



HAL
open science

Extension au cadre spatial de l'estimation non paramétrique par noyaux récurrents

Mohamed Yahaya

► **To cite this version:**

Mohamed Yahaya. Extension au cadre spatial de l'estimation non paramétrique par noyaux récurrents. Mathématiques générales [math.GM]. Université Charles de Gaulle - Lille III, 2016. Français. NNT : 2016LIL30066 . tel-01610867

HAL Id: tel-01610867

<https://theses.hal.science/tel-01610867>

Submitted on 5 Oct 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ LILLE III - CHARLES-DE-GAULLE, VILLENEUVE
D'ASCQ, FRANCE

École Doctorale SESAM

THÈSE DE DOCTORAT

Mention : Mathématiques et Applications

présentée par

Mohamed YAHAYA

Extension au cadre spatial de l'estimation non
paramétrique par noyaux récurrents

Soutenue publiquement le 15 décembre 2016
devant le jury composé de :

Mme. Sophie DABO	Professeure, Université de Lille 3	Directrice
Mr. Aboubacar AMIRI	Maître de conférences, Université de Lille 3	Co-encadrant
Mr. Mohamed RACHADI IBRAHIM	Maître de conférences, Université des Comores	Co-encadrant
Mme. Ghislaine GAYRAUD	Professeure, Université de Compiègne	Présidente
Mr. Salim BOUZEBDA	Professeur, Université de Compiègne	Rapporteur
Mr. Célestin KOKONENDJI	Professeur, Université de Franche-Comté	Rapporteur
Mr. Noureddine RHOMARI	Professeur, Université Mohamed Premier	Examineur
Mme. Gwladys TOULEMONDE	Maître de conférences, Université de Montpellier 2	Examinatrice

Remerciements

C'est un moment très fort en émotion, fermer ce chapitre de ma vie en remerciant toutes les personnes qui ont eu le courage de m'accompagner durant toutes ces longues années de recherche.

Je tiens à remercier vivement ma directrice de thèse Sophie Dabo-Niang et à lui exprimer ma profonde gratitude pour son soutien constant durant ces années de recherche. Je tiens à lui dire qu'elle a réussi par ses conseils et ses encouragements à développer le plaisir de la recherche.

Sincèrement, je ne sais pas comment exprimer ma reconnaissance envers mon co-directeur Aboubabcar Amiri pour le soutien moral et pédagogique qu'il m'a apporté depuis mon arrivée au Laboratoire LEM de l'Université Charles de Gaulle-Lille 3 jusqu'à ce jour. Ses idées et sa facilité à poser des questions simples et intéressantes m'ont fortement guidé et m'ont fait aimer davantage ce métier difficile qu'est la recherche. Je mesure la chance qui m'a été offerte de travailler avec Sophie Dabo-Niang et Aboubabcar Amiri et aussi d'avoir appris avec eux des mathématiques belles et variées. Je leur en suis infiniment reconnaissant pour leurs précieux conseils, leurs suggestions et aussi leurs encouragements.

Je remercie Mohamed Ibrahim Rachadi pour son soutien administratif, financier et aussi surtout pour l'honneur qu'il m'a offert ce jour.

Je remercie Monsieur Salim BOUZEBDA, Professeur à l'université de Compiègne et Monsieur Célestin KOKONENDJI, Professeur à l'université de Franche-Comté pour avoir accepté d'être rapporteurs de cette thèse.

Je remercie très chaleureusement Madame Ghislaine GAYRAUD, Professeure à l'université de Compiègne et Madame Gwladys TOULEMONDE, Maître de conférences à l'université de Montpellier 2 d'avoir accepté d'être examinatrices de cette thèse. Je leur suis très reconnaissant du temps qu'elles ont consacré à l'évaluation de cette thèse et de leur présence à ce jury.

Je tiens à exprimer ma profonde gratitude à Monsieur Noureddine RHOMARI, Professeur à l'université Mohamed Premier, qui m'a fait l'honneur de venir siéger en tant qu'examinateur parmi les membres du jury.

Je remercie toute l'équipe du Laboratoire LEM de l'université Charles de Gaulle-Lille 3 pour les moments précieux qu'on a passé ensemble. J'avoue que cette dernière m'a donnée le goût de la recherche et j'espère continuer à travailler davantage avec elle après cette thèse de doctorat. Une petite pensée aussi à tous les membres du laboratoire LEM : certains m'ont aidé lorsque j'avais des problèmes avec Latex ou R, certains m'ont donné des conseils lorsque je me posais des questions sur mes choix d'orientation et d'autres sont tout simplement devenus des ami(e)s... Merci également à mes collègues de bureau, de Lille 3, avec qui nous avons pu discuter Mathématiques mais aussi de nos différentes cultures et de nos petites histoires du quotidien. Il me faut également remercier l'ensemble des personnes présentes à l'UFR MIME pour leur gentillesse et le travail fourni pour permettre le bon déroulement de chaque année universitaire. Je ne pourrais jamais oublier les moments chaleureux que j'ai passé au sein de ce laboratoire.

Finalement, ce rêve n'aurait pas pu aussi se réaliser sans l'amour, la confiance, et le soutien des gens que je vais les remercier maintenant car, sans eux je ne serais pas où j'en suis aujourd'hui. Je remercie mes parents (paix à leurs âmes) qui ont fait grandir mon imagination dès mon enfance en m'initiant sur les calculs et aussi de m'avoir appris à ne jamais baisser les bras, à rester toujours debout face aux difficultés, à aimer la recherche et la science.

À ma femme Naïla Ahmed qui a vécu toutes ses longues années de doutes et d'angoisses de la recherche scientifique. Merci pour ton soutien et ta patience. Je remercie mes frères et soeurs de m'avoir donné tellement de chaleur et de joie de vivre. À mon cousin Adam Mohamed et sa femme Hinnie qui n'ont ménagé aucun effort pour me soutenir jusqu'à ce jour. Je remercie aussi Rastami Ahamadi, Fatoumia Oussen, Malizandro Salim, Salma Salim, Abdouroihamane Maamoune, Nourdine Bourhane, Abdouroihamane Mahamoud et mes collègues de travail de la Faculté des Sciences et Techniques de l'université des Comores (UDC) pour leur soutien moral et pédagogique.

Je remercie enfin l'université des Comores(UDC) et ses partenaires(AUF), le SCAC et Campus France pour leur soutien administratif et financier.

Je dédie ce travail à tous ceux que j'aime et en particulier à mon fils Nayad et à ma fille Ibtissam. Je leur souhaite une longue vie de bonheur, de prospérité et de plein succès.

Résumé

Résumé en français

Dans cette thèse, nous nous intéressons aux méthodes dites récursives qui permettent une mise à jour des estimations séquentielles de données spatiales ou spatio-temporelles et qui ne nécessitent pas un stockage permanent de toutes les données. Traiter et analyser des flux des données, *Data Stream*, de façon effective et efficace constitue un défi actif en statistique. En effet, dans beaucoup de domaines d'applications, des décisions doivent être prises à un temps donné à la réception d'une certaine quantité de données et mises à jour une fois de nouvelles données disponibles à une autre date. Nous proposons et étudions ainsi des estimateurs à noyau de la fonction de densité de probabilité et la fonction de régression de flux de données spatiales ou spatio-temporelles. Plus précisément, nous adaptons les estimateurs à noyau classiques de Parzen-Rosenblatt et Nadaraya-Watson. Pour cela, nous combinons la méthodologie sur les estimateurs récursifs de la densité et de la régression et celle d'une distribution de nature spatiale ou spatio-temporelle. Nous donnons des applications et des études numériques des estimateurs proposés. La spécificité des méthodes étudiées réside sur le fait que les estimations prennent en compte la structure de dépendance spatiale des données considérées, ce qui est loin d'être trivial. Cette thèse s'inscrit donc dans le contexte de la statistique spatiale non-paramétrique et ses applications. Elle y apporte trois contributions principales qui reposent sur l'étude des estimateurs non-paramétriques récursifs dans un cadre spatial/spatio-temporel et s'articule autour des l'estimation récur-

sive à noyau de la densité dans un cadre spatial, l'estimation récursive à noyau de la densité dans un cadre spatio-temporel, et l'estimation récursive à noyau de la régression dans un cadre spatial.

Mots-clefs

Statistique spatiale; Flux de données; Données dépendantes; Processus faiblement mélangeant; Estimation non paramétrique; Estimateur à noyau; Convergence en moyenne quadratique; Convergence presque sûre.

Extension to spatial setting of kernel recursive estimation

Abstract

In this thesis, we are interested in recursive methods that allow to update sequentially estimates in a context of spatial or spatial-temporal data and that do not need a permanent storage of all data. Process and analyze *Data Stream*, effectively and efficiently is an active challenge in statistics. In fact, in many areas, decisions should be taken at a given time at the reception of a certain amount of data and updated once new data are available at another date. We propose and study kernel estimators of the probability density function and the regression function of spatial or spatial-temporal data-stream. Specifically, we adapt the classical kernel estimators of Parzen-Rosenblatt and Nadaraya-Watson. For this, we combine the methodology of recursive estimators of density and regression and that of a distribution of spatial or spatio-temporal data. We provide applications and numerical studies of the proposed estimators.

The specificity of the methods studied resides in the fact that the estimates take

into account the spatial dependence structure of the relevant data, which is far from trivial. This thesis is therefore in the context of non-parametric spatial statistics and its applications. This work makes three major contributions. which are based on the study of non-parametric estimators in a recursive spatial/space-time and revolves around the recursive kernel density estimate in a spatial context, the recursive kernel density estimate in a space-time and recursive kernel regression estimate in space.

Keywords

Spatial statistics ; Data stream; Dependent data; Weakly dependent mixing processes ; Nonparametric estimation ; Kernel estimator ; Mean squared error convergence ; Almost sure convergence.

Travaux et Publications

Amiri, A., Dabo-Niang, S., Yahaya, M. (2015). Nonparametric recursive density estimation for spatial data. *C. R. Acad. Sci. Paris, Ser. I 354 (2016) 205–210.*

Participation à des conférences, séminaires et journées d'études

- 1) **Méthodes séquentielles et leurs applications aux données spatiales.** Regroupement régional du projet Horizons francophones¹, Sciences fondamentales - informatique et mathématiques de l'AUF, Campus numérique francophonie d'Antananarivo (Madagascar), du 24 au 26 octobre 2012.
- 2) **Estimation non paramétrique de la densité spatiale par une approche récursive.** Regroupement inter-régional du projet Horizons francophones, Sciences fondamentales - informatique et mathématiques de l'AUF, Ecole Nationale Supérieure Polytechnique de Yaoundé (Cameroun), du 1 au 5 juillet 2013²
- 3) **Estimation de la densité d'un flux de données spatio-temporelles.** 2ème regroupement inter-régional du projet Horizons francophones, Sciences fondamentales : informatique et mathématiques de l'AUF, Campus numérique francophone d'Antananarivo (Madagascar), du 13 au 17 octobre 2014

1. Le projet Horizons francophones a pour but de renforcer les capacités du corps professoral des pays du Sud dans le domaine de la formation à l'enseignement et à la recherche. Les doctorants sont prioritairement des enseignants en poste au sein de l'université d'origine ou ayant une activité d'enseignement ou de soutien à l'enseignement, qui seront amenés, grâce à leur formation, à renforcer l'offre de formation et de recherche de leur université d'origine.

2. Pendant une semaine, ce sont réunis 24 doctorants pour une formation intensive à la recherche : une présentation de l'avancement des travaux suivi d'un débat, des communications en sessions plénière et des ateliers sur la rédaction d'articles scientifiques animés par des professeurs des universités.

Table des matières

1	Introduction générale	1
1.1	Généralités sur la statistique spatiale	1
1.1.1	L'estimation spatiale	3
1.1.2	L'estimation non paramétrique spatiale	4
1.2	L'estimation récursive	7
2	Estimateurs récursifs de la densité spatiale	13
2.1	Méthodologie	14
2.1.1	Notations et définitions	14
2.1.2	Présentation de l'estimateur	16
2.2	Résultats asymptotiques	18
2.2.1	Hypothèses	18
2.2.2	Convergence en moyenne quadratique	20
2.2.3	Convergence presque sûre	23
2.3	Application	24
2.3.1	Cadre général	24
2.3.2	Estimation de la fonction de risque relatif	25
2.3.3	Propriétés asymptotiques	26
2.3.4	Sélection des paramètres	27
2.4	Preuves	28
2.4.1	Preuve du Lemme 2.4.1	29
2.4.2	Preuve du Théorème 2.2.1	29

TABLE DES MATIÈRES

2.4.3	Preuve du Corollaire 2.2.1	32
2.4.4	Preuve du Théorème 2.2.2	33
2.4.5	Preuve du Théorème 2.2.3	35
2.4.6	Preuve du Théorème 2.3.1	38
3	Estimateurs récurrents de la densité d'un flux de données spatio-temporelles	39
3.1	Flux de données spatio-temporelles	40
3.2	Présentation de l'estimateur	42
3.3	Hypothèses et résultats asymptotiques	43
3.4	Résultats numériques	48
3.4.1	Sélection des paramètres	49
3.4.2	Résultats numériques	51
3.5	Preuves	53
3.5.1	Preuve du Théorème 3.3.1	54
3.5.2	Preuve du Théorème 3.3.2	59
4	Estimation de la régression spatiale par une approche récursive	65
4.1	Cadre du travail	66
4.1.1	Notations et hypothèses	70
4.1.2	Preuve du Théorème 4.1.1	71
5	Conclusion générale et perspectives	75
5.1	Conclusion générale	75
5.2	Perspectives	76

Table des figures

2.1	Asymptotique intensive (grille irrégulière)	15
2.2	Asymptotique extensive (grille régulière)	15
3.1	Une réalisation du processus spatiotemporel $\mathbf{X}_t, t = 1, \dots, 6$ et $n = 10^4$	50

Liste des tableaux

3.1	MSE pour comparer les estimateurs récursif et non-récursif	52
3.2	Temps de calcul en secondes pour comparer les estimateurs récursif et non-récursif	52

Introduction générale

Sommaire

1.1 Généralités sur la statistique spatiale	1
1.1.1 L'estimation spatiale	3
1.1.2 L'estimation non paramétrique spatiale	4
1.2 L'estimation récursive	7

Cette thèse est une contribution à la statistique spatiale non-paramétrique et ses applications. Le thème principal repose sur l'étude des estimateurs non-paramétriques récursifs dans un cadre spatial/spatio-temporel et s'articule autour des points suivants :

- Estimation récursive à noyau de la densité dans un cadre spatial
- Estimation récursive à noyau de la densité dans un cadre spatio-temporel
- Estimation récursive à noyau de la régression dans un cadre spatial

1.1 Généralités sur la statistique spatiale

La statistique spatiale inclut toutes les techniques statistiques qui étudient des phénomènes observés sur des ensembles spatiaux. De tels phénomènes apparaissent dans une variété de domaines incluant entre autres la géologie, l'océanographie, l'économie, l'épidémiologie, les sciences de l'environnement. La statistique spatiale étudie ainsi des phénomènes observés sur un ensemble spatial $\mathcal{S} \in \mathbb{R}^N$,

$N \geq 2$ de sites. La statistique spatiale se différencie de la statistique classique par le fait que les observations sont dépendantes dans l'espace. On appelle données spatiales ou géoréférencées, les réalisations d'une famille de variables aléatoires indexées dans \mathbb{R}^N . Cette famille sera appelée champ aléatoire.

Les trois types de données géoréférencées.

On peut regrouper les données spatiales en trois types (voir [Cressie \(1991\)](#), [Chiles et Delfiner \(1999\)](#)) : les données géostatistiques, les données latticielles et les données ponctuelles.

- **Les données géostatistiques** : Les données géostatistiques sont définies sur un ensemble spatial \mathcal{S} qui est un sous-espace continu de \mathbb{R}^N . Ces données sont alors mesurées en un nombre fini de sites échantillonnés : le champ $Z = (Z_{\mathbf{s}}, \mathbf{s} \in \mathcal{S})$ est observé en n sites fixés $\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n\} \subset \mathcal{S}$ avec $Z_{\mathbf{s}} \in \mathbb{R}^d$.

D.G. Krige, ingénieur des mines en Afrique du Sud, est à l'origine de la géostatistique. Il s'est intéressé à la prédiction de la qualité du minerai dans un bloc minier, à partir d'échantillons observés.

On peut citer comme exemples de données géo-statistiques, des cumuls de précipitations dans une région $\mathcal{S} \subset \mathbb{R}^2$, les propriétés des sols, les caractéristiques géologiques d'une zone de prospection $\mathcal{S} \subset \mathbb{R}^2$,... Les questions types que l'on se pose dans l'étude de ces données concernent la caractérisation de la variabilité spatiale en fonction de la distance (géographique) entre deux lieux (sites), l'interpolation (cartographie ou Krigeage) de la variable étudiée entre les points mesurés.

- **Les données latticielles** : Dans ce cas le champ est observé sur un ensemble spatial $\mathcal{S} \subset \mathbb{R}^2$ discret et fixé. Les sites \mathbf{s} représentent en général des unités géographiques d'un réseau structuré à partir par exemple d'un graphe de voisinage. On peut citer par exemple des zones liées par un réseau de transport, des pixels d'une image IRM, etc.

On utilise dans ce types de données, des extensions au cadre spatial des modèles ARMA de séries temporelles.

Les questions types que l'on se pose dans l'étude de ces données concernent la

caractérisation de la variabilité spatiale, l'indépendance entre voisins, l'explication de la distribution des caractéristiques en fonction des distributions dans un voisinage donné.

• **Les données ponctuelles** : Les processus ponctuels surviennent lorsque les sites où ont lieu les observations sont aléatoires. Ce type de processus sont des extensions des processus ponctuels indexés dans \mathbb{R} au cadre \mathbb{R}^N . La question que l'on se pose dans l'étude de ces types de données est : la localisation des sites est-elle homogène ? Est-elle plutôt régulière ? Ou présente-t-elle des agrégats ? Par exemple, on retrouve ce type de données dans l'étude de la répartition spatiale d'une espèce.

Les analyses de ces trois catégories de données vont de la visualisation, de l'exploration de la dépendance spatiale à la modélisation.

Un des problèmes fondamentaux en statistique spatiale est la prédiction en un lieu non observé à l'aide d'observations effectuées dans d'autres sites à l'aide de la structure de dépendance entre les sites. Un problème non négligeable est la non-existence des notions de passé et futur et d'un ordre naturel en spatial contrairement au cadre temporel. Ainsi, les modèles classiques en séries temporelles ne peuvent pas être étendus facilement aux données spatiales. Pour en savoir plus sur les différences et les similarités existantes entre les séries spatiales et chronologiques, nous référons le lecteur aux travaux de [Tjøstheim \(1987\)](#).

1.1.1 L'estimation spatiale

Soient \mathcal{S} un ensemble spatial et $Z = \{Z_{\mathbf{s}}, \mathbf{s} \in \mathcal{S}\}$ un champ aléatoire spatial indexé dans \mathcal{S} où \mathbf{s} est considéré comme un site. Les premiers outils de la statistique spatiale furent paramétriques. Il existe une vaste littérature sur les modèles paramétriques en statistique spatiale allant de l'estimation de la variabilité spatiale et des tests associés à la prédiction (cf. [Guyon \(1995\)](#), [Cressie \(1991\)](#), [Chiles et Delfiner \(1999\)](#), [Stein \(1999\)](#), [Gaetan et Guyon \(2008\)](#) etc). Les méthodes non paramétriques spatiales sont plus récentes mais il existe une dynamique assez forte dans ce cadre. On peut citer entre autres [Tran \(1990\)](#), [Carbon *et al.* \(1997\)](#), [Biau et Cadre \(2004\)](#), [Menezes *et al.* \(2010\)](#), [Ould-Abdi et al. \(2010\)](#), [Dabo-Niang et Thiam \(2010\)](#), [García-Soidán et Menezes \(2012\)](#), [Dabo-Niang et](#)

1.1. GÉNÉRALITÉS SUR LA STATISTIQUE SPATIALE

Yao (2013), Ternynck (2014),...dans le cadre de l'estimation d'une densité, de la régression ou la prédiction spatiale non-paramétrique.

Dans le cadre des données géostatistiques, principalement considérées dans cette thèse, on considère \mathcal{S} comme un sous-ensemble fixé de \mathbb{R}^N avec $N > 1$. On dénote par $(\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n)$ des sites localisés dans $\mathcal{S} \subseteq \mathbb{R}^N$ et on suppose qu'on dispose d'observations $Z_{\mathbf{s}_1}, \dots, Z_{\mathbf{s}_n}$.

Dans les estimations paramétriques et non-paramétriques, lorsque qu'on s'intéresse aux propriétés asymptotiques des estimateurs, on étudie généralement leurs comportements lorsque la taille de l'échantillon converge vers l'infini. Deux structures sont généralement utilisées pour l'étude asymptotique (cf. Gaetan et Guyon (2008)) dans un cadre spatial :

- L'asymptotique extensive (*increasing domain asymptotics*) : c'est la situation où le nombre d'observations disponible croît avec le domaine d'observation \mathcal{S} .
- L'asymptotique intensive (*infill asymptotics*) : c'est la situation où les observations augmentent dans un domaine \mathcal{S} fixé et borné (on a de plus en plus de sites d'observation tandis que la région reste fixe).

Dans nos travaux, l'asymptotique extensive est considérée.

Dans la suite, nous commençons par donner quelques définitions, puis nous donnons quelques méthodes d'estimation non-paramétriques utiles dans le cadre de nos travaux.

