Meteorology*, **32**, 783-793.
- BRONIATOWSKI, M. (1993), Cross validation in kernel nonparametric density estimation : a survey, *Publi. de l'I.S.U.P.*, vol XXXVII, 3, 3-28.
- BROOKS, M.M. and MARRON, J.S. (1990), Asymptotic optimality of the least-squares cross-validation bandwidth for kernel estimates of intensity functions, *Stochastic Processes and their Applications*, **38**, 157-165, North-Holland.
- COX, D.R. and ISHAM, V. (1980), *Point processes*, London : Chapman and Hall.
- CRETOIS, E. (1994), Utilisation de la méthode des pas aléatoires en estimation dans les processus ponctuels, *Thèse de doctorat 3<sup>ème</sup> cycle*, Université de Rouen.
- DIGGLE, P. (1985), A kernel method for smoothing point process data, *Appl. Statist.*, **34**, 138-147.
- DIGGLE, P. and MARRON, J.S. (1988), Equivalence of smoothing parameters selectors in density and intensity estimation, *J. Amer. Statist. Assoc.*, **83**, 793-800.
- GREGOIRE, G. (1993), Least Squares Cross Validation for Counting Process Intensities, *Scand. J. Statist.*, **20**, 343-360.
- HALL, P. and SCHUCANY, W. (1989), A local Cross-Validation algorithm. *Statist. and Prob. Letters*, **8**, 109-117.
- MARRON, J.S. (1988), Automatic smoothing parameter selection : a survey, *Empirical Economics*, **13**, 187-208.
- MIELNICZUK, J., SARDA, P. and VIEU, P. (1989), Local data-driven bandwidth choice for density estimation, *J. Statist., Plann. and Inf.*, **23**, 53-69.

- RAMLAU-HANSEN, H. (1983), Smoothing counting process intensities by means of kernel functions, *Ann.Statist.*, **11**, 453-466.
- ROSENBLATT, M. (1956), Remarks on some nonparametric estimates of a density function, *Ann. Math. Statist.*, **27**, 832-835.
- VIEU, P. (1993), Bandwidth selection for kernel regression : a survey. In «*Computer Intensive Methods in Statistics*». Ed. W.HÄRDLE and L.SIMAR, Statistics and Computing, Physica Verlag, **1**, 134-149.
- WU, J.S. and CHU, C.K. (1993), Kernel type estimators of jump points and values of a regression function, *Ann. Statist.*, **21**, 3, 1545-1566. 17