1.1.2 L'estimation non paramétrique spatiale

Comme mentionné ci-dessus, la statistique spatiale constitue l'ensemble des techniques formelles qui étudient les entités utilisant leurs propriétés topologiques, géométriques ou géographiques. L'analyse spatiale se réfère ainsi à une variété de techniques, et se pose dans une variété de domaines. Des questions complexes se posent dans l'analyse spatiale, dont beaucoup ne sont ni clairement défini ni complètement résolu, mais constituent la base de la recherche actuelle. Le plus fondamental d'entre eux est le problème de la reconstruction d'un phénomène

sur son domaine à partir d'un ensemble de valeurs observées. Un intérêt principal en modélisation spatiale est ainsi la prédiction de valeurs non observées d'un champ aléatoire. Dans ce but, des techniques linéaires de prévision (krigeage) ont été développées. Les méthodes de krigeage standard conviennent bien lorsque la variable spatiale est gaussienne parce que le prédicteur linéaire est le prédicteur optimal dans le cas gaussien. Cependant, ces techniques donnent des résultats médiocres de prédiction lorsque la variable considérée est loin d'être gaussienne. D'autres approches telles que le krigeage lognormal ou le krigeage trans-gaussien qui transforment les données d'origine en des données réalisées comme étant gaussienne, ont été introduites. Néanmoins, avec ces dernières techniques, il pourrait y avoir des valeurs aberrantes comme un effet des données extrêmes qui ne peuvent pas être supprimées. Cela exige une considération particulière pour les inférences. Ainsi, une alternative du krigeage pourrait être un modèle non paramétrique.

Dans ce mémoire de thèse, nous nous intéressons ainsi à l'estimation non-paramétrique récursive de la densité et de la régression de flux de données spatiales ou spatio/temporelles. Les premières méthodes non-paramétriques spatiales concernent l'estimation de la densité et de la régression. La plupart de ces méthodes font appel à l'estimateur à noyau.

Estimation de la densité spatiale par la méthode à noyau.

Les premières contributions concernant l'estimation à noyau de [Parzen \(1962\)](#) et [Rosenblatt \(1985\)](#) de la densité spatiale sont dues à [Tran \(1990\)](#). Il se place dans le cas où on observe un champ aléatoire strictement stationnaire $(X_{\mathbf{i}})_{\mathbf{i} \in \mathbb{Z}^N} \in \mathbb{R}^d$, de densité marginale f et estime cette dernière en un point $x \in \mathbb{R}^d$ par

$$f_{\mathbf{n}}(x) = \frac{1}{\widehat{\mathbf{n}}h_{\mathbf{n}}^d} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K\left(\frac{x - X_{\mathbf{i}}}{h_{\mathbf{n}}}\right)$$

où $K(\cdot)$ est un noyau, $h_{\mathbf{n}}$ est la fenêtre, $\mathcal{I}_{\mathbf{n}} = \{\mathbf{i} \in \mathbb{Z}^N : 1 \leq i_k \leq n_k, k = 1, \dots, N\}$ et $\widehat{\mathbf{n}} = n_1 \times \dots \times n_N$. Il montre, sous certaines conditions de dépendance (mélange fort), la normalité asymptotique de $f_{\mathbf{n}}$. Ce dernier est étendu par [Tran et Yakowitz \(1993\)](#) qui proposent un estimateur de f basé sur les k -plus proches voisins et montre sa normalité asymptotique via la technique de décomposition par blocs utilisée. [Carbon et al. \(1997\)](#) ont étudié les convergences uniformes en

probabilité et presque sûre de $(f_{\mathbf{n}})$. [Hallin et al. \(2001\)](#) ont étudié l'estimateur à noyau de la densité d'un processus spatial linéaire et ont montré sa normalité asymptotique. [El Machkouri \(2014\)](#) a considéré la normalité asymptotique de $(f_{\mathbf{n}})$ via la méthode de Lindeberg. [Dabo-Niang et al. \(2014\)](#) proposent une nouvelle version de l'estimateur de [Tran \(1990\)](#) qui tient compte à la fois de la valeur des observations mais aussi de la position des sites où ont lieu les observations. Il existe d'autres contributions dans la littérature spatiale que nous ne pouvons citer de manière exhaustive ([Biau \(2003\)](#), [Fazekas et Chuprunov \(2006\)](#),...) À partir de l'estimateur de la densité proposé par [Tran \(1990\)](#), nous construisons un estimateur récursif de la densité.

Estimation par la méthode à noyau d'une régression spatiale.

Soit $Z_{\mathbf{s}} = (X_{\mathbf{s}}, Y_{\mathbf{s}})_{\mathbf{s} \in \mathcal{I}_{\mathbf{n}}}$ un échantillon d'un champ aléatoire $Z = \{(X_{\mathbf{s}}, Y_{\mathbf{s}}), \mathbf{s} \in \mathbb{R}^N\}$ strictement stationnaire obéissant au modèle de régression $r(x) = \mathbb{E}[Y_{\mathbf{s}} | X_{\mathbf{s}} = x]$, c'est-à-dire basé sur l'espérance conditionnelle de $Y_{\mathbf{s}}$ sachant $X_{\mathbf{s}}$ où $X_{\mathbf{s}} \in \mathbb{R}^d$. L'estimateur à noyau de la fonction de régression $r(\cdot)$ est donné par :

$$r_{\mathbf{n}}(x) = \frac{\frac{1}{\widehat{\mathbf{n}}b_{\mathbf{n}}^d} \sum_{\mathbf{s} \in \mathcal{I}_{\mathbf{n}}} Y_{\mathbf{s}} K\left(\frac{x - X_{\mathbf{s}}}{b_{\mathbf{n}}}\right)}{f_{\mathbf{n}}(x)}.$$

Cet estimateur est au cœur d'une dynamique de recherche depuis le début des années 2000 (cf. [Lu et Chen \(2002\)](#), [Hallin et al. \(2004\)](#), [Biau et Cadre \(2004\)](#), [Carbon et al. \(2007\)](#), [Dabo-Niang et Yao \(2007\)](#), [Li et Tran \(2009\)](#), [Gheriballah et al. \(2010\)](#), [Menezes et al. \(2010\)](#), [Karácsony et Filzmoser \(2010\)](#), [Robinson \(2011\)](#) dans le cas de données réelles). Dans le cadre du "fixed-design setting", le modèle de régression suivant a été considéré :

$$Y_{\mathbf{i}} = r(x_{\mathbf{i}}) + \epsilon_{\mathbf{i}}, \quad \mathbf{i} \in \mathbb{N}^N$$

où $(x_{\mathbf{i}}, Y_{\mathbf{i}})_{\mathbf{i} \in \mathbb{N}^N}$ est observé sur $\mathcal{I}_{\mathbf{n}}$, les $x_{\mathbf{i}}$ sont déterministes, et $\{\epsilon_{\mathbf{i}}\}_{\mathbf{i} \in \mathbb{N}^N}$ est un champ aléatoire stationnaire de moyenne nulle. Lorsque $x_{\mathbf{i}} = \frac{\mathbf{i}}{n}$ avec $n_1 = \dots = n_N = n$, [El Machkouri \(2007\)](#) a proposé un estimateur basé sur l'estimateur de

Nadaraya-Watson et défini comme suit

$$\hat{r}(x) = \frac{\sum_{i \in \mathcal{I}_n} Y_i K\left(\frac{x-i/n}{h}\right)}{\sum_{i \in \mathcal{I}_n} K\left(\frac{x-i/n}{h}\right)}, \quad x \in [0, 1]$$

où K est la fonction noyau et h est la fenêtre. Il a montré la convergence uniforme en probabilité de cet estimateur. La normalité asymptotique a été étudié par [El Machkouri \(2010\)](#). L'estimateur local linéaire de la régression a été étendu récemment dans un cadre spatial; [Francisco-Fernandez et Opsomer \(2005\)](#) ont développé un estimateur local linéaire à noyau de la régression dans un cadre de design aléatoire spatial. Ils ont prouvé la convergence faible de leur estimateur. [Wang et Wang \(2009\)](#) ont proposé un estimateur local linéaire de la régression dans le cadre de données réelles spatio-temporelles. Ils établissent la normalité asymptotique de leur estimateur via la technique de décomposition par blocs de Bernstein. D'autres formes de régressions basées sur d'autres caractéristiques conditionnelles (quantile et mode) qui utilisent la méthode par noyau sont proposées dans la littérature (cf. [Hallin et al. \(2009\)](#), [Dabo-Niang et Thiam \(2010\)](#), [Ould-Abdi et al. \(2010\)](#), [Dabo-Niang et al. \(2014\)](#),...).

1.2 L'estimation récursive

Bien que des estimateurs non paramétriques tels que l'estimateur de la densité de [Parzen \(1962\)](#) et [Rosenblatt \(1985\)](#) et l'estimateur de la régression de [Nadaraya \(1964\)](#) et [Watson \(2014\)](#) peuvent être utilisés pour estimer les fonctions de densité et de régression, ces estimateurs présentent cependant un inconvénient lié à leur complexité de calcul dans un contexte de grande dimension. En effet, si la densité ou la régression est estimée par un estimateur non récursif, ce dernier doit être recalculé complètement à chaque fois qu'une nouvelle donnée apparaît. Cette succession d'opérations de plus en plus longues peut devenir coûteuse en temps de calcul et en espace de stockage des données. Les méthodes récursives procurent alors un grand avantage sur la gestion de ces données volumineuses. En effet avec ces méthodes, la mise à jour des estimations s'effectuent alors séquentiellement et ne nécessite pas un stockage permanent de toutes les données.

1.2. L'ESTIMATION RÉCURSIVE

Des approches récursives ont été introduites et étudiées dans le cas non spatial. On parle alors d'estimation en temps réel. Dans ce cadre, l'approche récursive est très intuitive du fait de la nature séquentielle de l'écoulement du temps. Cependant, dans le cas spatial l'absence de définitions universelles de la notion de passé et de futur ne facilite pas l'interprétation des méthodes séquentielle dans ce contexte.

Dans ce mémoire de thèse, nous nous intéressons aux versions récursives de l'estimateur de type [Parzen \(1962\)](#) et [Rosenblatt \(1985\)](#), dans un cadre spatial. Une première contribution a été proposée dans un cadre non spatial par [Wagner et Wolverton \(1969\)](#) sous la forme :

$$f_n^{\text{WW}}(x) := \frac{1}{n} \sum_{i=1}^n \frac{1}{h_i^d} K\left(\frac{x - X_i}{h_i}\right), \forall x \in \mathbb{R}^d.$$

Ces auteurs ont montré que cet estimateur possède les mêmes propriétés asymptotiques que celui de [Parzen \(1962\)](#) et [Rosenblatt \(1985\)](#). Il a également l'avantage de pouvoir s'écrire sous une forme autoregressive permettant une mise à jour de l'estimateur avec n .

De nombreuses variantes récursives ont également été proposées et étudiées depuis. En particulier, [Deheuvels \(1974\)](#) s'est intéressé à la famille générale suivante :

$$f_n^H(x) := \left[\sum_{i=1}^n h_i H(h_i) \right]^{-1} \sum_{i=1}^n H(h_i) K\left(\frac{x - X_i}{h_i}\right).$$

Il a étudié ses propriétés asymptotiques pour des données i.i.d et uni-dimensionnelles dans une classe de fonctions $H(\cdot)$. Plus récemment, [Amiri \(2010\)](#) introduit la famille d'estimateurs récursifs à noyau indexée par un paramètre ℓ , définie par :

$$f_n^\ell(x) := \frac{1}{\sum_{i=1}^n h_i^{d(1-\ell)}} \sum_{i=1}^n \frac{1}{h_i^{d\ell}} K\left(\frac{x - X_i}{h_i}\right), \quad x \in \mathbb{R}^d \quad (\ell \in [0, 1]) \quad (1.1)$$

qui correspond pour $d = 1$ au cas $H(u) = u^{-\ell}$ dans $f_n^H(x)$. Cette famille d'estimateurs généralisent les estimateurs récursives de la densité étudiés auparavant par [Wagner et Wolverton \(1969\)](#), pour $\ell = 0$, [Davies et Wegman \(1979\)](#) pour $\ell = 1/2$ et [Deheuvels \(1974\)](#) pour $\ell = 1$. La famille (1.1) possède également l'avantage de s'exprimer sous une forme simple permettant la comparaison des

différents estimateurs qui la composent.

Pour une bibliographie sur l'estimation de la densité par une approche récursive, nous reprenons ci-dessous celle détaillée dans [Amiri \(2010\)](#).

Dans le cas iid, [Davies \(1973\)](#), [Deheuvels \(1974\)](#), [Davies et Wegman \(1979\)](#) et [Wertz \(1985\)](#) étudient la famille $f_n^H(x)$ et les cas $\ell = 0$, $\ell = 1/2$ et $\ell = 1$. En particulier, en dimension $d = 1$, [Deheuvels \(1974\)](#) établit la convergence en moyenne quadratique de la famille $f_n^H(x)$ et donne des conditions nécessaires et suffisantes pour sa convergence presque sûre sous certaines conditions vérifiées par la fonction H . [Roussas et Tran \(1992\)](#), [Davies et Wegman \(1979\)](#) établissent les vitesses de convergence presque sûre exactes dans les cas $\ell = 0$, $\ell = 1/2$ et $\ell = 1$. [Isogai \(1994\)](#) établit sous certaines conditions, la normalité asymptotique pour $\ell = 1$ dans le cas iid.

Dans le cas dépendant, la convergence en moyenne quadratique et la normalité asymptotique pour $\ell = 1/2$ et $\ell = 1$ sont établis par [Masry \(1986\)](#), pour des processus stationnaires fortement mélangeants. La vitesse de convergence presque sûre ponctuelle pour $\ell = 1/2$ et $\ell = 1$ est étudiée par [Takahata \(1980\)](#), [Masry et Györfi \(1987\)](#), d'abord sous des conditions de ρ -mélangeance, ensuite pour $\ell = 1$, par [Masry \(1987\)](#), pour des processus fortement mélangeants. Pour les valeurs fixées $\ell = 1/2$ et $\ell = 1$, la normalité asymptotique pour des processus temporels négativement associés a été traitée par [Liang et Beak \(2004\)](#), tandis que la convergence en moyenne quadratique et la normalité asymptotique ont été étudiées par [Masry \(1986\)](#) sous une condition de forte mélangeance. Les approches utilisées dans tous ces travaux, notamment pour la convergence en moyenne quadratique et la normalité asymptotique, ne se généralisent pas aisément, en dimension supérieure pour des valeurs plus petites de ℓ , alors qu'en particulier, le cas $\ell = 0$ est intéressant du fait de la faible variance de l'estimateur. [Amiri \(2010\)](#) a ainsi généralisé les travaux précédents par l'introduction de la famille (1.1). Il a étudié le biais, variance et l'erreur quadratique moyenne (EQM) asymptotiques, ainsi que le comportement presque sûr de la famille (1.1) pour des observations iid. Ensuite, ces résultats sont généralisés à des processus fortement mélangeants. [Mezhoud et al. \(2014\)](#) ont étendu les résultats obtenus par [Amiri \(2010\)](#) au cas des processus temporels faiblement dépendants.

S'agissant de l'estimation de la fonction de régression, une vaste littérature traite le problème dans le cas non spatial par une approche récursive. Les premiers estimateurs récursifs considérés dans la littérature sont de la forme :

$$\frac{\sum_{i=1}^n Y_i K_{h_i}^*(x, X_i)}{\sum_{i=1}^n K_{h_i}^*(x, X_i)},$$

et les plus populaires d'entre eux sont proposés et par [Devroye et Wagner \(1980\)](#) et par [Ahmed et Lin \(1976\)](#). Ils sont définis respectivement par :

$$r_n^{DW}(x) := \frac{\sum_{i=1}^n \frac{Y_i}{h_i^d} K\left(\frac{x - X_i}{h_i}\right)}{\sum_{i=1}^n \frac{1}{h_i^d} K\left(\frac{x - X_i}{h_i}\right)}$$

et

$$r_n^{AL}(x) := \frac{\sum_{i=1}^n Y_i K\left(\frac{x - X_i}{h_i}\right)}{\sum_{i=1}^n K\left(\frac{x - X_i}{h_i}\right)}.$$

Ces deux estimateurs sont largement étudiés dans la littérature. Dans le cas iid on peut citer les travaux de [Ahmed et Lin \(1976\)](#), [Devroye et Wagner \(1980\)](#).

Pour des processus φ -mélangeants, [Wang et Liang \(2004\)](#) a étudié la convergence presque sûre uniforme de versions tronquées de r_n^{DW} et r_n^{AL} . Dans le cas α -mélangeant [Roussas \(1990\)](#) a établi la convergence uniforme presque sûre de r_n^{DW} , tandis que la normalité est démontrée par [Roussas et Tran \(1992\)](#). Ces résultats sont ensuite généralisés par [Amiri \(2010\)](#) par l'introduction de la famille d'estimateur de la régression :

$$r_n^\ell(x) := \frac{\varphi_n^\ell(x)}{f_n^\ell(x)} \text{ avec } \varphi_n^\ell(x) := \frac{1}{\sum_{i=1}^n h_i^{(1-\ell)d}} \sum_{i=1}^n \frac{Y_i}{h_i^{d\ell}} K\left(\frac{x - X_i}{h_i}\right), (\ell \in [0, 1]).$$

La convergence en moyenne quadratique et presque sûre ainsi que la normalité asymétrique de $r_n^\ell(x)$ ont été établit dans [Amiri \(2012, 2013\)](#) dans le cas iid et pour des processus α -mélangeants.

Contribution du Chapitre 2

Le chapitre 2 présente une estimation de la fonction de densité de probabilité d'une variable aléatoire spatiale multivariée. L'approche que nous utilisons est

un lissage par noyaux récurrents qui généralise la méthode de Parzen (1962). Le problème abordé dans ce chapitre a été largement étudié dans la littérature dans différents contextes. Des approches récurrentes ont été introduites et étudiées dans le cas non spatial. Nous renvoyons le lecteur aux travaux pionniers de Wagner et Wolverton (1969), Deheuvels (1974), Davies et Wegman (1979) sur l'estimation récurrente non paramétrique de la densité de processus temporels. De récentes contributions se sont intéressées aux propriétés asymptotiques de la densité de données dépendantes par une approche récurrente. Amiri (2010) a généralisé les travaux précédents par l'introduction d'une famille générale d'estimateurs récurrents, utilisant différentes valeurs d'un paramètre $\ell \in [0, 1]$ associé à chaque estimateur. Pour les valeurs fixées $\ell = 1/2$ et $\ell = 1$, la normalité asymptotique pour des processus temporels négativement associés a été traitée par Liang et Beak (2004), tandis que la convergence en moyenne quadratique et la normalité asymptotique ont été étudiées par Masry (1986) sous une condition de forte mélangeance. Mezhoud *et al.* (2014) ont étendu les résultats obtenus par Amiri (2010) au cas des processus temporels faiblement dépendants.

Dans ce chapitre, nous étendons les travaux de Amiri (2010) au cas de données spatiales dépendantes et présentons une version récurrente de l'estimateur à noyau de la densité spatiale classique étudié auparavant par Tran (1990), Carbon *et al.* (1997). Nous donnons des résultats asymptotiques d'un tel estimateur ainsi qu'une application à l'estimation d'une fonction de risque.

Contribution du Chapitre 3

Traiter et analyser des flux des données, *Data Stream*, de façon effective et efficace constitue un défi actif en statistique. En effet, dans beaucoup de domaines d'applications, des décisions doivent être prises à un temps donné à la réception d'une certaine quantité de données et mises à jour une fois de nouvelles données disponibles à une autre date. Dans ce chapitre, nous nous intéressons à l'estimation non-paramétrique de la densité de probabilité d'un "Data Stream" de nature spatio-temporelle. Nous proposons et étudions ainsi un estimateur à noyau de la fonction de densité de probabilité d'un flux de données spatio-temporelles. Plus précisément, nous adaptons l'estimateur à noyau classique de Parzen-Rosenblatt. Pour cela, nous proposons de combiner la méthodologie sur l'estimateur récurrent

de la densité et celle d'une distribution de nature spatiale ([Tran \(1990\)](#)). Nous donnons une étude de simulations.

Contribution du Chapitre 4

La statistique spatiale permet de répondre à l'un des besoins fondamentaux qui sont étudiés dans de nombreux domaines à savoir la prévision spatiale. Il arrive que des données soient enregistrées dans des sites spécifiques et il est crucial parmi ces sites, d'étudier l'influence qui s'y exerce entre deux sites. Dans ce chapitre, nous nous intéressons à l'estimation de la fonction de régression spatiale par le biais des méthodes non paramétriques. Nous étendons les travaux effectués dans le chapitre 2 au cadre de la régression. Nous présentons une version récursive de l'estimateur à noyau de la régression spatiale classique étudié auparavant par [Biau \(2003\)](#). Nous donnons des résultats asymptotiques d'un tel estimateur.

Organisation de la thèse

Les chapitres 2, 3 et 4 donnent les contributions de la thèse énoncées ci-dessus. Le document se termine par une conclusion et quelques perspectives de recherche.

Chapitre 2

Estimateurs récursifs de la densité spatiale

Sommaire

2.1	Méthodologie	14
2.1.1	Notations et définitions	14
2.1.2	Présentation de l'estimateur	16
2.2	Résultats asymptotiques	18
2.2.1	Hypothèses	18
2.2.2	Convergence en moyenne quadratique	20
2.2.3	Convergence presque sûre	23
2.3	Application	24
2.3.1	Cadre général	24
2.3.2	Estimation de la fonction de risque relatif	25
2.3.3	Propriétés asymptotiques	26
2.3.4	Sélection des paramètres	27
2.4	Preuves	28
2.4.1	Preuve du Lemme 2.4.1	29
2.4.2	Preuve du Théorème 2.2.1	29
2.4.3	Preuve du Corollaire 2.2.1	32
2.4.4	Preuve du Théorème 2.2.2	33

2.1. MÉTHODOLOGIE

2.4.5 Preuve du Théorème 2.2.3	35
2.4.6 Preuve du Théorème 2.3.1	38

Dans ce chapitre, nous nous intéressons à l'estimation de la fonction de densité de probabilité d'une variable aléatoire spatiale multivariée. L'approche que nous utilisons est un lissage par noyaux récursifs qui généralise la méthode de [Parzen \(1962\)](#). Le problème abordé dans ce chapitre a été largement étudié dans la littérature dans différents contextes. La méthode et les résultats obtenus ici généralisent les travaux existants dans le cadre d'une estimation non récursive de la densité spatiale.

Bien que des estimateurs non paramétriques tels que l'estimateur de la densité de [Parzen \(1962\)](#) peuvent être utilisés pour estimer la fonction de densité, ces estimateurs présentent cependant un inconvénient lié à leur complexité de calcul dans le contexte de grande dimension.

Dans ce chapitre, nous proposons d'étendre les travaux de [Amiri \(2010\)](#) au cas des données spatiales dépendantes. Ainsi, nous présentons une version récursive de l'estimateur à noyau de la densité spatiale classique étudié auparavant par [Tran \(1990\)](#), [Carbon et al. \(1997\)](#). Nous donnons des résultats asymptotiques d'un tel estimateur. Ce dernier sera ensuite adapté au cas des données spatio-temporelles dans le chapitre suivant.

Dans la suite de cette thèse, on considère un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$. Soient d, N et n des entiers naturels non nuls. Nous introduisons quelques notations et définitions qui seront utilisées dans la suite de ce document.

2.1 Méthodologie

2.1.1 Notations et définitions

Dans la définition suivante nous rappelons la notion de processus spatial.

Définition 2.1.1 *Soit $\mathcal{S} \subset \mathbb{R}^N$ un ensemble spatial (appelé aussi ensemble de sites). On appelle champ aléatoire (ou processus spatial) sur \mathcal{S} toute famille de variables aléatoires indexée par \mathcal{S} et on note alors : $X = \{X_{\mathbf{s}}, \mathbf{s} \in \mathcal{S}\}$.*

Dans la suite, nous considérons un ensemble fini de n sites, $\mathcal{I}_n \subset \mathbb{N}^N$, muni de la distance uniforme définie par :

$$\|\mathbf{i} - \mathbf{j}\| := \max_{1 \leq k \leq N} |i_k - j_k|,$$

pour tous sites $\mathbf{i} = (i_1, \dots, i_N)$, $\mathbf{j} = (j_1, \dots, j_N) \in \mathbb{N}^N$ et nous notons :

$$\hat{\mathbf{i}} = \text{Card} \{ \mathbf{j} \in \mathcal{I}_n : \|\mathbf{j}\| \leq \|\mathbf{i}\| \}.$$

Contrairement au cas temporel, en statistique spatiale, il existe deux façons de caractériser le comportement asymptotique d'un domaine spatial (Cressie (1991), p. 480). Les sites peuvent augmenter dans un domaine borné qui devient de plus en plus dense (Figure 2.1) ou bien le domaine lui-même peut être non borné et croître suivant les directions de l'espace (Figure 2.2). Dans le premier cas, le nombre de sites d'observation augmente de manière dense au sein d'une région bornée, on parle alors d'asymptotique intensive, tandis que dans le second, l'augmentation de la taille d'échantillon est due à l'expansion de la région (non bornée), on parle alors d'asymptotique extensive.

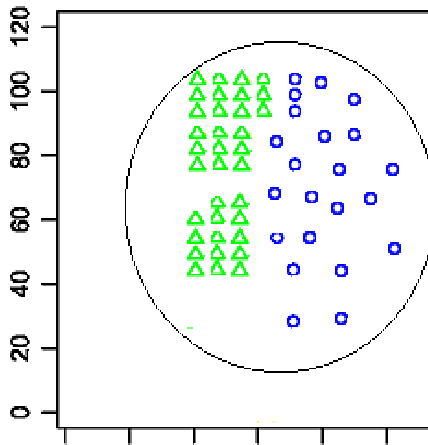


FIGURE 2.1 – Asymptotique intensive (grille irrégulière)

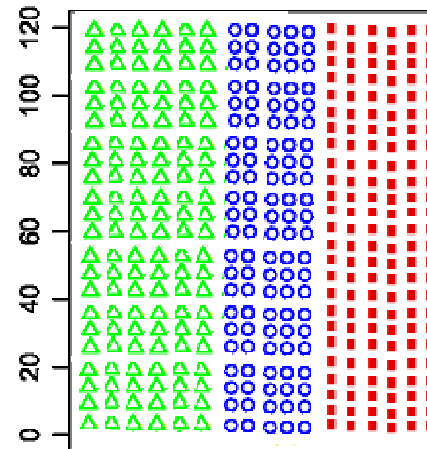


FIGURE 2.2 – Asymptotique extensive (grille régulière)

2.1. MÉTHODOLOGIE

Comme dans [Robison \(1926\)](#), dans la suite de cette thèse, dire que le N -uplet $\mathbf{n} = (n_1, \dots, n_N)$ tend vers l'infini signifie que $\min_{j=1, \dots, N} n_j$ tend vers l'infini. Observons que dans le cas d'une asymptotique extensive, le plus grand site (au sens de l'ordre lexicographique) du domaine tend vers l'infini, alors que pour une asymptotique intensive ce site est fixe. Enfin, lorsque le domaine est régulier et de forme rectangulaire, alors on a : $n = \hat{\mathbf{n}}$ et $n \rightarrow \infty \Leftrightarrow \mathbf{n} \rightarrow \infty$.

2.1.2 Présentation de l'estimateur

Considérons un processus spatial $X = \{X_{\mathbf{s}}, \mathbf{s} \in \mathcal{I}_n\}$, strictement stationnaire mesurable sur $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans \mathbb{R}^d , tel que pour tout $\mathbf{s} \in \mathcal{I}_n$, $X_{\mathbf{s}}$ admet la même distribution qu'un vecteur aléatoire X , de densité de probabilité inconnue $p(\cdot)$ relativement à la mesure de Lebesgue λ_d sur \mathbb{R}^d .

Nous cherchons donc à estimer $p(\cdot)$. Pour cela, nous introduisons une version spatiale de la famille générale d'estimateurs récursifs à noyau introduite dans le cadre temporel par [Amiri \(2010\)](#), définie par :

$$p_n^\ell(x) := \frac{1}{S_{n,\ell}} \sum_{\mathbf{s} \in \mathcal{I}_n} \frac{1}{h_{\mathbf{s}}^{d\ell}} K\left(\frac{X_{\mathbf{s}} - x}{h_{\mathbf{s}}}\right), \quad x \in \mathbb{R}^d \quad (\ell \in [0, 1]), \quad (2.1)$$

où $K(\cdot)$ est une fonction de noyau, $\{h_{\mathbf{s}}\}_{\mathbf{s} \in \mathcal{I}_n}$ une suite de fenêtres de lissage et

$$S_{n,\ell} := \sum_{\mathbf{s} \in \mathcal{I}_n} h_{\mathbf{s}}^{d(1-\ell)} \quad (2.2)$$

est une constante de normalisation. Ici, contrairement au cas d'une estimation par noyau classique, la fenêtre de lissage est adaptative. Le choix de celle-ci tient compte la localisation de l'observation $X_{\mathbf{s}}$. Comme cela a été démontré dans le cas temporel, cette prise en compte du comportement local du processus étudié a pour avantage de réduire la variance de l'estimateur, mais augmente le biais d'estimation. Le paramètre ℓ joue alors un rôle de régulateur de l'efficacité de l'estimateur selon des critères de biais et de variance minimaux. Soulignons toutefois que ces considérations sont d'ordre pratiques (lorsqu'on se place dans le cas d'une taille d'échantillon finie). En effet, nous montrerons que tout comme l'estimateur à noyau classique étudié par [Tran \(1990\)](#) et [Carbon *et al.* \(1997\)](#), l'estimateur défini en (2.1) est asymptotiquement sans biais et que sa variance tend également

vers zéro. Dans le cas temporel, [Amiri \(2010\)](#) a montré que le paramètre ℓ n'a pas d'influence majeure dans la qualité de l'estimation et que seuls les cas $\ell = 0$ et $\ell = 1$ présentent un intérêt pratique (la variance de l'estimateur est minimale pour $\ell = 0$ et le biais ainsi que l'erreur quadratique sont minimaux pour $\ell = 1$). Ainsi, afin de simplifier la rédaction, nous nous intéressons uniquement à ces cas. Dans cette thèse, on se place dans le cas où l'ensemble spatial \mathcal{I}_n est régulier, de forme rectangulaire et évolutif dans le sens où il croît de manière récursive (avec le temps par exemple). Dans cette configuration, nous utilisons l'ordre lexicographique et numérotions les sites pris dans cette grille rectangulaire. Par exemple, dans le cas bivarié ($N = 2$), la numérotation se fait de la façon suivante : l'observation du site $\mathbf{i} = (i_1, i_2)$ peut être indexée par $t = n_2(i_1 - 1) + i_2$ de la grille, voir [Robinson \(2011\)](#). Dans ce cas, la nouvelle observation du site est $(n_1 + 1, 1)$ et donc indexée par $n + 1$. Soit donc \mathbf{s}_{n+1} un nouveau site d'observation qui vient s'ajouter à \mathcal{I}_n , à savoir :

$$\mathcal{I}_{n+1} = \mathcal{I}_n \cup \{\mathbf{s}_{n+1}\}.$$

Si l'on note $X_{\mathbf{s}_{n+1}}$ la nouvelle observation du processus collectée au site \mathbf{s}_{n+1} , alors le choix d'une suite de fenêtres adaptatives permet une mise à jour de l'estimateur (2.1) via la formule autoregressive suivante :

$$p_{n+1}^\ell(x) = \frac{S_{n,\ell}}{S_{n+1,\ell}} p_n^\ell(x) + K_{n+1}^\ell (X_{\mathbf{s}_{n+1}} - x) \quad (2.3)$$

avec

$$K_i^\ell(\cdot) := \frac{1}{h_{\mathbf{s}_i}^{d_\ell} S_{i,\ell}} K\left(\frac{\cdot}{h_{\mathbf{s}_i}}\right).$$

Notons que dans (2.3), $p_{n+1}^\ell(x)$ est l'estimateur de la densité basé sur les observations collectées dans le domaine $\mathcal{I}_n \cup \{\mathbf{s}_{n+1}\}$. La formule (2.3) indique alors comment corriger l'estimation de la densité lorsqu'une donnée (une information) nouvelle est enregistrée sans recourir à tout "l'historique" du processus.

Remarquons également que la définition de l'estimateur (2.3) peut être motivée par un algorithme stochastique de type [Robbins et Monro \(1951\)](#). Pour mieux argumenter notre propos, rappelons d'abord cet algorithme.

Soit $M(\cdot)$ une fonction telle que l'équation $M(z) = 0$ admette une solution unique notée $z = \vartheta$. On se propose d'estimer ϑ . [Robbins et Monro \(1951\)](#) ont proposé

2.2. RÉSULTATS ASYMPTOTIQUES

une solution itérative $\widehat{\vartheta}_n$ définie à chaque “étape $n + 1$ ” par :

$$\widehat{\vartheta}_{n+1} = \widehat{\vartheta}_n - a_n \eta_n,$$

où a_n est une suite de “pas” asymptotiquement faible, telle que $\sum_{i=1}^{\infty} a_n^2 < \infty$ et η_n une estimation adéquate de la fonction $M(\cdot)$, qui dépend uniquement de $\widehat{\vartheta}_n$ et de l’information collectée à l’étape $n + 1$. En posant

$$\gamma_n = 1 - \frac{S_{n,\ell}}{S_{n+1,\ell}},$$

on déduit aisément de (2.3) que $p_n^\ell(x)$ est une combinaison convexe de la forme :

$$p_{n+1}^\ell(x) = (1 - \gamma_n)p_n^\ell(x) + \gamma_n \tilde{p}_{n+1}(x), \quad (2.4)$$

avec

$$\tilde{p}_{n+1}(x) = \frac{1}{h_{\widehat{\mathbf{s}}_{n+1}}^d} K\left(\frac{X_{s_{n+1}} - x}{h_{\widehat{\mathbf{s}}_{n+1}}}\right).$$

On peut remarquer que dans la représentation (2.4), la quantité $\tilde{p}_{n+1}(x)$ est un estimateur de la densité $p(x)$ basée uniquement sur l’information collectée au site \mathbf{s}_{n+1} . Ainsi, l’estimateur (2.3) peut donc se déduire d’un algorithme stochastique de type [Robbins et Monro \(1951\)](#) en choisissant $M(z) = z - p(x)$ avec un pas égale à γ_n .

Dans la littérature des algorithmes stochastiques, plusieurs choix de γ_n sont possibles. Plusieurs auteurs préconisent de prendre $\gamma_n = 1/(n + 1)$. Ce dernier correspond à un choix de paramètre $\ell = 1$, dans notre cas. Observons qu’avec ce choix, si $h_{\widehat{\mathbf{s}}} = h_n$ pour tout $\mathbf{s} \in \mathcal{I}_n$ (choix d’une fenêtre unique pour l’ensemble de la zone \mathcal{I}_n) on retrouve l’estimateur classique étudié par [Tran \(1990\)](#) et [Carbon et al. \(1997\)](#).

2.2 Résultats asymptotiques

2.2.1 Hypothèses

Afin d’établir nos résultats asymptotiques, nous avons besoin des hypothèses suivantes.

Hypothèse H. 2.2.1 \mathcal{I}_n est régulier et rectangulaire

$$\mathcal{I}_n = \{\mathbf{i} = (i_1, \dots, i_N), 1 \leq i_j \leq n_j, j = 1, \dots, N\}.$$

Hypothèse H. 2.2.2

(i) $h_n \downarrow 0$ et $nh_n^d \rightarrow \infty$ lorsque $n \rightarrow \infty$;

(ii) Pour tout $r \in]-\infty, 2 + d]$, on a :

$$B_{n,r} := \frac{1}{n} \sum_{i=1}^n \left(\frac{h_i}{h_n} \right)^r \rightarrow B_r > 0 \text{ lorsque } n \rightarrow \infty.$$

La première condition de l'hypothèse H.2.2.2 est classique en estimation non paramétrique, tandis que la seconde est très utile dans nos calculs et est également propre à la récursivité. Elle est souvent utilisée dans la littérature dans le cas temporel, notamment par Masry (1986), Davies et Wegman (1979), Yamato (1972), Mugisha et Samanta (1981) et Isogai (1994). Des exemples de choix de fenêtre vérifiant H.2.2.2 dans le cas temporel sont discutés dans Amiri (2010) et sont facilement transposables au cas spatial, notamment les fenêtres $h_n = A_n n^{-q}$ avec $A_n \downarrow A > 0$, $0 < q < 1$ ou $h_n = A_n \left(\frac{\ln n}{n} \right)^q$, avec $A_n \downarrow A > 0$ et $0 < q < \frac{1}{2+d}$.

Hypothèse H. 2.2.3 $K : \mathbb{R}^d \rightarrow \mathbb{R}$ est une densité bornée symétrique telle que :

$$\int_{\mathbb{R}^d} uK(u)du = 0 \text{ et } \int_{\mathbb{R}^d} \|u\|^2 K(u)du < \infty$$

avec $\|\cdot\|$ une norme de \mathbb{R}^d .

Hypothèse H. 2.2.4 La densité p est bornée et de classe C^2 sur \mathbb{R}^d .

Les hypothèses H.2.2.3 et H.2.2.4 sont usuelles en estimation non paramétrique et sont vérifiées en particulier par le noyau d'Epanechnikov et le noyau Gaussien, etc.

Hypothèse H. 2.2.5 Pour tous $\mathbf{i}, \mathbf{j} \in \mathcal{I}_n$ tels que $\mathbf{i} \neq \mathbf{j}$, le vecteur aléatoire $(X_{\mathbf{i}}, X_{\mathbf{j}})$ admet une densité jointe $f_{\mathbf{i}, \mathbf{j}}(\cdot, \cdot)$ telle que :

$$\sup_{\mathbf{i} \neq \mathbf{j}} \|g_{\mathbf{i}, \mathbf{j}}\|_{\infty} < \infty, \text{ où } g_{\mathbf{i}, \mathbf{j}}(\cdot, \cdot) = f_{\mathbf{i}, \mathbf{j}}(\cdot, \cdot) - p(\cdot) \times p(\cdot).$$

2.2. RÉSULTATS ASYMPTOTIQUES

Hypothèse H. 2.2.6

Le champ $(X_{\mathbf{i}})_{\mathbf{i} \in \mathcal{I}_n}$ est α -mélangeant : il existe une fonction

$$\phi : \mathbb{R}^+ \rightarrow \mathbb{R}^+ \text{ avec } \phi(z) \downarrow 0 \text{ quand } z \rightarrow \infty,$$

telle que pour tous $E, E' \subset \mathbb{R}^N$ dont les cardinaux $\text{Card}(E)$ et $\text{Card}(E')$ sont finis on ait :

$$\begin{aligned} \alpha(\sigma(E), \sigma(E')) &:= \sup_{A \in \sigma(E), B \in \sigma(E')} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \\ &\leq \psi(\text{Card}(E), \text{Card}(E')) \phi(\text{dist}(E, E')), \end{aligned}$$

où

$$\sigma(E) = \{X_{\mathbf{i}}, \mathbf{i} \in E\} \text{ et } \sigma(E') = \{X_{\mathbf{i}}, \mathbf{i} \in E'\},$$

$\text{dist}(E, E')$ est la distance entre E et E' et $\psi(\cdot)$ est une fonction positive symétrique croissante sur l'ensemble des deux variables.

On suppose de plus que les fonctions $\phi(\cdot)$ et $\psi(\cdot)$ vérifient les inégalités suivantes :

$$\phi(z) \leq Cz^{-\theta}, \quad \theta > 0 \text{ et } \psi(z, t) \leq C \min(z, t)$$

La mesure de dépendance utilisée dans H.2.2.6 est usuelle en statistique spatiale. On suppose généralement dans ce cadre que la fonction $\phi(\cdot)$ tend vers 0 avec une vitesse polynomiale ou bien

$$\phi(z) \leq C \exp(-\mu z), \text{ pour tous } C, \mu > 0,$$

c'est à dire que $\phi(\cdot)$ tend vers 0 avec une vitesse exponentielle. Nos résultats peuvent être prouvés sous ces conditions.

2.2.2 Convergence en moyenne quadratique

Le Théorème 2.2.1 suivant établit la convergence de l'erreur quadratique moyenne de l'estimateur (2.1). Soit p_n (resp. p_n^*) l'estimateur donné par (2.1) avec $\ell = 1$ (resp. $\ell = 0$).

Théorème 2.2.1 *Soit x tel que $p(x) > 0$. Sous les hypothèses H.2.2.1 - H.2.2.4, on a :*

$$h_n^{-2} [\mathbb{E}(p_n(x)) - p(x)] \rightarrow B_2 \int_{\mathbb{R}^d} \frac{K(u)}{2} \sum_{j_1, j_2=1}^d u_{j_1} u_{j_2} \frac{\partial^2 p(x)}{\partial x_{j_1} \partial x_{j_2}} du, \quad (2.5)$$

lorsque $n \rightarrow \infty$, avec $u = (u_1, \dots, u_d)$.

De plus si H.2.2.3-H.2.2.6 sont vraies et $\theta > 1 + N$, on a :

$$nh_n^d \nabla(p_n(x)) \rightarrow p(x) B_{-d} \int_{\mathbb{R}^d} K^2(u) du \quad (2.6)$$

Le Théorème 2.2.1 est établi sous l'hypothèse H.2.2.1 mais il peut être généralisé moyennant quelques conditions supplémentaires au cas d'un domaine irrégulier notamment lorsque le domaine d'observation satisfait l'hypothèse H.2.2.7 ci-dessous.

Hypothèse H. 2.2.7 *La région d'observation \mathcal{D} est un ensemble irrégulier infini et dénombrable, tel que tous deux points de \mathcal{D} sont distants d'au moins $\delta > 1$. Dans ce cas (comme dans Conley (1999)), nous considérons une suite finie de régions fermées et convexes $\{R_u\}$, de surfaces croissantes lorsque $u \rightarrow +\infty$. On suppose alors que le domaine d'échantillonnage est obtenu par l'intersection de chacune des régions convexes et $\mathcal{D} : \mathcal{I}_n = \mathcal{D} \subset R_u$.*

Ici, nous ne spécifions aucun ordre, ni toute autre hypothèse sur la configuration et l'augmentation de la taille de l'échantillon, exceptée l'hypothèse de croissance uniforme des surfaces de R_u dans au moins deux directions non opposées pour augmenter la taille de l'échantillon n (lorsque $u \rightarrow +\infty$) voir Conley (1999). Ainsi, une nouvelle observation correspond à n'importe quel nouveau site \mathbf{s}_{n+1} de \mathcal{D} qui se trouve en dehors de \mathcal{I}_n .

Le Théorème 2.2.1 est une extension au cas spatial des résultats obtenus dans le cas de données temporelles par Amiri (2010) sous la condition α -mélangeante. Nous avons également le corollaire suivant :

Corollaire 2.2.1 *Sous les hypothèses du Théorème 2.2.1, p_n^* vérifie (2.5), lorsque B_2 est remplacé par $B_d^{-1} B_{d+2}$.*

De plus si $\theta > \frac{Nd+2}{2}$, alors p_n^ vérifie (2.6) lorsque B_{-d} est remplacé par B_d^{-1} .*

2.2. RÉSULTATS ASYMPTOTIQUES

Notons en particulier que si $h_n = n^{-\frac{1}{4+d}}$, on obtient le comportement asymptotique de l'erreur quadratique moyenne :

$$n^{\frac{4}{4+d}} \text{EQM}(p_n^*(x)) \rightarrow \left[\int_{\mathbb{R}^d} K(u) \sum_{j_1, j_2=1}^d u_{j_1} u_{j_2} \frac{\partial^2 p(x)}{\partial x_{j_1} \partial x_{j_2}} du \right]^2 + \frac{4}{4+d} p(x) \int_{\mathbb{R}^d} K^2(u) du$$

lorsque $n \rightarrow \infty$ pour tout x tel que $p(x) > 0$. En effet, si $h_n = n^{-\frac{1}{4+d}}$, alors pour tout $r \leq d+2$, on a $B_r = \frac{d+4}{d+4-r}$. Il vient que

$$B_d^{-1} = \frac{d+4}{4} \text{ et } B_{d+2} = \frac{d+4}{2}.$$

Donc $B_d^{-1} B_{d+2} = 2$. Puisque l'erreur quadratique moyenne se décompose comme la somme du biais au carré et de la variance, on obtient donc le résultat précédent. Une représentation similaire peut être établie pour l'estimateur (2.1), pour tout $\ell \in [0, 1]$.

Enfin, soulignons que les résultats du Théorème 2.2.1 sont établis en supposant que la suite des fenêtres $\{h_{\mathfrak{s}}\}_{\mathfrak{s}}$ est déterministe. Plutôt que de faire dépendre les fenêtres uniquement des localisations des observations, on peut choisir des fenêtres qui dépendent des observations elles-mêmes. On obtient ainsi une suite de fenêtres $\{h_{\mathfrak{s}}\}_{\mathfrak{s}}$ aléatoires. C'est en particulier le cas lorsque les sites sont répartis aléatoirement dans l'espace et que l'on s'intéresse à la distribution de ceux-ci. Les observations sont alors les coordonnées des sites eux-mêmes. Dans le cas non spatial, ce choix de fenêtres est suggéré par Abramson (1982) et étudié également par Hall et Marron (1988). Motivé par les références ci-dessus, nous obtenons le Théorème suivant qui établit la représentation biais-variance de notre estimateur avec un choix de fenêtres aléatoires.

Théorème 2.2.2 *Sous les hypothèses du Théorème 2.2.1 si $h_{\mathfrak{s}} = \lambda_{\mathfrak{s}} f(X_{\mathfrak{s}})^{-1/2}$, avec $\{\lambda_{\mathfrak{s}}\}_{\mathfrak{s}}$ une suite vérifiant l'hypothèse H.2.2.2, alors*

$$\mathbb{E}(p_n(x)) \rightarrow p(x)$$

lorsque $n \rightarrow \infty$ et

$$n \lambda_n^d \mathbb{V}(p_n(x)) \rightarrow f(x)^{d/2} p(x) B_{-d} \int_{\mathbb{R}^d} K^2(u) du$$

lorsque $n \rightarrow \infty$ pour tout x tel que $p(x) > 0$.

Moyennant quelques hypothèses supplémentaires sur $p(\cdot)$ et une complexité accrue au niveau de la preuve, on peut obtenir un biais asymptotique exact de l'estimateur (2.1) et établir la même vitesse de convergence qu'au Théorème 2.2.1.

2.2.3 Convergence presque sûre

En plus des hypothèses précédentes nous aurons besoin de l'hypothèse suivante :

Hypothèse H. 2.2.8 Pour tout $\epsilon > 0$ (et pour $n = n_1 \times \dots \times n_N$), soit

$$u_n = \prod_{i=1}^N (\log n_i) (\log \log n_i)^{1+\epsilon}.$$

On suppose que

$$n \log n^{\frac{\theta-2N}{4N-\theta}} h_n^{\frac{d\theta}{\theta-4N}} u_n^{\frac{2N}{4N-\theta}} \rightarrow \infty, \text{ où } \theta > \max \left\{ 4N, \frac{Nd+2}{2} \right\}.$$

La définition de u_n de l' hypothèse H.2.2.8 implique que

$$\sum_{(n_1, \dots, n_N) \in \mathbb{N}^N} \frac{1}{n u_n} < \infty$$

(rappelons que $n = n_1 \times \dots \times n_N$).

Le théorème suivant établit la convergence presque sûre de p_n dans le cas $\ell = 1$.

Théorème 2.2.3 Sous les hypothèses H.2.2.1, H.2.2.2- H.2.2.6, on a :

$$|p_n(x) - p(x)| = O \left(\sqrt{\frac{\log n}{n h_n^d}} \right), \quad p.s \tag{2.7}$$

lorsque $n \rightarrow \infty$ pour tout x tel que $p(x) > 0$.

L'hypothèse H.2.2.1 permet de se placer dans le cadre d'une grille régulière où le type d'asymptotique choisi est extensif. Les hypothèses H.2.2.2- H.2.2.6 sont cruciales pour la convergence presque sûre de l'estimateur p_n . Les Théorèmes 2.2.1 et 2.2.3 sont également obtenus dans le cas $\ell = 0$ (à partir de l'équation (2.1)) par Amiri et al. (2015) et peuvent aisément se généraliser au cas général $\ell \in [0, 1]$.

2.3 Application

2.3.1 Cadre général

En épidémiologie, l'utilisation de l'analyse spatiale pour l'aide à la décision est de plus en plus fréquente, notamment pour évaluer des risques sanitaires en association avec des pollutions environnementales. Pour cela, on combine des méthodes épidémiologiques, statistiques et les systèmes d'information géographiques. Par exemple, en santé environnementale, on s'intéresse souvent à l'analyse de la répartition spatiale d'indicateurs de santé. Cette démarche comporte plusieurs objectifs. D'une part, on cherche à décrire les variations des indicateurs de santé et modéliser leur structure. D'autre part, on a besoin de mettre en évidence des associations entre ces variations et celles des expositions à des facteurs de risque environnementaux.

Les variations spatiales des indicateurs de santé et des facteurs d'expositions environnementaux sont souvent étudiés en épidémiologie dans un cadre descriptif. Il existe différents types d'analyses faisant intervenir une approche géographique : la représentation cartographique du risque de maladie, la détection d'agrégats spatiaux autour d'un point source ou encore l'évaluation de l'association entre risque et exposition environnementale en fonction de facteurs de risque connus. Pour effectuer ces analyses, des informations spatialisées sont utilisées. Ces informations peuvent être des données sanitaires. Par exemple, on peut utiliser les cas, pour une épidémie donnée, géolocalisés à partir des adresses des malades, ainsi que des informations contextuelles.

Pour comprendre l'évolution spatiale du risque lié à une épidémie dans une zone géographique donnée, il est nécessaire d'étudier la distribution spatiale des cas

d'épidémie, en comparaison avec la distribution spatiale de la population. On examine alors la fonction de risque relative liée à l'épidémie. Celle-ci représente le rapport entre la densité des cas d'épidémie et la densité de contrôle (densité de la distribution spatiale de la population exposée à l'épidémie).

2.3.2 Estimation de la fonction de risque relatif

Supposons que nous avons les coordonnées géographiques (exemple latitude, longitude) d'un nombre n_1 de cas d'épidémie que nous notons par $(X_i)_{i \in \mathcal{U}_{n_1}}$ et d'un ensemble de n_2 cas de contrôle représentant des populations à risque $(Y_j)_{j \in \mathcal{U}_{n_2}}$ où

$$\mathcal{U}_j = \{s_1, \dots, s_{n_j}\}, \quad j = 1, 2.$$

On suppose que les X_i (resp. les Y_j) sont identiquement distribués avec une densité commune p (resp. q), telles que p et q sont définies dans un compact $\mathcal{D} \subset \mathbb{R}^2$ (une quelconque région géographique de \mathbb{R}^2). La fonction de risque relatif est définie par (Bithell (1990)) :

$$r(\mathbf{z}) = \frac{p(\mathbf{z})}{q(\mathbf{z})} \text{ pour tout } \mathbf{z} \in \mathcal{D}.$$

Plutôt que d'étudier directement la fonction $r(\cdot)$, on peut lui appliquer une transformation de type Box-Cox. Par exemple, une transformation logarithmique, conduit au risque logarithmique défini par (Kelsall et Diggle (1995)) :

$$\rho(\mathbf{z}) = \log r(\mathbf{z}) = \log \left[\frac{p(\mathbf{z})}{q(\mathbf{z})} \right].$$

Davies et Hazelton (2010) utilisent des estimateurs à noyau pour l'estimation de la fonction de risque relatif liée à *la cirrhose biliaire primitive*¹ dans le nord-ouest de l'Angleterre.

En s'inspirant de Davies et Hazelton (2010), nous pouvons proposer un estimateur de la fonction de risque relatif basé sur des noyaux récursifs. Nous utilisons

1. c'est une maladie hépatique touchant le plus souvent la femme.

2.3. APPLICATION

pour cela l'estimateur (2.1) pour estimer p et q . Soient \hat{p} et \hat{q} les estimateurs associés, à savoir :

$$\hat{p}(\mathbf{z}) = \frac{1}{S_p} \sum_{\mathbf{i} \in \mathcal{U}_{n_1}} \frac{1}{h_{\mathbf{i}}^\ell} K\left(\frac{X_{\mathbf{i}} - \mathbf{z}}{h_{\mathbf{i}}}\right) \text{ et } \hat{q}(\mathbf{z}) = \frac{1}{S_q} \sum_{\mathbf{j} \in \mathcal{U}_{n_2}} \frac{1}{\kappa_{\mathbf{j}}^\ell} K\left(\frac{Y_{\mathbf{j}} - \mathbf{z}}{\kappa_{\mathbf{j}}}\right), \quad (2.8)$$

avec

$$S_p := \sum_{\mathbf{i} \in \mathcal{U}_{n_1}} h_{\mathbf{i}}^{2-\ell} \text{ et } S_q := \sum_{\mathbf{j} \in \mathcal{U}_{n_2}} \kappa_{\mathbf{j}}^{2-\ell} \quad \text{où } \ell \in [0, 2].$$

En remplaçant p et q par leur estimations, on obtient un estimateur de la (log) fonction de risque relatif :

$$\hat{\rho}_\ell(\mathbf{z}) = \log \hat{r}(\mathbf{z}) = \log \left[\frac{\hat{p}(\mathbf{z})}{\hat{q}(\mathbf{z})} \right], \quad \ell \in [0, 2]. \quad (2.9)$$

2.3.3 Propriétés asymptotiques

On pose

$$n = \min(n_1, n_2)$$

et pour toute fonction à deux variables ψ , on définit :

$$\Lambda_\psi(\mathbf{z}) := \frac{1}{\psi(\mathbf{z})} \int_{\mathbb{R}^2} \frac{K(u)}{2} \left[u_1^2 \frac{\partial^2 \psi}{\partial x_1^2}(\mathbf{z}) + 2u_1 u_2 \frac{\partial^2 \psi}{\partial x_1 \partial x_2}(\mathbf{z}) + u_2^2 \frac{\partial^2 \psi}{\partial x_2^2}(\mathbf{z}) \right] du,$$

avec $u = (u_1, u_2)$.

Hypothèse H. 2.3.1

- (i) \mathcal{D} vérifie l'hypotèse H.2.2.1
- (ii) Les fenêtres h et κ vérifient l'hypotèse H.2.2.2
- (iii) Le noyau K vérifie l'hypotèse H.2.2.3
- (iv) Les densités p et q sont bornées et de classe C^2 sur \mathbb{R}^2

Hypothèse H. 2.3.2

- (i) Pour tous $\mathbf{i}, \mathbf{j} \in \mathcal{U}_{n_1}$ tels que $\mathbf{i} \neq \mathbf{j}$, le vecteur aléatoire $(X_{\mathbf{i}}, X_{\mathbf{j}})$ admet une densité jointe $p_{\mathbf{i}, \mathbf{j}}^*(\cdot, \cdot)$ telle que :

$$\sup_{\mathbf{i} \neq \mathbf{j}} \|\mathcal{P}_{\mathbf{i}, \mathbf{j}}\|_\infty < \infty, \text{ où } \mathcal{P}_{\mathbf{i}, \mathbf{j}}(\cdot, \cdot) = p_{\mathbf{i}, \mathbf{j}}^* - p(\cdot) \times p(\cdot).$$

- (ii) Pour tous $\mathbf{i}, \mathbf{j} \in \mathcal{U}_{n_2}$ tels que $\mathbf{i} \neq \mathbf{j}$, le vecteur aléatoire $(Y_{\mathbf{i}}, Y_{\mathbf{j}})$ admet une densité jointe $q_{\mathbf{i}, \mathbf{j}}^*(\cdot, \cdot)$ telle que :

$$\sup_{\mathbf{i} \neq \mathbf{j}} \|\mathcal{Q}_{\mathbf{i}, \mathbf{j}}\|_{\infty} < \infty, \text{ où } \mathcal{Q}_{\mathbf{i}, \mathbf{j}}(\cdot, \cdot) = q_{\mathbf{i}, \mathbf{j}}^* - q(\cdot) \times q(\cdot).$$

- (iii) Pour tous $\mathbf{i} \in \mathcal{U}_{n_1}$ et $\mathbf{j} \in \mathcal{U}_{n_2}$, les variables $X_{\mathbf{i}}$ et $Y_{\mathbf{j}}$ sont indépendantes.
 (iv) Les champs $(X_{\mathbf{i}})_{\mathbf{i} \in \mathcal{U}_{n_1}}$ et $(Y_{\mathbf{i}})_{\mathbf{i} \in \mathcal{U}_{n_2}}$ vérifient l'hypothèse H 2.2.6.

Théorème 2.3.1 *Soit $\mathbf{z} \in \mathcal{D}$ tel que $p(\mathbf{z}) > 0$ et $q(\mathbf{z}) > 0$. Sous l'hypothèse H.2.3.1, si $\hat{\rho}$ est l'estimateur donné par (2.9) avec $\ell = 2$, alors on a :*

$$\mathbb{E}(\hat{\rho}(\mathbf{z})) - \rho(\mathbf{z}) = \frac{B_{4-\ell}}{B_{2-\ell}} \{h_{n_1}^2 \Lambda_p(\mathbf{z}) - \kappa_{n_2}^2 \Lambda_q(\mathbf{z})\} + o_{\mathbb{P}}(1) \quad (2.10)$$

lorsque $n \rightarrow \infty$.

De plus si H.2.3.2 est vraie, alors pour $\theta > 3$, on a :

$$\mathbb{V}(\hat{\rho}(\mathbf{z})) = \frac{B_{2(1-\ell)}}{B_{2-\ell}^2} \int_{\mathbb{R}^2} K^2(u) du \left[\frac{1}{n_1 h_{n_1}^2 p(\mathbf{z})} + \frac{1}{n_2 \kappa_{n_2}^2 q(\mathbf{z})} \right] + o_{\mathbb{P}}(1) \quad (2.11)$$

lorsque $n \rightarrow \infty$.

2.3.4 Sélection des paramètres

En estimation non-paramétrique la sélection des fenêtres de lissage $h_{\mathbf{i}}$ et $\kappa_{\mathbf{j}}$ est un problème crucial qui affecte considérablement le degré de lissage de l'estimateur. De trop grandes valeurs de la fenêtre donnent un estimateur avec un grand biais et une variance très faible, alors qu'une trop petite fenêtre donne un estimateur avec un biais trop faible et une variance trop grande. En pratique on peut par exemple utiliser la méthode de sélection des fenêtres introduite par [Davies et Hazelton \(2010\)](#) dont voici la description.

Soit \tilde{p} (resp. \tilde{q}) un estimateur de p (resp. q) calculé avec une fenêtre auxiliaire \tilde{h} appelée *fenêtre pilote*. Afin de calculer la valeur de $\hat{p}(\mathbf{z})$, on utilise une suite de fenêtres adaptatives de la forme :

$$h_{\mathbf{i}} = \frac{h_0}{\gamma \tilde{p}(X_{\mathbf{i}})^{1/2}}, \quad \mathbf{i} \in \mathcal{U}_{n_1}, \quad (2.12)$$

2.4. PREUVES

où h_0 est une constante multiplicative appelée *fenêtre globale*, commune à tous les sites d'observation, et

$$\gamma = \exp \left\{ \sum_{\mathbf{i} \in \mathcal{U}_{n_1}} \log [\tilde{p}(X_{\mathbf{i}})^{-1/2}] + \sum_{\mathbf{j} \in \mathcal{U}_{n_2}} \log [\tilde{q}(X_{\mathbf{j}})^{-1/2}] \right\}^{\frac{1}{n_1 + n_2}}$$

est la moyenne géométrique groupée des séries $\tilde{p}(X_{\mathbf{i}})$, $\mathbf{i} \in \mathcal{U}_{n_1}$ et $\tilde{q}(X_{\mathbf{j}})$, $\mathbf{j} \in \mathcal{U}_{n_2}$. Le calcul de $\hat{q}(\mathbf{z})$ se fait en sélectionnant la suite des fenêtres $\kappa_{\mathbf{j}}$, $\mathbf{j} \in \mathcal{U}_{n_2}$ de la même manière. La même fenêtre globale h_0 est utilisée dans les deux procédures de sélection. Celle-ci est donnée par :

$$h_0 = \overline{IQR}/1.34 \left\{ \frac{10^4 \pi}{128 n^* \Gamma(5)} \int K^2(z) dz \right\}^{\frac{1}{3}},$$

où n^* est la moyenne géométrique de n_1 et n_2 et \overline{IQR} est la moyenne des écarts interquartile des $X_{\mathbf{i}}$ et des $Y_{\mathbf{j}}$.

Cette méthode de sélection des fenêtres est implémentée dans le package *sparr*.²

2.4 Preuves

Rappelons que pour $\ell = 1$, on a d'après (2.1) :

$$p_n(x) = \frac{1}{n} \sum_{\mathbf{i} \in \mathcal{I}_n} h_{\mathbf{i}}^{-d} K \left(\frac{X_{\mathbf{i}} - x}{h_{\mathbf{i}}} \right), \quad x \in \mathbb{R}^d. \quad (2.13)$$

Pour prouver les résultats de la Section 2.2, nous utiliserons souvent le Lemme 2.4.1.

Lemme 2.4.1 *Soit $(w_n)_{n \geq 1}$ une suite de nombres réels tendant vers w . Si l'hypothèse H. 2.2.2 est vérifiée pour tout $r \in]-\infty, d + 2]$, alors :*

$$\frac{1}{n} \sum_{i=1}^n \left(\frac{h_i}{h_n} \right)^r w_i \rightarrow B_r w$$

lorsque $n \rightarrow \infty$.

2. <https://cran.r-project.org/web/packages/sparr/sparr.pdf>

2.4.1 Preuve du Lemme 2.4.1

Le Lemme 2.4.1 est une conséquence directe du lemme de Toeplitz (voir Stout (1974)). Sa preuve est une adaptation directe du calcul de Amiri (2010), p. 28. ■

2.4.2 Preuve du Théorème 2.2.1

Pour contrôler le terme de variance de $p_n(x)$, nous le décomposons en deux termes principaux. On a :

$$\mathbb{V}(p_n(x)) = \mathbb{V} \left[\frac{1}{n} \sum_{\mathbf{i} \in \mathcal{I}_n} h_{\mathbf{i}}^{-d} K \left(\frac{X_{\mathbf{i}} - x}{h_{\mathbf{i}}} \right) \right] = I_1 + I_2, \quad (2.14)$$

où

$$I_1 = \frac{1}{n^2} \sum_{\mathbf{i} \in \mathcal{I}_n} h_{\mathbf{i}}^{-2d} \mathbb{V}(Z_{\mathbf{i}}) \text{ et } I_2 = \frac{1}{n^2} \sum_{\substack{\mathbf{i}, \mathbf{j} \in \mathcal{I}_n \\ \mathbf{i} \neq \mathbf{j}}} h_{\mathbf{i}}^{-d} h_{\mathbf{j}}^{-d} \text{Cov}(Z_{\mathbf{i}}, Z_{\mathbf{j}}),$$

avec

$$Z_{\mathbf{i}} = K \left(\frac{X_{\mathbf{i}} - x}{h_{\mathbf{i}}} \right).$$

D'une part, sous l'hypothèse H.2.2.2, on a :

$$\begin{aligned} nh_n^d I_1 &= \frac{h_n^d}{n} \sum_{\mathbf{i} \in \mathcal{I}_n} h_{\mathbf{i}}^{-2d} [\mathbb{E}(Z_{\mathbf{i}}^2) - \mathbb{E}^2(Z_{\mathbf{i}})] \\ &= \frac{h_n^d}{n} \sum_{i=1}^n h_i^{-2d} \left\{ \int_{\mathbb{R}^d} K^2 \left(\frac{t-x}{h_i} \right) p(t) dt - \left[\int_{\mathbb{R}^d} K \left(\frac{t-x}{h_i} \right) p(t) dt \right]^2 \right\}. \end{aligned}$$

En faisant le changement de variable $u = \frac{t-x}{h_i}$, on obtient :

$$\begin{aligned} nh_n^d I_1 &= \frac{h_n^d}{n} \sum_{i=1}^n h_i^{-d} \left\{ \int_{\mathbb{R}^d} K^2(u) p(x + uh_i) du - h_i^d \left[\int_{\mathbb{R}^d} K(u) p(x + uh_i) du \right]^2 \right\} \\ &:= \frac{1}{n} \sum_{i=1}^n \left(\frac{h_i}{h_n} \right)^{-d} w_i, \end{aligned} \quad (2.15)$$

avec

$$w_i := \int_{\mathbb{R}^d} K^2(u) p(x + uh_i) du - h_i^d \left[\int_{\mathbb{R}^d} K(u) p(x + uh_i) du \right]^2 \rightarrow p(x) \int_{\mathbb{R}^d} K^2(u) du,$$

2.4. PREUVES

lorsque $h_i \rightarrow 0$, grâce à la continuité de p (hypothèse H.2.2.4). On a donc par application du Lemme 2.4.1 et grâce à l'hypothèse H.2.2.2 :

$$nh_n^d I_1 \rightarrow p(x) B_{-d} \int_{\mathbb{R}^d} K^2(u) du \quad (2.16)$$

lorsque $n \rightarrow \infty$.

D'autre part, on décompose le terme de covariance I_2 comme suit :

$$I_2 = J_1 + J_2, \quad (2.17)$$

où

$$J_1 = \frac{1}{n^2} \sum_{\substack{\mathbf{i}, \mathbf{j} \in \mathcal{I}_n \\ \|\mathbf{i}-\mathbf{j}\| \leq c_n}} h_{\mathbf{i}}^{-d} h_{\mathbf{j}}^{-d} \text{Cov}(Z_{\mathbf{i}}, Z_{\mathbf{j}}) \text{ et } J_2 = \frac{1}{n^2} \sum_{\substack{\mathbf{i}, \mathbf{j} \in \mathcal{I}_n \\ c_n < \|\mathbf{i}-\mathbf{j}\|}} h_{\mathbf{i}}^{-d} h_{\mathbf{j}}^{-d} \text{Cov}(Z_{\mathbf{i}}, Z_{\mathbf{j}}),$$

c_n étant une suite qui tend vers l'infini lorsque n tend vers l'infini.

Les changements de variables

$$u_1 = \frac{t_1 - x}{h_{\mathbf{i}}} \text{ et } u_2 = \frac{t_2 - x}{h_{\mathbf{j}}}$$

et la densité jointe définie dans l'hypothèse H.2.2.5, conduisent à :

$$\text{Cov}(Z_{\mathbf{i}}, Z_{\mathbf{j}}) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} h_{\mathbf{i}}^d h_{\mathbf{j}}^d K(u_1) K(u_2) g_{\mathbf{i}, \mathbf{j}}(x + h_{\mathbf{i}} u_1, x + h_{\mathbf{j}} u_2) du_1 du_2.$$

Par conséquent :

$$|\text{Cov}(Z_{\mathbf{i}}, Z_{\mathbf{j}})| \leq h_{\mathbf{i}}^d h_{\mathbf{j}}^d \|g_{\mathbf{i}, \mathbf{j}}\|_{\infty}.$$

Maintenant puisque \mathcal{I}_n est une grille régulière, un raisonnement similaire à celui de [Tran \(1990\)](#), donne

$$\text{Card} \{(\mathbf{u}, \mathbf{v}) \in \mathcal{I}_n^2 : \|\mathbf{u} - \mathbf{v}\| \leq c_n\} = O(n c_n^N).$$

Il vient alors que :

$$\begin{aligned} nh_n^d |J_1| &\leq \frac{h_n^d}{n} \sum_{\mathbf{i} \in \mathcal{I}_n} h_{\mathbf{i}}^{-d} \sum_{\mathbf{j} \in \mathcal{I}_n : \|\mathbf{i}-\mathbf{j}\| \leq c_n} h_{\mathbf{j}}^{-d} |\text{Cov}(Z_{\mathbf{i}}, Z_{\mathbf{j}})| \\ &\leq \frac{Ch_n^d}{n} \sum_{\mathbf{i} \in \mathcal{I}_n} \sum_{\mathbf{j} \in \mathcal{I}_n : \|\mathbf{i}-\mathbf{j}\| \leq c_n} \sup_{\mathbf{i} \neq \mathbf{j}} \|g_{\mathbf{i}, \mathbf{j}}\|_{\infty} \\ &\leq Ch_n^d c_n^N \sup_{\mathbf{i} \neq \mathbf{j}} \|g_{\mathbf{i}, \mathbf{j}}\|_{\infty} = O(h_n^d c_n^N). \end{aligned}$$

Concernant le terme J_2 , une adaptation de l'inégalité de Billingsley (voir [Blanke et Bosq \(2007\)](#), dans un cadre non spatial), permet d'obtenir l'inégalité suivante :

$$|\text{Cov}(Z_{\mathbf{i}}, Z_{\mathbf{j}})| \leq 4\phi(\|\mathbf{i} - \mathbf{j}\|) \|K\|_{\infty}^2.$$

Par conséquent grâce à la décroissance de la suite $\{h_s\}_s$ et l'hypothèse [H.2.2.6](#) :

$$\begin{aligned} nh_n^d |J_2| &\leq \frac{h_n^d}{n} \sum_{\mathbf{i} \in \mathcal{I}_n} h_{\mathbf{i}}^{-d} \sum_{\mathbf{j} \in \mathcal{I}_n: \|\mathbf{i}-\mathbf{j}\| > c_n} h_{\mathbf{j}}^{-d} |\text{Cov}(Z_{\mathbf{i}}, Z_{\mathbf{j}})| \\ &\leq \frac{4 \|K\|_{\infty}^2 h_n^d}{n} \sum_{\mathbf{i} \in \mathcal{I}_n} h_{\mathbf{i}}^{-d} \sum_{\mathbf{j} \in \mathcal{I}_n: \|\mathbf{i}-\mathbf{j}\| > c_n} h_{\mathbf{j}}^{-d} \phi(\|\mathbf{i} - \mathbf{j}\|) \\ &\leq \frac{4 \|K\|_{\infty}^2}{nh_n^d} \sum_{\mathbf{j} \in \mathcal{I}_n: t_{\mathbf{j}} > c_n} t_{\mathbf{j}}^{-\theta} \end{aligned}$$

avec $t_{\mathbf{j}} := \|\mathbf{i} - \mathbf{j}\|$ pour tout $\mathbf{i} \in \mathcal{I}_n$ fixé. Ainsi

$$nh_n^d |J_2| \leq \frac{4 \|K\|_{\infty}^2 h_n^{-d}}{(\theta - 1)} c_n^{1-\theta} = O(c_n^{1-\theta} h_n^{-d}).$$

En choisissant

$$c_n = \left[h_n^{-\frac{2d}{N+\theta-1}} \right], \text{ avec } \theta > 1 + N,$$

on a :

$$nh_n^d |I_2| \rightarrow 0 \text{ lorsque } n \rightarrow \infty, \quad (2.18)$$

Le résultat [\(3.4\)](#) est une conséquence de [\(2.16\)](#) et [\(2.18\)](#).

Concernant le terme de biais, on a

$$\begin{aligned} \mathbb{E}(p_n(x)) - p(x) &= \frac{1}{n} \sum_{\mathbf{i} \in \mathcal{I}_n} h_{\mathbf{i}}^{-d} \mathbb{E} \left[K \left(\frac{X_{\mathbf{i}} - x}{h_{\mathbf{i}}} \right) \right] - p(x) \\ &= \frac{1}{n} \sum_{\mathbf{i} \in \mathcal{I}_n} h_{\mathbf{i}}^{-d} \int_{\mathbb{R}^d} K \left(\frac{t - x}{h_{\mathbf{i}}} \right) p(t) dt - p(x) \end{aligned} \quad (2.19)$$

En faisant un changement de variable, $u = \frac{t - x}{h_{\mathbf{i}}}$, on obtient :

$$\begin{aligned} \mathbb{E}(p_n(x)) - p(x) &= \frac{1}{n} \sum_{\mathbf{i} \in \mathcal{I}_n} \int_{\mathbb{R}^d} K(u) p(x + h_{\mathbf{i}} u) du - p(x) \\ &= \frac{1}{n} \sum_{\mathbf{i} \in \mathcal{I}_n} \int_{\mathbb{R}^d} K(u) [p(x + h_{\mathbf{i}} u) - p(x)] du. \end{aligned} \quad (2.20)$$

2.4. PREUVES

Grâce à l'hypothèse H.2.2.3 et par application d'un développement de Taylor à l'ordre 2, il existe un réel τ , avec $0 < \tau < 1$ tel que :

$$h_n^{-2} [\mathbb{E}(p_n(x)) - p(x)] = \frac{1}{n} \sum_{i=1}^n \left(\frac{h_i}{h_n} \right)^2 \omega_i,$$

avec

$$\omega_i = \int_{\mathbb{R}^d} \frac{K(u)}{2} \sum_{j_1, j_2=1}^d u_{j_1} u_{j_2} \frac{\partial^2 p(x + \tau h_i u)}{\partial x_{j_1} \partial x_{j_2}} du \rightarrow \int_{\mathbb{R}^d} \frac{K(u)}{2} \sum_{j_1, j_2=1}^d u_{j_1} u_{j_2} \frac{\partial^2 p(x)}{\partial x_{j_1} \partial x_{j_2}} du$$

lorsque $h_{s_i} \rightarrow 0$. Grâce aux hypothèses H.2.2.3 et H.3.3.4, au théorème de la convergence dominée et au Lemme 2.4.1, on obtient (2.5). \blacksquare

2.4.3 Preuve du Corollaire 2.2.1

La preuve est identique à celle du Théorème 2.2.1. Le seul point de différence repose sur le contrôle du terme de covariance. Considérons donc la décomposition (2.17), qui cette fois s'écrit avec

$$J_1 = \frac{1}{B_{n,d}^2 n^2 h_n^{2d}} \sum_{\substack{\mathbf{i}, \mathbf{j} \in \mathcal{I}_n \\ \|\mathbf{i} - \mathbf{j}\| \leq c_n}} \text{Cov}(Z_{\mathbf{i}}, Z_{\mathbf{j}}) \text{ et } J_2 = \frac{1}{B_{n,d}^2 n^2 h_n^{2d}} \sum_{\substack{\mathbf{i}, \mathbf{j} \in \mathcal{I}_n \\ c_n < \|\mathbf{i} - \mathbf{j}\|}} \text{Cov}(Z_{\mathbf{i}}, Z_{\mathbf{j}}).$$

Écrivons :

$$|\text{Cov}(Z_{\mathbf{i}}, Z_{\mathbf{j}})| \leq h_{\mathbf{i}}^d h_{\mathbf{j}}^d \|g_{\mathbf{i}, \mathbf{j}}\|_{\infty} \leq \frac{h_{\mathbf{i}}^{2d} + h_{\mathbf{j}}^{2d}}{2} \sup_{\mathbf{i} \neq \mathbf{j}} \|g_{\mathbf{i}, \mathbf{j}}\|_{\infty}. \quad (2.21)$$

Ici le Lemme 2.4.1 n'est pas directement applicable du fait que la majoration de la covariance fait apparaître les termes $h_{\mathbf{i}}^{2d}$ et $h_{\mathbf{j}}^{2d}$. En effet, $2d > d+2$ lorsque $d > 3$.

Ainsi, si $d > 3$ alors on choisit un réel ζ tel que

$$\frac{N}{\theta - 1} < \zeta \leq \frac{2}{d}.$$

Puisque la suite des fenêtres $(h_n)_{n \geq 1}$ est décroissante, alors on obtient à partir de l'inegalité (2.21) :

$$|\text{Cov}(Z_{\mathbf{i}}, Z_{\mathbf{j}})| \leq \frac{h_{\mathbf{i}}^{d(\zeta+1)} h_1^{d(1-\zeta)} + h_{\mathbf{j}}^{d(\zeta+1)} h_1^{d(1-\zeta)}}{2} \sup_{\mathbf{i} \neq \mathbf{j}} \|g_{\mathbf{i}, \mathbf{j}}\|_{\infty}.$$

Ainsi :

$$\begin{aligned}
 nh_n^d |J_1| &\leq \frac{1}{B_{n,d}^2} \frac{1}{nh_n^d} \sum_{\mathbf{i} \in \mathcal{I}_n} \sum_{\mathbf{j} \in \mathcal{I}_n: \|\mathbf{i}-\mathbf{j}\| \leq c_n} |\text{Cov}(Z_{\mathbf{i}}, Z_{\mathbf{j}})| \\
 &\leq \frac{1}{B_{n,d}^2} \frac{1}{n} \sum_{i=1}^n \left(\frac{h_i}{h_n} \right)^{d(\zeta+1)} h_1^{d(1-\zeta)} h_n^{d\zeta} \sum_{\mathbf{j} \in \mathcal{I}_n: \|\mathbf{i}-\mathbf{j}\| \leq c_n} \sup_{\mathbf{i} \neq \mathbf{j}} \|g_{\mathbf{i},\mathbf{j}}\|_{\infty} \\
 &\leq \frac{B_{n,d}^{d(\zeta+1)}}{B_{n,d}^2} h_1^{d(1-\zeta)} h_n^{d\zeta} c_n^N \sup_{\mathbf{i} \neq \mathbf{j}} \|g_{\mathbf{i},\mathbf{j}}\|_{\infty} \\
 &= O(h_n^{d\zeta} c_n^N).
 \end{aligned}$$

Si $1 \leq d \leq 3$, on prend $\zeta = 1$ et on applique un raisonnement identique.

Concernant le terme J_2 , on a :

$$\begin{aligned}
 nh_n^d |J_2| &\leq \frac{1}{nB_{n,d}^2 h_n^d} \sum_{\mathbf{i} \in \mathcal{I}_n} \sum_{\mathbf{j} \in \mathcal{I}_n: \|\mathbf{i}-\mathbf{j}\| > c_n} |\text{Cov}(Z_{\mathbf{i}}, Z_{\mathbf{j}})| \\
 &\leq \frac{4 \|K\|_{\infty}^2}{nB_{n,d}^2 h_n^d} \sum_{\mathbf{i} \in \mathcal{I}_n} \sum_{\mathbf{j} \in \mathcal{I}_n: \|\mathbf{i}-\mathbf{j}\| > c_n} \phi(\|\mathbf{i}-\mathbf{j}\|) \\
 &\leq \frac{4 \|K\|_{\infty}^2}{B_{n,d}^2 h_n^d} \sum_{\mathbf{j} \in \mathcal{I}_n: t_{\mathbf{j}} > c_n} t_{\mathbf{j}}^{-\theta} \leq \frac{4 \|K\|_{\infty}^2}{B_{n,d}^2 h_n^d (\theta-1)} c_n^{1-\theta} = O\left(\frac{c_n^{1-\theta}}{h_n^d}\right). \quad (2.22)
 \end{aligned}$$

En choisissant $c_n = \left\lfloor h_n^{-\frac{d(\zeta+1)}{N+\theta-1}} \right\rfloor$, on obtient :

$$nh_n^d |I_2| = O\left(h_n^{\frac{d(\zeta(\theta-1)-N)}{N+\theta-1}}\right) \rightarrow 0 \text{ lorsque } n \rightarrow \infty.$$

D'où le résultat. ■

2.4.4 Preuve du Théorème 2.2.2

Pour $\ell = 1$, on obtient à partir de la formule (2.1) :

$$p_n(x) = \frac{1}{n} \sum_{\mathbf{i} \in \mathcal{I}_n} \lambda_{\mathbf{i}}^{-d} f(X_{\mathbf{i}})^{d/2} K\left(\frac{f(X_{\mathbf{i}})^{1/2}(x - X_{\mathbf{i}})}{\lambda_{\mathbf{i}}}\right).$$

2.4. PREUVES

Ainsi, pour le terme de biais on a :

$$\mathbb{E}(p_n(x)) = \frac{1}{n} \sum_{\mathbf{i} \in \mathcal{I}_n} \lambda_{\mathbf{i}}^{-d} \int_{\mathbb{R}^d} f(t)^{d/2} p(t) K \left(\frac{f(t)^{1/2}(x-t)}{\lambda_{\mathbf{i}}} \right) dt.$$

En effectuant le changement de variable

$$u = \frac{f(x)^{1/2}(x-t)}{\lambda_{\mathbf{i}}}$$

$$\begin{aligned} \mathbb{E}(p_n(x)) &= \frac{f(x)^{-d/2}}{n} \sum_{\mathbf{i} \in \mathcal{I}_n} \int_{\mathbb{R}^d} f(x - \eta_{\mathbf{i}} u)^{d/2} p(x - \eta_{\mathbf{i}} u) K \left(u \sqrt{\frac{f(x - \eta_{\mathbf{i}} u)}{f(x)}} \right) du \\ &= \frac{p(x)}{n} \sum_{\mathbf{i} \in \mathcal{I}_n} \int_{\mathbb{R}^d} \tau_{x,f}(\eta_{\mathbf{i}} u)^d \tau_{x,p}(\eta_{\mathbf{i}} u)^2 K(u \tau_x(\eta_{\mathbf{i}} u)) du. \end{aligned}$$

avec pour toute fonction ψ et x fixé :

$$\eta_{\mathbf{i}} := \lambda_{\mathbf{i}} f(x)^{-1/2} \text{ et } \tau_{x,\psi}(\cdot) := \sqrt{\frac{\psi(x - \cdot)}{\psi(x)}}.$$

Ce qui donne la convergence du terme de biais vers 0, du fait que $\eta_{\mathbf{i}} \rightarrow 0$ lorsque $\hat{\mathbf{i}} \rightarrow \infty$. Pour le terme de variance, on a :

$$\mathbb{V}(p_n(x)) = \frac{1}{n^2} \sum_{\mathbf{i} \in \mathcal{I}_n} \lambda_{\mathbf{i}}^{-2d} \mathbb{V}(Z_{\mathbf{i}}) + \frac{1}{n^2} \sum_{\substack{\mathbf{i}, \mathbf{j} \in \mathcal{I}_n \\ \mathbf{i} \neq \mathbf{j}}} \lambda_{\mathbf{i}}^{-d} \lambda_{\mathbf{j}}^{-d} \text{Cov}(Z_{\mathbf{i}}, Z_{\mathbf{j}}) := I_1 + I_2, \quad (2.23)$$

où

$$Z_{\mathbf{i}} := f(X_{\mathbf{i}})^{d/2} K \left(\frac{f(X_{\mathbf{i}})^{1/2}(x - X_{\mathbf{i}})}{\lambda_{\mathbf{i}}} \right).$$

Pour le premier terme de la décomposition précédente, nous écrivons :

$$\begin{aligned} n \lambda_n^d I_1 &= \frac{\lambda_n^d}{n} \sum_{\mathbf{i} \in \mathcal{I}_n} \lambda_{\mathbf{i}}^{-2d} \left\{ \int_{\mathbb{R}^d} f(t)^d p(t) K^2 \left(\frac{p(t)^{1/2}(x-t)}{\lambda_{\mathbf{i}}} \right) dt - \mathbb{E}^2(Z_{\mathbf{i}}) \right\} \\ &:= \frac{1}{n} \sum_{i=1}^n \left(\frac{\lambda_i}{\lambda_n} \right)^{-d} w_i, \end{aligned} \quad (2.24)$$

avec

$$\begin{aligned} w_i &:= f(x)^{-d/2} \int_{\mathbb{R}^d} f(x - \eta_i u)^d p(x - \eta_i u) K^2(u \tau_x(\eta_i u)) du - \lambda_i^d \mathbb{E}^2(Z_{\mathbf{s}_i}) \\ &\rightarrow f(x)^{d/2} p(x) \int_{\mathbb{R}^d} K^2(u) du, \end{aligned}$$

lorsque $\eta_i \rightarrow 0$, grâce à la continuité de p et τ_x . On a donc par application du Lemme 2.4.1 et grâce à l'hypothèse H.2.2.2 :

$$n\lambda_n^d I_1 \rightarrow B_{-d} f(x)^{d/2} p(x) \int_{\mathbb{R}^d} K^2(u) du \quad (2.25)$$

lorsque $n \rightarrow \infty$. ■

2.4.5 Preuve du Théorème 2.2.3

Posons

$$G_n(x) = p_n(x) - \mathbb{E}p_n(x) = \sum_{\mathbf{i} \in \mathcal{I}_n} \Delta_{\mathbf{i}} \text{ avec } \Delta_{\mathbf{i}} := \frac{1}{S_{n,0}} (Z_{\mathbf{i}} - \mathbb{E}Z_{\mathbf{i}}),$$

où les $Z_{\mathbf{i}}$ sont définis en (2.14). Dans la suite, a_n est un entier. On utilise la méthode des blocs en regroupant les variables aléatoires $\Delta_{\mathbf{i}}$ comme suit :

$$\begin{aligned} U(1, n, \mathbf{j}, x) &= \sum_{\substack{i_k=2j_k a_n+1 \\ k=1, \dots, N}}^{(2j_k+1)a_n} \Delta_{\mathbf{i}}; \\ U(2, n, \mathbf{j}, x) &= \sum_{\substack{i_k=2j_k a_n+1 \\ k=1, \dots, N-1}}^{(2j_k+1)a_n} \sum_{i_N=(2j_N+1)a_n+1}^{2(j_N+1)a_n} \Delta_{\mathbf{i}}; \\ U(3, n, \mathbf{j}, x) &= \sum_{\substack{i_k=2j_k a_n+1 \\ k=1, \dots, N-2}}^{(2j_k+1)a_n} \sum_{i_{N-1}=(2j_{N-1}+1)a_n+1}^{2(j_{N-1}+1)a_n} \sum_{i_N=2j_N a_n+1}^{(2j_N+1)a_n} \Delta_{\mathbf{i}}; \\ U(4, n, \mathbf{j}, x) &= \sum_{\substack{i_k=2j_k a_n+1 \\ k=1, \dots, N-2}}^{(2j_k+1)a_n} \sum_{i_{N-1}=(2j_{N-1}+1)a_n+1}^{2(j_{N-1}+1)a_n} \sum_{i_N=(2j_N+1)a_n+1}^{2(j_N+1)a_n} \Delta_{\mathbf{i}}, \end{aligned} \quad (2.26)$$

etc. On note que

$$U(2^{N-1}, n, \mathbf{j}, x) = \sum_{\substack{i_k=2j_k a_n+1 \\ k=1, \dots, N-1}}^{(2j_k+1)a_n} \sum_{i_N=2j_N a_n+1}^{2(j_N+1)a_n} \Delta_{\mathbf{i}}; \quad U(2^N, n, \mathbf{j}, x) = \sum_{\substack{i_k=2j_k a_n+1 \\ k=1, \dots, N}}^{(2j_k+1)a_n} \Delta_{\mathbf{i}}.$$

Et pour tout entier $1 \leq i \leq 2^N$, on définit

$$E(n, i, x) = \sum_{\substack{j_k=0 \\ k=1, \dots, N}}^{r_k-1} U(i, n, \mathbf{j}, x),$$

2.4. PREUVES

alors on peut écrire

$$G_n(x) = \sum_{i=1}^{2^{N+1}} E(n, i, x).$$

Sans perte de généralité, on montre le résultat pour $i = 1$. Puisque K est bornée, on obtient

$$U(1, n, \mathbf{j}, x) \leq \frac{a_n^N \|K\|_\infty}{nh_n^d}.$$

On considère les suites

$$\lambda_n = \left\lfloor (nh_n^d \log n)^{1/2} \right\rfloor, \text{ et } r_i = n_i / (2a_n), \quad i = 1, \dots, N.$$

On en déduit que, pour n assez grand,

$$\lambda_n |U(1, n, \mathbf{j}, x)| = O\left(\frac{\log n}{\sqrt{nh_n^d}}\right). \quad (2.27)$$

On numérote les variables aléatoires $U(1, n, \mathbf{j}, x)$ dans $E(n, 1, x)$ de façon arbitraire avec leurs représentants respectifs $\hat{U}_1, \dots, \hat{U}_M$ tel que $M = r_1 \dots r_N$.

Par application de l'inégalité de Markov et du Lemme 4.5 de couplage de [Carbon et al. \(1997\)](#), il existe des variables aléatoires réelles notées $\tilde{U}_1, \dots, \tilde{U}_M$ et qui sont indépendantes de $\hat{U}_1, \dots, \hat{U}_M$ telle que pour tout i , $\hat{U}_i = \hat{U}_i(x)$ a la même distribution que $\tilde{U}_i = \tilde{U}_i(x)$. De plus :

$$\mathbb{P}\left(|\hat{U}_i - \tilde{U}_i| > \epsilon\right) \leq 18 \left(\frac{\|\hat{U}_i\|_\infty}{\epsilon}\right) \psi(n, a_n^N) \phi(a_n). \quad (2.28)$$

Soit $\epsilon_n = \eta \sqrt{\frac{\log n}{nh_n^d}}$, où η est un nombre positif. Il vient que

$$\mathbb{P}(|E(n, 1, x)| > \epsilon_n) \leq \mathbb{P}\left(\sum_{j=1}^M |\tilde{U}_j| > \frac{\epsilon_n}{2}\right) + \mathbb{P}\left(\sum_{j=1}^M |\hat{U}_j - \tilde{U}_j| > \frac{\epsilon_n}{2}\right). \quad (2.29)$$

Les \tilde{U}_i étant indépendantes, l'inégalité de Bernstein entraîne que :

$$\begin{aligned} \mathbb{P}\left(\sum_{j=1}^M |\tilde{U}_j| > \frac{\epsilon_n}{2}\right) &\leq 2 \exp\left(-\frac{\lambda_n \epsilon_n}{2} + \frac{\lambda_n^2}{4} \sum_{i=1}^M \mathbb{E}\left(\hat{U}_i^2(x)\right)\right) \\ &\leq 2 \exp\left(-\frac{\eta \log n}{2} + \frac{\lambda_n^2}{4} \sum_{i=1}^M \mathbb{E}\left(\hat{U}_i^2(x)\right)\right). \end{aligned}$$

Il est clair que,

$$\frac{\lambda_n^2}{4} \sum_{i=1}^M \mathbb{E} \left(\hat{U}_i^2(x) \right) \leq \frac{\lambda_n^2}{4} \mathbb{V}(p_n(x)) \leq \frac{\log(n)}{4} n h_n^d \mathbb{V}(p_n(x))$$

De l'égalité 2.27, on en déduit que pour n assez grand,

$$n h_n^d \mathbb{V}(p_n(x)) \leq p(x) B_{-d} \int_{\mathbb{R}^d} K^2(u) du.$$

Il s'en suit donc que

$$\frac{\lambda_n^2}{4} \sum_{i=1}^M \mathbb{E} \left(\hat{U}_i^2(x) \right) \leq \log(n) p(x) \frac{B_{-d}}{4} \int_{\mathbb{R}^d} K^2(u) du.$$

D'où

$$\mathbb{P} \left(\sum_{j=1}^M |\tilde{U}_j| > \frac{\epsilon_n}{2} \right) \leq n^{-\delta}$$

où $\delta > 1$, pour η assez grand. Concernant le second terme du second membre de (2.29), on a

$$\left\| \hat{U}_i(x) \right\|_{\infty} \leq \left(\frac{a_n^N \|K\|_{\infty}}{n h_n^d} \right).$$

On obtient grâce à l'inégalité de Markov :

$$\mathbb{P} \left(\sum_{i=1}^M \left| \hat{U}_i - \tilde{U}_i \right| > \frac{\epsilon_n}{2} \right) \leq C M \left(\frac{a_n^N \|K\|_{\infty}}{n h_n^d} \right) \epsilon_n^{-1} \psi(n, a_n^N) \phi(a_n).$$

Ce qui donne

$$\begin{aligned} \mathbb{P} \left(\sum_{i=1}^M \left| \hat{U}_i - \tilde{U}_i \right| > \frac{\epsilon_n}{2} \right) &\leq C \left(\frac{1}{h_n^d} \right) \epsilon_n^{-1} \psi(n, a_n^N) (\phi(a_n)) \\ &\leq C \left(\frac{1}{h_n^d} \right) \epsilon_n^{-1} a_n^N a_n^{-\theta}. \end{aligned}$$

En posant $a_n = \left(\frac{\log n}{n h_n^d} \right)^{-1/(2N)}$, on obtient

$$\mathbb{P} \left(\sum_{i=1}^M \left| \hat{U}_i - \tilde{U}_i \right| > \frac{\epsilon_n}{2} \right) \leq C n^{\frac{2N-\theta}{2N}} \log n^{\frac{\theta-2N}{2N}} h_n^{\frac{-d\theta}{2N}} = C \beta_n.$$

2.4. PREUVES

On rappelle que $n = \hat{n} = n_1 \times n_2 \times \dots \times n_N$ et de l'hypothèse H.2.2.8 on déduit que

$$nu_n \beta_n = \left(n \log n^{\frac{\theta-2N}{4N-\theta}} h_n^{\frac{d\theta}{\theta-4N}} u_n^{\frac{2N}{4N-\theta}} \right)^{\frac{4N-\theta}{2N}} \rightarrow 0 \text{ lorsque } n \rightarrow \infty.$$

Étant donné $\epsilon > 0$, on peut trouver $n_0 \in \mathbb{N}$ tel que si $n \geq n_0$ alors

$$|nu_n \beta_n| \leq \epsilon.$$

Les suites étant à termes positifs, on obtient :

$$\beta_n \leq \frac{\epsilon}{nu_n}.$$

Or d'après l'hypothèse H. 2.2.8,

$$\sum_{n \geq 1} C \beta_n = \sum_{(n_1, \dots, n_N) \in \mathbb{N}^N} \frac{\epsilon C}{nu_n} < \infty.$$

Le Théorème 2.2.3 découle alors du lemme de Borel-Cantelli. ■

2.4.6 Preuve du Théorème 2.3.1

On a :

$$\hat{\rho}(\mathbf{z}) = \log \frac{p(\mathbf{z})[1 + \delta_p(\mathbf{z})]}{q(\mathbf{z})[1 + \delta_q(\mathbf{z})]} = \rho(\mathbf{z}) + \log \frac{1 + \delta_p(\mathbf{z})}{1 + \delta_q(\mathbf{z})}$$

avec

$$\delta_p(\mathbf{z}) := \frac{\hat{p}(\mathbf{z}) - p(\mathbf{z})}{p(\mathbf{z})} \text{ et } \delta_q(\mathbf{z}) := \frac{\hat{q}(\mathbf{z}) - q(\mathbf{z})}{p(\mathbf{z})}.$$

D'après le Théorème 2.2.3 on a $\delta_p(\mathbf{z}) \rightarrow 0$ et $\delta_q(\mathbf{z}) \rightarrow 0$ p.s, lorsque $n_1, n_2 \rightarrow 0$. Ainsi, un développement limité de la fonction log permet d'écrire :

$$\hat{p}(\mathbf{z}) = \rho(\mathbf{z}) + \delta_p(\mathbf{z}) - \delta_q(\mathbf{z}) + O_{\mathbb{P}} [\delta_p^2(\mathbf{z})] + O_{\mathbb{P}} [\delta_q^2(\mathbf{z})].$$

Par conséquent pour n suffisamment grand, on obtient grâce au Théorème 2.2.1,

$$\mathbb{E}(\hat{p}(\mathbf{z})) = \rho(\mathbf{z}) + \mathbb{E}(\delta_p(\mathbf{z})) - \mathbb{E}(\delta_q(\mathbf{z})) + o_{\mathbb{P}}(1)$$

et grâce à l'hypothèse H.2.3.2, on a aussi

$$\text{Var}(\hat{p}(\mathbf{z})) = \rho(\mathbf{z}) + \text{Var}(\delta_p(\mathbf{z})) - \text{Var}(\delta_q(\mathbf{z})) + o_{\mathbb{P}}(1)$$

On en déduit les résultats du Théorème 2.3.1 par application de ceux obtenus au Théorème 2.2.1. ■

Chapitre 3

Estimateurs récursifs de la densité d'un flux de données spatio-temporelles

Sommaire

3.1	Flux de données spatio-temporelles	40
3.2	Présentation de l'estimateur	42
3.3	Hypothèses et résultats asymptotiques	43
3.4	Résultats numériques	48
3.4.1	Sélection des paramètres	49
3.4.2	Résultats numériques	51
3.5	Preuves	53
3.5.1	Preuve du Théorème 3.3.1	54
3.5.2	Preuve du Théorème 3.3.2	59

Dans ce chapitre, nous nous intéressons à l'estimation non-paramétrique de la densité de probabilité d'un "Data Stream" de nature spatio-temporelle. Nous proposons et étudions un estimateur à noyau de la fonction de densité de probabilité d'un flux de données spatio-temporelles. Plus précisément, nous adaptons l'estimateur à noyau classique de Parzen-Rosenblatt. Pour cela, nous proposons de combiner la méthodologie sur l'estimateur récursif de la densité et celle d'une distribution de nature spatiale ([Tran \(1990\)](#)).

3.1 Flux de données spatio-temporelles

Soient $d, N \geq 1$ et X une variable aléatoire à valeurs dans \mathbb{R}^d , définie sur l'espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$. Un flux de données spatio-temporelles est un processus spatiotemporel pour lequel les données d'entrée peuvent être décomposées en une suite de matrices de la forme

$$W_{(\mathbf{s},t)} := \left\{ X_{(\mathbf{s},t)1}, \dots, X_{(\mathbf{s},t)k(\mathbf{s},t)} \right\}, \quad \mathbf{s} \in \mathbb{N}^N, \quad t \in \mathbb{Z},$$

indexées par l'espace et le temps, où le sous-échantillon $X_{(\mathbf{s},t)1}, \dots, X_{(\mathbf{s},t)k(\mathbf{s},t)}$ est une suite de vecteurs aléatoires de \mathbb{R}^d de même distribution qu'une variable aléatoire X .

Dans le cadre de flux de données spatio-temporelles considéré ici, $W_{(\mathbf{s},t)}$ est désigné comme une fenêtre et $k(\mathbf{s}, t)$ sa largeur, considérée comme un nombre entier.

En s'inspirant du cas traité par [Hall *et al.* \(2005\)](#) dans le cadre non spatial, on peut considérer le cas où les tailles des lots $W_{(\mathbf{s},t)}$ sont aléatoires ; c'est-à-dire supposer que $k(\mathbf{s}, t)$ un processus spatio-temporel également. En effet, dans un cadre temporel, ces derniers ont considéré une estimation de la densité dans laquelle les tailles des lots sont aléatoires, lorsque la distribution sous-jacente évolue avec le temps et les observations sont indépendantes.

Dans ce chapitre, nous traitons le cas d'un flux de données spatio-temporelles stationnaire, ce qui signifie que la distribution des $X_{(\mathbf{s},t)j}$ est invariant dans l'espace et le temps. L'objectif est de fournir un estimateur de la densité f des $X_{(\mathbf{s},t)j}$ qui peut être mis à jour facilement quand un nouveau lot de données est disponible, tout en gardant de bonnes propriétés théoriques.

Nous supposons qu'un utilisateur recueille de manière séquentielle un flux de données spatio-temporelles à travers le temps dans $\{1, \dots, T\}$ et l'espace dans un ensemble spatial \mathcal{I}_n de cardinal n , d'une région $\mathcal{D} \subset \mathbb{Z}^N$, où \mathbb{Z}^N est muni de la métrique uniforme définie au Chapitre 2.

Posons

$$\mathcal{D}_{n,T} = \mathcal{I}_n \times \{1, \dots, T\},$$

l'ensemble d'observations du processus $\{W_{(\mathbf{s},t)}\}_{\mathbf{s},t}$.

Dans ce chapitre, nous nous intéressons au cas où le processus satisfait une condi-

CHAPITRE 3. ESTIMATEURS RÉCURSIFS DE LA DENSITÉ D'UN FLUX DE DONNÉES SPATIO-TEMPORELLES

tion de dépendance faible qui est plus générale que le mélange fort dans la mesure où la faible dépendance couvre une grande classe de processus. En effet, le mélange fort est utile pour caractériser la dépendance d'un certain nombre de processus, et permet d'obtenir les mêmes résultats asymptotiques comme dans le cas indépendant. Cependant, certains processus qui sont d'intérêt pour les statistiques ne sont pas fortement mélangeants, mais remplissent des conditions de dépendance faible. Pour une introduction plus exhaustive sur la faible dépendance, nous nous référons à [Doukhan et Louhichi \(1999\)](#) et [Dedecker et al. \(2007\)](#). On rappelle d'abord quelques définitions d'une suite doublement indexée.

Définition 3.1.1 Soit $u : \mathbb{N}^2 \rightarrow \mathbb{R}$ une suite doublement indexée de nombres réels. On dit que $(u(n, m))$ converge vers $a \in \mathbb{R}$ et on écrit

$$\lim_{n, m \rightarrow \infty} u(n, m) = a$$

si la condition suivante est satisfaite : $\forall \epsilon > 0, \exists N, M$ tel que si $\forall n \geq N, m \geq N$, on ait $|u(n, m) - a| < \epsilon$.

Définition 3.1.2 Une suite doublement indexée $(u(n, m))$ est dite bornée s'il existe une constante $C > 0$ tel que $|u(n, m)| < C$ pour tout $n, m \in \mathbb{N}$.

Définition 3.1.3 Soit $(u(n, m))$ une suite doublement indexée de nombres réels :

- Si $u(n, m) \leq u(n', m') \quad \forall (n, m) \leq_{lex} (n', m') \in \mathbb{N}^2$, on dit que la suite $(u(n, m))$ est décroissante avec \leq_{lex} l'ordre lexicographique dans \mathbb{N}^2 défini par :

$$(n, m) \leq_{lex} (n', m') \Leftrightarrow n \leq n' \text{ ou } (n = n' \text{ et } m \leq m'),$$

- Si $u(n, m) \geq u(j, k) \quad \forall (n, m) \leq_{lex} (j, k) \in \mathbb{N}^2$, on dit que la suite $(u(n, m))$ est croissante.

Il existe différents coefficients qui sont définis pour décrire la faible dépendance, voir [Dedecker et al. \(2007\)](#).

Nous utiliserons les notions de η -dépendance introduites par [Doukhan et Louhichi \(1999\)](#) et que nous adapterons au cadre spatio-temporel qui nous intéresse.

3.2. PRÉSENTATION DE L'ESTIMATEUR

Définition 3.1.4 Soit une fonctions $\psi : \mathbb{N}^2 \times \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$, la suite spatio-temporelle

$$\left\{ X_{(\mathbf{s},t)}, (\mathbf{s}, t) \in \mathcal{D}_{n,T} \right\}$$

de variables aléatoires est dite ψ -faiblement-dépendante s'il existe une suite décroissante $\{\epsilon(r, m), (r, m) \in \mathbb{N}^2\}$ tendant vers 0 quand $r, m \rightarrow \infty$ telle que, pour tous sous-ensembles

$$\mathcal{S} = \{\mathbf{s}_1, \dots, \mathbf{s}_j\} \text{ et } \mathcal{S}' = \{\mathbf{s}'_1, \dots, \mathbf{s}'_k\}$$

de \mathcal{I}_n et pour tout (t_1, t_2, \dots, t_j) et (t'_1, \dots, t'_k) avec $t_1 \leq t_2 \leq \dots \leq t_k \leq t_j + z = t'_1 \leq \dots \leq t'_k$, l'inégalité suivante est vérifiée :

$$\begin{aligned} & \left| \text{Cov} \left(g_0 \left(X_{(\mathbf{s}_1,t)}, \dots, X_{(\mathbf{s}_j,t)} \right), g_1 \left(X_{(\mathbf{s}'_1,t)}, \dots, X_{(\mathbf{s}'_k,t)} \right) \right) \right| \\ & \leq \psi \left(j \|g_1\|_\infty, k \|g_0\|_\infty, \text{Lip}(g_0), \text{Lip}(g_1) \right) \epsilon(z, w), \end{aligned} \quad (3.1)$$

où $w := \text{dist}(\mathcal{S}, \mathcal{S}')$ est la distance entre \mathcal{S} et \mathcal{S}' et $g_p : \mathbb{R}^{d(j(1-p)+kp)} \rightarrow \mathbb{R}$, $p = 0, 1$ sont des fonctions arbitraires bornées, telles que

$$\|g_p\|_\infty < \infty \text{ et } \text{Lip}(g_p) := \sup_{u \neq v} \frac{|g_p(u) - g_p(v)|}{\|u - v\|_1} < \infty, \text{ avec } \|x\|_1 = \sum_{i=1}^{d(j(1-p)+kp)} |x_i|. \quad (3.2)$$

Définition 3.1.5 Dans l'équation (3.1), quand

$$\psi(j, k, u, v) = ju + kv,$$

le coefficient ϵ sera noté η . Dans ce cas, on dira simplement que le processus

$$\left\{ X_{(\mathbf{s},t)}, (\mathbf{s}, t) \in \mathcal{D}_{n,T} \right\}$$

est η -faiblement-dépendant.

3.2 Présentation de l'estimateur

Pour simplifier la présentation, et puisque le cas général est plus complexe en terme de notations, nous étudierons l'estimateur de la densité simplifié lorsque $k(\mathbf{s}, t) = 1$. Ainsi, nous considérons l'estimateur de la densité suivant :

$$f_{n,T}(x) := \frac{1}{S_{n,T}} \sum_{\mathbf{i} \in \mathcal{I}_n} \sum_{t=1}^T K \left(\frac{X_{(\mathbf{i},t)} - x}{h_{(\mathbf{i},t)}} \right), \quad x \in \mathbb{R}^d, \quad (3.3)$$

où

$$S_{n,T} := \sum_{i \in \mathcal{I}_n} \sum_{t=1}^T h_{(i,t)}^d.$$

3.3 Hypothèses et résultats asymptotiques

Pour évaluer le biais et la variance de (3.3), nous supposons les hypothèses suivantes.

Hypothèse H. 3.3.1 *Le processus*

$$\left\{ X_{(\mathbf{s},t)}, (\mathbf{s}, t) \in \mathcal{D}_{n,T} \right\}$$

est η -dépendant . De plus, il existe η_1 et η_2 avec $\eta_1(0) = \eta_2(0) = 1$ tel que pour tout $z, w \in \mathbb{R}$,

$$\eta(z, w) = \eta_1(z)\eta_2(w); \lim_{m \rightarrow \infty} m^\nu \sum_{r>m} \eta_1(r) = 0 \quad \text{et} \quad \lim_{m \rightarrow \infty} m^{\mu-N} \sum_{r>m} r^N \eta_2(r) = 0.$$

où $\nu > \frac{d+1}{2}$ et $\mu > Nd$.

Hypothèse H. 3.3.2 *K est une densité bornée et symétrique telle que*

$$\int_{\mathbb{R}^d} uK(u)du = 0, \quad \text{et} \quad \int_{\mathbb{R}^d} \|u\|^2 |K(u)| du < \infty.$$

Hypothèse H. 3.3.3 (i) $h_{(n,T)} \downarrow 0$, quand $n, T \rightarrow \infty$.

(ii) Pour tout $\rho \in]-\infty, 2+d]$,

$$B_{n,T,\rho} := \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \left(\frac{h_{(i,t)}}{h_{(n,T)}} \right)^\rho \rightarrow B_\rho \quad \text{quand} \quad (n, T) \rightarrow \infty.$$

Hypothèse H. 3.3.4 (i) *Le processus spatio-temporel*

$$\left\{ X_{(\mathbf{s},t)}, (\mathbf{s}, t) \in \mathcal{D}_{n,T} \right\}$$

est strictement stationnaire de densité marginale f .

3.3. HYPOTHÈSES ET RÉSULTATS ASYMPTOTIQUES

(ii) Pour tous $(\mathbf{i}, t), (\mathbf{j}, \tau) \in \mathcal{D}_{n,T}$ tels que $(\mathbf{i}, t) \neq (\mathbf{j}, \tau)$, le vecteur aléatoire $(X_{(\mathbf{i},t)}, X_{(\mathbf{j},\tau)})$ admet une densité jointe $f_{(X_{(\mathbf{i},t)}, X_{(\mathbf{j},\tau)})}$ telle que :

$$M := \sup_{\|\mathbf{i}-\mathbf{j}\| \geq 1, |t-\tau| \geq 1} |f_{(X_{(\mathbf{i},t)}, X_{(\mathbf{j},\tau)})} - f \times f| < \infty.$$

Hypothèse H. 3.3.5 La densité f est bornée et 2 fois différentiable sur \mathbb{R}^d .

La condition **H. 3.3.1** est la version spatio-temporelle de l'hypothèse de dépendance faible considérée par [Doukhan et Louhichi \(1999\)](#) pour l'estimation de densité dans le cadre temporel. Les conditions de convergence ci-dessus sur les coefficients de dépendance sont considérées pour simplifier les calculs et sont semblables à celles proposées par [Doukhan et Louhichi \(2001\)](#).

A titre d'exemples de fenêtres vérifiant les hypothèses **H. 3.3.1**, $h_{(\hat{\mathbf{s}},t)} = (\hat{\mathbf{s}}t)^{-\nu}$, $0 < \nu < 1$ avec $\hat{\mathbf{s}} = \text{Card}(\{\mathbf{i} \in \mathcal{I}_n, \mathbf{i} <_{lex} \mathbf{s}\})$ où $<_{lex}$ est l'ordre lexicographique de \mathbb{N}^N . On voit dans le cadre d'un rectangle où les observations se portent sur n sites, $\hat{\mathbf{n}} = n$, $h_{(\mathbf{n},t)} = (nt)^{-\nu}$ et pour le cas spatial, il suffit de prendre $t = 1$.

L'hypothèse **H.3.3.2** est courante pour l'estimation à noyau de la densité.

H. 3.3.3 (ii) est une hypothèse technique utilisée dans la méthode récursive à noyau, c'est la version spatio-temporelle des conditions utilisés par [Amiri \(2010\)](#) dans le cas temporel et par [Amiri et al. \(2015\)](#) dans le cas spatial. Elle vient du fait que nous devons contrôler asymptotiquement la mise à jour de la fenêtre et est remplie par les choix habituels de cette dernière. **H. 3.3.4** concerne la stationnarité des flux de données. Cette condition est moins réaliste en pratique (elle peut être assouplie moyennant une complexité supplémentaire sur les preuves), puisqu'en pratique la majorité des flux de données est générée par des procédés qui peuvent changer considérablement au fil du temps et/ou de l'espace. Dans ce dernier contexte plus général, nous pouvons considérer l'hypothèse que $X_{(\mathbf{s},t)1}, \dots, X_{(\mathbf{s},t)k(\mathbf{s},t)}$ est une séquence de vecteurs aléatoires de \mathbb{R}^d , identiquement distribués suivant une densité de probabilité $f_{(\mathbf{s},t)}$. Dans ce cas, $f_{(\mathbf{s},t)}$ reflète le comportement local et à court terme des phénomènes décrits par le flux, alors que le comportement à long terme des phénomènes sur un site peut être décrit par une densité globale tenant compte du passé du flux jusqu'au temps t (voir par exemple [Hall et al. \(2005\)](#), [Cao et al. \(2012\)](#)). Nos résultats théoriques peuvent alors être étendus dans ce cas général par une estimation de la densité globale

CHAPITRE 3. ESTIMATEURS RÉCURSIFS DE LA DENSITÉ D'UN FLUX DE DONNÉES SPATIO-TEMPORELLES

prenant en compte le passé du flux.

Avec les hypothèses précédentes, nous sommes en mesure de donner des résultats théoriques. Le Théorème 3.3.1 établit la convergence en moyenne quadratique de l'estimateur récursif (3.3).

Théorème 3.3.1 *Sous les hypothèses H. 3.3.1 - H. 3.3.5 si*

$$nTh_{(n,T)}^d \rightarrow \infty \text{ quand } (n, T) \rightarrow \infty,$$

alors

$$nTh_{(n,T)}^d \mathbb{V}(f_{n,T}(x)) \rightarrow f(x) B_d^{-1} \int_{\mathbb{R}^d} K^2(u) du \quad (3.4)$$

et

$$h_{(n,T)}^{-2} \left(\mathbb{E}(f_{n,T}(x)) - f(x) \right) \rightarrow B_d^{-1} B_{d+2} \int_{\mathbb{R}^d} \frac{K(u)}{2} \sum_{j_1, j_2=1}^d u_{j_1} u_{j_2} \frac{\partial^2 f(x)}{\partial x_{j_1} \partial x_{j_2}} du \quad (3.5)$$

quand $(n, T) \rightarrow \infty$, pour tous x tels que $f(x) > 0$.

Le Théorème 3.3.1 est une extension des résultats obtenus dans le cas temporel par Amiri (2010) avec un processus α -mélangeant. En outre, un résultat similaire a été prouvé par Mezhoud *et al.* (2014) pour des processus temporels η -dépendants. Dans le cas particulier où $h_{(\mathbf{i}, t)} = h_{(n, T)}$ pour tout $\mathbf{i} \in \mathcal{I}_n$, $t \in \{1, \dots, T\}$, nous obtenons du Théorème 3.3.1, la convergence en moyenne quadratique de l'estimateur à noyau de Parzen-Rosenblatt, $f_{n,T}^{PR}(x)$ dans le cadre spatio-temporel. Notons que les expressions du biais et la variance de notre estimateur dépendent de la constante B_r . Ceci permet une comparaison avec l'estimateur Parzen-Rosenblatt en termes de variance. Avec le choix usuel de la fenêtre

$$h_{(n,T)} = (nT)^{-1/(d+4)},$$

nous avons

$$\frac{\mathbb{V}f_{n,T}(x)}{\mathbb{V}f_{n,T}^{PR}(x)} \rightarrow B_d^{-1} = \frac{4}{d+4}.$$

3.3. HYPOTHÈSES ET RÉSULTATS ASYMPTOTIQUES

Par conséquent, la variance de l'estimateur récursif de la densité est réduite. La comparaison est également possible par rapport au biais asymptotique. Les relations (3.4) et (3.5) nous permettent de déduire une borne exacte de l'erreur quadratique moyenne (MSE) et les fenêtres optimales de l'estimateur récursif. Par conséquent, si les hypothèses du Théorème 3.3.1 sont vérifiées, alors pour (n, T) assez grands, on obtient :

$$\begin{aligned} MSE\left(f_{n,T}(x)\right) &\sim h_{(n,T)}^4 \left(B_d^{-1} B_{d+2} \int_{\mathbb{R}^d} \frac{K(u)}{2} \sum_{j_1, j_2=1}^d u_{j_1} u_{j_2} \frac{\partial^2 f(x)}{\partial x_{j_1} \partial x_{j_2}} du \right)^2 \\ &\quad + \frac{B_d^{-1}}{nTh_{(n,T)}^d} f(x) \int_{\mathbb{R}^2} K(u) du, \end{aligned}$$

où la fenêtre optimale est donnée par

$$h_{n,T_{opt}} = \frac{\left(dB_d^{-1} f(x) \int_{\mathbb{R}^d} K(u) du \right)^{\frac{1}{d+4}}}{\left(B_d^{-1} B_{d+2} \int_{\mathbb{R}^d} K(u) \sum_{j_1, j_2=1}^d u_{j_1} u_{j_2} \frac{\partial^2 f(x)}{\partial x_{j_1} \partial x_{j_2}} du \right)^{\frac{2}{d+4}}} (nT)^{-\frac{1}{d+4}}. \quad (3.6)$$

Notons que pour obtenir nos résultats asymptotiques dans le contexte d'un réseau irrégulier, on peut considérer des hypothèses supplémentaires sur le nombre d'unités spatiales sur une boule fermée de \mathcal{D} où la région \mathcal{D} est un réseau irrégulier et infini dénombrable. En outre, tous les éléments de \mathcal{D} sont censés se situer à des distances d'au moins $\delta > 1$ de l'autre. Dans ce cas, nous considérons une séquence de régions finies, convexes et fermées $\{R_u\}$ du *increasing domain* quand $u \rightarrow +\infty$ et supposons que l'échantillon est constitué de l'intersection de l'une des régions convexes et $\mathcal{D} : \mathcal{I}_n = \mathcal{D} \subset R_u$. À part l'hypothèse d'une zone de plus en plus uniforme de $\{R_U\}$ dans au moins deux directions non opposées pour augmenter la taille de l'échantillon n (quand $u \rightarrow +\infty$) (voir Conley (1999)), nous n'avons pas spécifié un ordre et toute autre hypothèse sur la configuration et la croissance de la taille de l'échantillon.

Le théorème suivant donne la convergence presque sûre de (3.3). Pour cela, on utilise une version spatio-temporelle de la notion de τ -faible-dépendance introduite par Dedecker et Prieur (2004), puisque la η -dépendance est plus générale

CHAPITRE 3. ESTIMATEURS RÉCURSIFS DE LA DENSITÉ D'UN FLUX DE DONNÉES SPATIO-TEMPORELLES

mais insuffisante pour atteindre la convergence uniforme presque sûre dans notre contexte.

Définition 3.3.1 Soit $\mathcal{L}(\mathbb{R}^d)$ l'ensemble des fonctions

$$g : \mathbb{R}^d \rightarrow \mathbb{R} \text{ telles que } \text{Lip}(g) < 1.$$

Pour toute fonction aléatoire intégrable, X à valeurs dans \mathbb{R}^d , et définie sur $(\Omega, \mathcal{A}, \mathbb{P})$ et pour tout σ -algèbre \mathcal{M} dans \mathcal{A} , soit

$$\tau(\mathcal{A}, X) = \mathbb{E} \left(\sup_{g \in \mathcal{L}(\mathbb{R}^d)} \left| \mathbb{P}_{X|\mathcal{M}}(g) - \mathbb{P}_X(g) \right| \right),$$

où \mathbb{P}_X et $\mathbb{P}_{X|\mathcal{M}}$ sont les distributions de X et $X|\mathcal{M}$ respectivement. Définissons pour tous sous-ensemble de localisations $\mathcal{S} = \{\mathbf{s}_1, \dots, \mathbf{s}_j\}$ et $\mathcal{S}' = \{\mathbf{s}'_1, \dots, \mathbf{s}'_k\}$ de \mathcal{I}_n et pour tout (t_1, t_2, \dots, t_j) et (t'_1, \dots, t'_k) tel que $t_1 \leq t_2 \leq \dots \leq t_j + z = t'_1 \leq \dots \leq t'_k$:

$$\tau(z, w) = \sup_{n > 0} \max_{1 \leq k \leq n} \frac{1}{k} \Theta_k$$

avec

$$\Theta_k := \sup_{\substack{\mathcal{S} = \{\mathbf{s}_1, \dots, \mathbf{s}_j\}, \mathcal{S}' = \{\mathbf{s}'_1, \dots, \mathbf{s}'_k\} \\ t_1 \leq t_2 \leq \dots \leq t_k \leq t_j + z = t'_1 \leq \dots \leq t'_k}} \left\{ \tau \left(\sigma \left(X_{(\mathbf{s}_1, t)}, \dots, X_{(\mathbf{s}_j, t)} \right), X_{(\mathbf{s}'_1, t)}, \dots, X_{(\mathbf{s}'_k, t)} \right) \right\}.$$

On dira que la suite

$$\left\{ X_{(\mathbf{s}, t)}, (\mathbf{s}, t) \in \mathcal{D}_{n, T} \right\}$$

de variables aléatoires est τ -faiblement-dépendant si pour tout $\tau(z, w) \rightarrow 0$ quand $z, w \rightarrow \infty$.

Notons que, selon [Dedecker et al. \(2007\)](#), on peut écrire que pour tout $z, w > 0$,

$$\eta(z, w) \leq \tau(z, w).$$

Théorème 3.3.2 Soit D un sous-ensemble compact de \mathbb{R}^d . Supposons que

$$\left\{ X_{(\mathbf{s}, t)} (\mathbf{s}, t) \in \mathcal{D}_{n, T} \right\}$$

3.4. RÉSULTATS NUMÉRIQUES

est τ -dépendant et qu'il existe $\tau_1(\cdot)$ et $\tau_2(\cdot)$, avec $\tau_1(0) = \tau_2(0) = 1$ tel que $\tau(z, w) = \tau_1(z)\tau_2(w)$ pour tout $z, w \in \mathbb{R}$:

$$\tau_1(z) = O(\exp(-\rho_1 z)) \text{ et } \tau_2(w) = O(\exp(-\rho_2 w)) \quad (3.7)$$

où $\rho_1, \rho_2 > 0$. Si les hypothèses **H. 3.3.2** - **H. 3.3.5** sont vraies, alors le choix $h_{n,T} = O((nT)^{-q'})$ avec

$$\rho_1 > \frac{dq' + 3}{2}, \rho_2 > \frac{dq' + 3}{2} \text{ et } 0 < q' < \frac{1}{d},$$

entraîne que

$$\limsup_{(n,T) \rightarrow \infty} \sqrt{\frac{(Tn)^{1-dq'}}{\log nT}} \sup_{x \in D} |f_{n,T}(x) - f(x)| \leq \frac{1}{2} \left(\sup_{x \in D} f(x) B_d^{-1} \int_{\mathbb{R}^d} K^2(u) du + 2 \right) \quad p.s. \quad (3.8)$$

Le Théorème 3.3.2 est une extension des travaux de [Carbon et al. \(1997\)](#) dans le cas d'un processus spatial α -mélangeant au cadre d'un processus spatio-temporel τ -faiblement dépendant.

3.4 Résultats numériques

Cette section fournit une étude de simulation pour analyser le comportement de notre estimateur récursif de la densité pour certaines données spatio-temporelles faiblement dépendantes. Le travail effectué dans ce chapitre concerne le cas de flux de données spatio-temporelles, dans lequel les données sont capturées continuellement au fil du temps en $n = 36$ sites d'observation fixés et pour lesquelles des mises à jour en temps réel de l'estimation sont nécessaires. Le scénario utilisé a été fourni par le package **stream**¹ disponible dans l'environnement **R**. Cet outil a un support pour la lecture et l'écriture de flux de données via une connexion **R** qui permet d'accéder à des données produites à l'extérieur de l'environnement **R**. Afin de générer un tableau de données spatio-temporelles, des vecteurs

$$\mathbf{X}_t = (X_{(1,t)}, \dots, X_{(36,t)})^T, t \geq 1$$

1. <https://cran.r-project.org/web/packages/stream/stream.pdf>

CHAPITRE 3. ESTIMATEURS RÉCURSIFS DE LA DENSITÉ D'UN FLUX DE DONNÉES SPATIO-TEMPORELLES

sont simulés grâce au processus Gaussien spatio-temporel défini comme étant la solution de l'équation stochastique donnée par (à l'aide du package **spate**² de **R**) :

$$\frac{\partial}{\partial t} X_{(\mathbf{i},t)} = - \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} \nabla X_{(\mathbf{i},t)} + \nabla S' S \nabla X_{(\mathbf{i},t)} - \zeta X_{(\mathbf{i},t)} + \epsilon_{(\mathbf{i},t)}, \quad X_{(1,1)} = 0,$$

avec $t > 0$ $\mathbf{i} \in \mathcal{I}_n$, ∇ est l'opérateur gradient et

$$S = \frac{1}{\rho_1} \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\gamma \sin \alpha & \gamma \cos \alpha \end{pmatrix},$$

où $\epsilon_{(\mathbf{i},t)}$ est un processus spatial Gaussien, bruit blanc (en temps) avec une structure de dépendance spatiale décrite par la fonction de covariance de Matérn de densité spectrale :

$$\hat{f}(\mathbf{k}) = \frac{\sigma^2}{2\pi} \left(\|\mathbf{k}\|^2 + \frac{1}{\rho_0^2} \right)^{-2},$$

où $\sigma > 0$, $\rho_0 > 0$ sont les paramètres de covariance, $\mathbf{k} \in \mathbb{N}^N$.

Une fois que la base de données simulées est créée et enregistrée sur le disque, nous ouvrons une connexion vers le fichier où elle a été écrite et nous la traitons comme un flux de données spatio-temporelles. Puis, à tout instant t , la ligne de données \mathbf{X}_t est lue en retour. Dans ce contexte, dans lequel il y a une taille d'échantillon variable dans le temps, notre méthode d'estimation démarre à l'instant $t_0 = 1$ et l'estimateur de la densité est continuellement mis à jour par rapport au temps jusqu'à atteindre l'instant final $T = 50$ (correspondant à la taille des échantillons : 36, 72, ..., 1800). En outre, afin d'éviter des calculs inutiles, l'algorithme est interrompu si la distance absolue entre deux valeurs de l'estimateur de la densité obtenus en deux étapes successives est inférieure à 10^{-5} .

3.4.1 Sélection des paramètres

Dans cette étude de simulation, nous avons choisi les valeurs suivantes des paramètres du processus spatio-temporel :

$$\rho_0 = 0.05, \sigma = 0.7, \zeta = -\log(0.99), \rho_1 = 0.06, \gamma = 3, \alpha = \pi/4, \mu_x = \mu_y = -0.1.$$

2. <https://cran.r-project.org/web/packages/spate/spate.pdf>

3.4. RÉSULTATS NUMÉRIQUES

En pratique, pour être en mesure de mettre en œuvre l'estimateur récursif étudié dans ce chapitre, le choix de certains paramètres est nécessaire. Le noyau Gaussien est considéré dans la mise en œuvre des deux algorithmes considérés (récursives et non récursives). Il est bien connu que le problème de la fenêtre optimale est fondamental dans le cadre non paramétrique et a fait l'objet de nombreuses recherches. L'estimation de ce paramètre est un problème crucial qui affecte notamment le niveau de lissage de l'estimateur qui en résulte. De grandes valeurs de la fenêtre donnent un estimateur avec un biais important et une faible variabilité, tandis que des valeurs trop petites donnent un estimateur avec un petit biais et une grande variabilité.

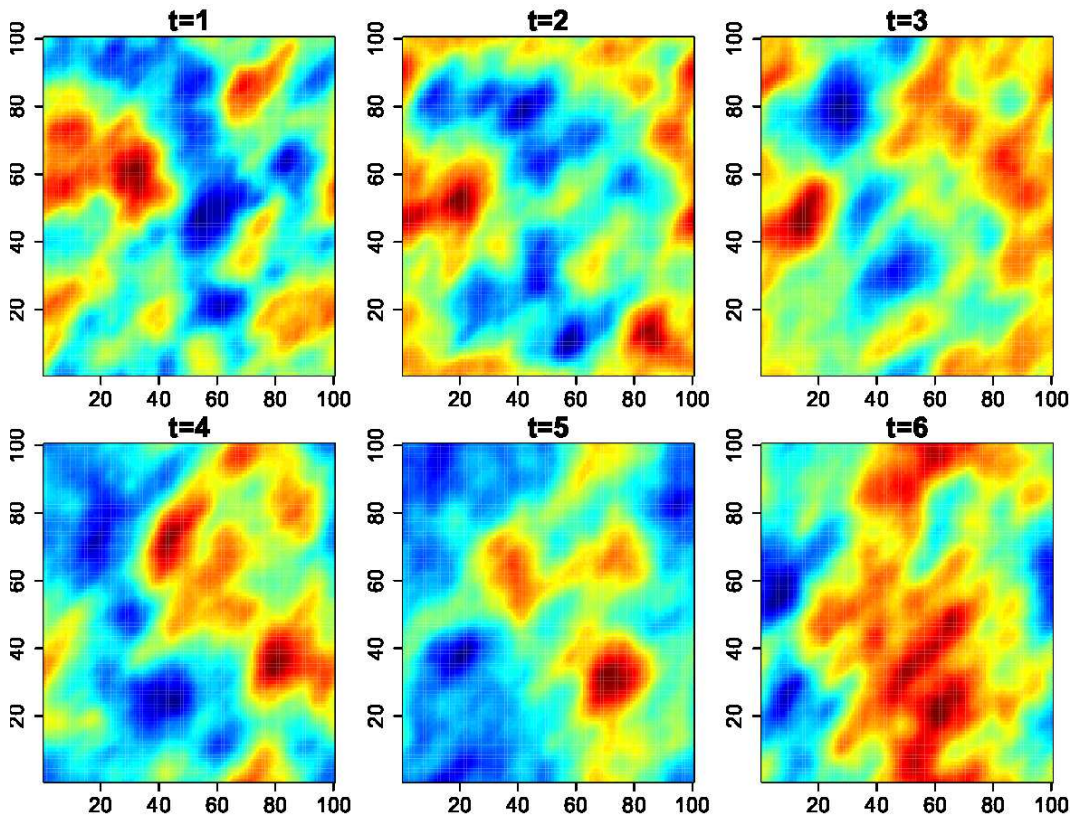


FIGURE 3.1 – Une réalisation du processus spatio-temporel $\mathbf{X}_t, t = 1, \dots, 6$ et $n = 10^4$.

Dans cette section, nous avons sélectionné une fenêtre spatialement et temporellement adaptative, conformément à l'algorithme suivant :

```

| pour tout  $k = t_0, \dots, T$  :
  | Fenêtre
    | considérer le vecteur  $\mathbf{X} = \{X_{(\mathbf{s},k)}, \mathbf{i} \in \mathcal{I}_n\}$  ;
    | calculer  $\sigma_k$ , l'écart-type de l'échantillon  $\{X_{(\mathbf{s},t)}, \mathbf{s} \in \mathcal{I}_n, t = 1, \dots, k\}$  ;
    | pour tout  $t = 1, \dots, k$  :
      | pour  $\mathbf{s} \in \mathcal{I}_n$  :
        | calculer les fenêtres adaptatives :  $h_{(\hat{\mathbf{s}},t)} = 1.06 \times \sigma_k(\hat{\mathbf{s}}k)^{-1/5}$ ;
        | fin de boucle
      | fin de boucle
    | sauvegarder la matrice de fenêtres  $h_k = \{h_{(\hat{\mathbf{s}},t)}, \mathbf{s} \in \mathcal{I}_n, t = 1, \dots, k\}$ ;
    | fin de boucle
  | fin de boucle

```

Le choix d'une fenêtre adaptative spatio-temporelle est motivé par le fait qu'une fenêtre fixe est moins efficace quand il y a une in-homogénéité locale de la distribution. Les fenêtres $h_{\hat{\mathbf{s}},t}$, $(\mathbf{s}, t) \in \mathcal{D}$ peuvent être sélectionnées en utilisant une autre approche, telle qu'une version récursive de la méthode de validation croisée basée sur la minimisation de l'erreur quadratique. Plus précisément, des estimateurs récursifs d'une suite de fenêtres sont donnés par une combinaison convexe : $\hat{h}_{\hat{\mathbf{s}},t} = (1 - \gamma_t)\hat{h}_{\hat{\mathbf{s}},t-1} + \gamma_t\hat{C}V(W_{(\mathbf{s},t)})$ où $\hat{C}V(W_{(\mathbf{s},t)})$ est la fenêtre sélectionnée par une validation croisée basée sur des données à jour des lots ($W_{(\mathbf{s},t)}$).

3.4.2 Résultats numériques

Pour évaluer la performance de notre estimateur, nous utilisons la moyenne des erreurs quadratiques (MSE) sur 100 points x_i de $[0, 1]$:

$$MSE = \frac{1}{100M} \sum_{m=1}^M \sum_{i=1}^{100} \left(f(x_i) - f_{n,T}^{[m]}(x_i) \right)^2$$

3.4. RÉSULTATS NUMÉRIQUES

où $f_{n,T}^{[m]}(x_i)$ est la valeur de l'estimateur $f(x_i)$ calculée de la $m^{\text{ième}}$ réplication par la procédure de Monte-Carlo aux point x_i , $i = 1, \dots, 100$ de l'échantillon W_t , $t = 1, \dots, T$. La Table 3.1 donne la moyenne de 100 réplifications de la MSE des estimateurs de la densité (récuratif et non-récuratif) évaluée en ces 100 points.

T	10	20	30	40	50
Estimateur récuratif	0.00129	0.00025	0.00015	0.00013	0.00013
Estimateur non-récuratif	0.00144	0.00034	0.00034	0.00021	0.00020

TABLE 3.1 – MSE pour comparer les estimateurs récuratif et non-récuratif

La performance est également évaluée par rapport au temps de calcul de l'estimateur de la densité à un point $x \in [0, 1]$. La Table 3.2 donne le temps moyen de calcul en secondes pour les estimateurs de la densité basé sur 100 simulations.

L'analyse des Tables 3.1 et 3.2 révèle que (i) les MSE sont en faveur de l'estimateur non récuratif mais les MSE des deux estimateurs coïncident si les résultats sont approximés à 10^{-3} près et (ii) le temps de calcul est clairement en faveur de l'estimateur récuratif. En outre, comme prévu, nous observons que les erreurs d'estimation diminuent et le temps de calcul diminue quand le nombre de fenêtres du flux de données augmente. Enfin, les résultats montrent que le temps de calcul cumulé de $T = 1$ à $T = 500$ est de 1.90 secondes pour l'estimateur récuratif et 2757.501 secondes pour le non-récuratif. L'avantage de notre approche récurative dans le cadre d'un flux de données spatio-temporelles réside dans le gain de temps de calcul considérable enregistré sans une augmentation substantielle des erreurs d'estimation.

T	100	200	300	400	500
Estimateur récuratif	0.003	0.003	0.004	0.004	0.005
Estimateur non-récuratif	0.382	1.107	2.193	3.783	57.919

TABLE 3.2 – Temps de calcul en secondes pour comparer les estimateurs récuratif et non-recursif

3.5 Preuves

Dans la suite, C est une constante dont la valeur peut varier d'une ligne à l'autre. Avant d'établir les preuves des résultats principaux, nous donnons un lemme préliminaire, qui est une conséquence d'un résultat de [Robison \(1926\)](#). Il s'agit d'une généralisation du lemme de Toeplitz dans le cadre d'une série doublement indexée.

Lemme 3.5.1 *Soit $(w_{n,T})_{n \geq 1, T \geq 1}$ une suite bornée de limite finie w . Sous l'hypothèse **H 3.3.3**,*

$$\frac{1}{nT} \sum_{k=1}^n \sum_{t=1}^T \left(\frac{h_{(k,t)}}{h_{(n,T)}} \right)^r w_{k,t} \rightarrow B_r w \text{ quand } (n, T) \rightarrow \infty.$$

Preuve

Si on pose :

$$a_{k,t}(n, T) = \begin{cases} \frac{1}{nT} \left(\frac{h_{(k,t)}}{h_{(n,T)}} \right)^r & \text{if } k \leq n \text{ et } t \leq T \\ 0 & \text{sinon} \end{cases},$$

alors, **H. 3.3.3**(i) permet d'avoir,

$$\text{pour } k \geq 1 \text{ et } t \geq 1, \quad a_{k,t}(n, T) \leq \frac{h_{(k,t)}^r}{nT h_{(n,T)}^r}.$$

Ainsi

$$\lim_{(n,T) \rightarrow \infty} a_{k,t}(n, T) = 0 \text{ pour tout } k \geq 1 \text{ et } t \geq 1. \quad (3.9)$$

À l'aide de **H. 3.3.3**(i), nous obtenons

$$\sum_{k \geq 1} |a_{k,t}(n, T)| \leq \sum_{k=1}^n \frac{h_{(1,t)}^r}{nT h_{(n,T)}^r} \leq \frac{h_{(1,t)}^r}{T h_{(n,T)}^r} \rightarrow 0 \quad (3.10)$$

et de manière similaire

$$\sum_{t \geq 1} |a_{k,t}(n, T)| \leq \frac{h_{(k,1)}^r}{n h_{(n,T)}^r} \rightarrow 0,$$

3.5. PREUVES

lorsque $(n, T) \rightarrow \infty$. De plus, **H. 3.3.3(ii)** permet de dire qu' il existe une constante $A > 0$ telle que pour $n \geq 1$ et $T \geq 1$,

$$\sum_{k \geq 1, t \geq 1} |a_{k,t}(n, T)| < A < \infty. \quad (3.11)$$

En utilisant une fois de plus **H. 3.3.3(ii)**, nous avons

$$\lim_{(n, T) \rightarrow \infty} \sum_{k=1}^n \sum_{t=1}^T a_{k,t}(n, T) = \lim_{(n, T) \rightarrow \infty} B_{n, T, r} = B_r < \infty. \quad (3.12)$$

Les équations (3.9)-(3.12) et le Théorème 3 de Robison (1926), p. 67 permettent de conclure. ■

Par simplicité, notons dans la suite

$$K_{(\mathbf{i}, t)} = K \left(\frac{X_{(\mathbf{i}, t)} - x}{h_{(\mathbf{i}, t)}} \right) \text{ pour } x \in \mathbb{R}^d, \mathbf{i} \in \mathcal{I}_n \text{ et } t = 1, \dots, T.$$

3.5.1 Preuve du Théorème 3.3.1

Considérons en premier lieu le terme de variance :

$$nTh_{(n, T)}^d \mathbb{V} \left(f_{n, T}(x) \right) = \frac{1}{nTh_{(n, T)}^d B_{n, T, d}^2} \left(\begin{array}{l} \sum_{\mathbf{i} \in \mathcal{I}_n} \sum_{t=1}^T \mathbb{V} \left(K_{(\mathbf{i}, t)} \right) + \\ \sum_{\mathbf{i} \in \mathcal{I}_n} \sum_{\substack{t=1 \\ t \neq \tau}}^T \sum_{\tau=1}^T \text{Cov} \left(K_{(\mathbf{i}, t)}, K_{(\mathbf{i}, \tau)} \right) + \\ \sum_{\substack{\mathbf{i}, \mathbf{j} \in \mathcal{I}_n \\ \mathbf{i} \neq \mathbf{j}}} \sum_{t=1}^T \text{Cov} \left(K_{(\mathbf{i}, t)}, K_{(\mathbf{j}, t)} \right) + \\ \sum_{\substack{\mathbf{i}, \mathbf{j} \in \mathcal{I}_n \\ \mathbf{i} \neq \mathbf{j}}} \sum_{\substack{t=1 \\ t \neq \tau}}^T \sum_{\tau=1}^T \text{Cov} \left(K_{(\mathbf{i}, t)}, K_{(\mathbf{j}, \tau)} \right) \end{array} \right) \\ := I_1 + I_2 + I_3 + I_4. \quad (3.13)$$

CHAPITRE 3. ESTIMATEURS RÉCURSIFS DE LA DENSITÉ D'UN FLUX DE DONNÉES SPATIO-TEMPORELLES

Étudions d'abord le terme I_1 :

$$I_1 = \frac{1}{B_{n,T,d}^2} \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \left(\frac{h_{(i,t)}}{h_{(n,T)}} \right)^d \left\{ \begin{array}{l} \int_{\mathbb{R}^d} K^2(u) f(x + uh_{(i,t)}) du \\ - \left(\int_{\mathbb{R}^d} K(u) f(x + uh_{(i,t)}) du \right)^2 h_{(i,t)}^d \end{array} \right\}.$$

Les hypothèses **H. 3.3.2**, **H. 3.3.3**, **H. 3.3.5** et le Lemme 3.5.1 permettent d'avoir

$$I_1 \rightarrow f(x) B_d^{-1} \int_{\mathbb{R}^d} K^2(u) du, \text{ quand } (n, T) \rightarrow \infty. \quad (3.14)$$

Considérons les termes de covariance dans (3.13). En ce qui concerne I_2 , nous avons

$$I_2 \leq \frac{2}{nT h_{(n,T)}^d B_{n,T,d}^2} \sum_{\mathbf{i} \in \mathcal{I}_n} \sum_{t=1}^T \left\{ \begin{array}{l} \sum_{\substack{\tau=1 \\ 1 \leq |\tau-t| \leq d_{(n,T)}}}^T \left| \text{Cov} (K_{(\mathbf{i},t)}, K_{(\mathbf{i},\tau)}) \right| \\ + \sum_{\substack{\tau=1 \\ |\tau-t| > d_{(n,T)}}}^T \left| \text{Cov} (K_{(\mathbf{i},t)}, K_{(\mathbf{i},\tau)}) \right| \end{array} \right\}$$

où

$$d_{(n,T)} = \left\lfloor h_{(n,T)}^{-\frac{d(\zeta+1)+1}{\nu+1}} \right\rfloor, \text{ avec } \frac{d+1}{\nu d} < \zeta \leq \frac{2}{d}.$$

A l'aide des hypothèses **H. 3.3.3** et **H. 3.3.4**, nous obtenons :

$$\begin{aligned} |I_{21}| &\leq \frac{2M d_{(n,T)}}{nT h_{(n,T)}^d B_{n,T,d}^2} \sum_{i=1}^n \sum_{t=1}^T h_{(i,t)}^{2d} \leq \frac{C d_{(n,T)} h_{n,T}^{d\zeta} B_{n,T,d(\zeta+1)}}{B_{n,T,d}^2} \\ &= O\left(d_{(n,T)} h_{(n,T)}^{d\zeta}\right) = o(1), \quad \text{lorsque } (n, T) \rightarrow \infty. \end{aligned}$$

Concernant le terme I_{22} , nous obtenons à l'aide de **H. 3.3.1**, **H. 3.3.3** (i) et (3.1), pour tout $\mathbf{i} \in \mathcal{I}_n$ et tout $t, \tau = 1, \dots, T$:

$$\left| \text{Cov} (K_{(\mathbf{i},t)}, K_{(\mathbf{i},\tau)}) \right| \leq \frac{\|K\|_\infty \text{Lip}(K)}{h_{(i,t)} + h_{i,\tau}} \eta_1(|\tau - t|) \leq \frac{2\|K\|_\infty \text{Lip}(K)}{h_{(n,T)}} \eta_1(|\tau - t|).$$

Cette dernière inéquation, combinée à **H. 3.3.1** et **H. 3.3.3** impliquent :

$$\begin{aligned} |I_{22}| &\leq \frac{4}{nT h_{(n,T)}^d B_{n,T,d}^2} \sum_{\mathbf{i} \in \mathcal{I}_n} \sum_{t=1}^T \sum_{r>d_{(n,T)}}^{T-1} \left| \text{Cov} (K_{(\mathbf{i},t)}, K_{(\mathbf{i},r+t)}) \right| \\ &\leq \frac{C}{B_{n,T,d}^2 h_{(n,T)}^{d+1}} \sum_{r>d_{(n,T)}} \eta_1(r) = O\left(\frac{d_{(n,T)}^{-\nu}}{h_{(n,T)}^{d+1}}\right) = o(1), \quad \text{quand } (n, T) \rightarrow \infty. \end{aligned}$$

3.5. PREUVES

Il s'en suit

$$I_2 = o(1) \quad \text{lorsque } (n, T) \rightarrow \infty. \quad (3.15)$$

Etudions maintenant I_3 . Pour cela, nous écrivons

$$I_3 \leq \frac{C}{nTh_{(n,T)}^d B_{n,T,d}^2} \sum_{\mathbf{i} \in \mathcal{I}_n} \left\{ \sum_{\substack{\mathbf{j} \in \mathcal{I}_n \\ 1 \leq \|\mathbf{i} - \mathbf{j}\| \leq c_{(n,T)}}} \sum_{t=1}^T h_{(\mathbf{j},t)}^{2d} + \sum_{\substack{\mathbf{j} \in \mathcal{I}_n \\ \|\mathbf{i} - \mathbf{j}\| > c_{(n,T)}}} \sum_{t=1}^T \left| \text{Cov} \left(K_{(\mathbf{i},t)}, K_{(\mathbf{j},t)} \right) \right| \right\} \\ := I_{31} + I_{32},$$

où

$$c_{(n,T)} = \left\lfloor h_{(n,T)}^{-\frac{d(\zeta'+1)+1}{N+\mu}} \right\rfloor, \quad \text{avec } \frac{N(d+1)}{\mu d} < \zeta' \leq \frac{2}{d}.$$

Par des arguments similaires à ceux utilisés pour étudier I_{21} et I_{22} , nous obtenons pour (n, T) assez grands

$$|I_{31}| \leq \frac{C c_{(n,T)}^N h_{(n,T)}^{d\zeta'} B_{n,T,d(\zeta'+1)}}{B_{n,T,d}^2} = O \left(c_{(n,T)}^N h_{(n,T)}^{d\zeta'} \right) = o(1)$$

et

$$|I_{32}| \leq \frac{2 \text{Lip}(K)}{nh_{(n,T)}^{d+1} B_{n,T,d}^2} \sum_{\substack{\mathbf{i}, \mathbf{j} \in \mathcal{I}_n \\ \|\mathbf{i} - \mathbf{j}\| > c_{(n,T)}}} \eta_2(\|\mathbf{i} - \mathbf{j}\|) \leq C \frac{c_{(n,T)}^{-N}}{h_{(n,T)}^{d+1} B_{n,T,d}^2} \sum_{r > c_{(n,T)}} r^N \eta_2(r) \\ = O \left(\frac{c_{(n,T)}^{-\mu}}{h_{(n,T)}^{d+1}} \right) = o(1).$$

Ainsi

$$I_3 = o(1) \quad \text{as } (n, T) \rightarrow \infty. \quad (3.16)$$

Pour étudier le dernier terme de (3.13), posons

$$b_{(n,T)} = \left\lfloor h_{(n,T)}^{-\frac{d(\zeta''+1)+1}{N+\mu}} \right\rfloor, \quad a_{(n,T)} = \left\lfloor h_{(n,T)}^{-\frac{d(\zeta''+1)+1}{\nu+1}} \right\rfloor, \quad (3.17)$$

avec

$$\frac{(d+1)(N(\nu+2)+\mu)}{d(\nu\mu-N)} < \zeta'' \leq \frac{2}{d},$$

CHAPITRE 3. ESTIMATEURS RÉCURSIFS DE LA DENSITÉ D'UN FLUX
DE DONNÉES SPATIO-TEMPORELLES

et

$$S = \left\{ (\mathbf{i}, t, \mathbf{j}, \tau) \in \mathcal{D}_{n,T}^2, \mathbf{i} \neq \mathbf{j}, t \neq \tau : 0 < \|\mathbf{i} - \mathbf{j}\| \leq b_{(n,T)}, \quad 0 < |t - \tau| \leq a_{(n,T)} \right\};$$

nous pouvons alors écrire :

$$|I_4| = \left| \frac{1}{nTh_{(n,T)}^d B_{n,T,d}^2} \left\{ \sum_{(\mathbf{i}, \mathbf{j}, \tau) \in S} \text{Cov} \left(K_{(\mathbf{i}, t)}, K_{(\mathbf{j}, \tau)} \right) + \sum_{(\mathbf{i}, \mathbf{j}, \tau) \in S^c} \text{Cov} \left(K_{(\mathbf{i}, t)}, K_{(\mathbf{i}, \tau)} \right) \right\} \right|$$

Donc :

$$|I_4| = \left| \frac{1}{nTh_{(n,T)}^d B_{n,T,d}^2} \left\{ \begin{aligned} & \sum_{\substack{\mathbf{i}, \mathbf{j} \in \mathcal{I}_n \\ \|\mathbf{i} - \mathbf{j}\| \leq b_{(n,T)}}} \sum_{t=1}^T \sum_{\tau=1}^T \left| \text{Cov} \left(K_{(\mathbf{i}, t)}, K_{(\mathbf{j}, \tau)} \right) \right| \\ & + \sum_{\substack{\mathbf{i}, \mathbf{j} \in \mathcal{I}_n \\ \|\mathbf{i} - \mathbf{j}\| \leq b_{(n,T)}}} \sum_{t=1}^T \sum_{\tau=1}^T \left| \text{Cov} \left(K_{(\mathbf{i}, t)}, K_{(\mathbf{j}, \tau)} \right) \right| \\ & + \sum_{\substack{\mathbf{i}, \mathbf{j} \in \mathcal{I}_n \\ \|\mathbf{i} - \mathbf{j}\| > b_{(n,T)}}} \sum_{t=1}^T \sum_{\tau=1}^T \left| \text{Cov} \left(K_{(\mathbf{i}, t)}, K_{(\mathbf{j}, \tau)} \right) \right| \\ & + \sum_{\substack{\mathbf{i}, \mathbf{j} \in \mathcal{I}_n \\ \|\mathbf{i} - \mathbf{j}\| > b_{(n,T)}}} \sum_{t=1}^T \sum_{\tau=1}^T \left| \text{Cov} \left(K_{(\mathbf{i}, t)}, K_{(\mathbf{i}, \tau)} \right) \right| \end{aligned} \right\} \right|$$

$$:= I_{41} + I_{42} + I_{43} + I_{44}.$$

En utilisant les mêmes arguments que pour l'étude de I_3 , nous obtenons

$$\begin{aligned} |I_{42}| &\leq \frac{4 \|K\|_\infty \text{Lip}(K) 2^N b_{(n,T)}^N}{h_{(n,T)}^{d+1} B_{n,T,d}^2} \sum_{r=a_{(n,T)}+1}^{T-1} \eta_1(r) \\ &= O \left(\frac{b_{(n,T)}^N}{a_{(n,T)}^\nu h_{(n,T)}^{d+1}} \right) = o(1) \end{aligned}$$

et

$$|I_{43}| \leq \frac{Ca_{(n,T)} b_{(n,T)}^{-N}}{h_{(n,T)}^{d+1} B_{n,T,d}^2} \sum_{r=b_{(n,T)}+1}^n r^N \eta_2(r) = O \left(\frac{a_{(n,T)} b_{(n,T)}^{-\mu}}{h_{(n,T)}^{d+1}} \right) = o(1).$$

3.5. PREUVES

Ces mêmes arguments, donnent pour n et T assez grands,

$$|I_{44}| = O\left(\frac{a_{(n,T)} b_{(n,T)}^{-\mu}}{h_{(n,T)}^{d+1}}\right) = o(1).$$

Finalement, nous obtenons pour le premier terme

$$\begin{aligned} |I_{41}| &\leq \frac{2M h_{1,1}^{d(1-\zeta'')} h_{(n,T)}^{d\zeta''}}{B_{n,T,d}^2} \frac{1}{nT} \sum_{t=1}^T \sum_{\substack{\mathbf{i}, \mathbf{j} \in \mathcal{I}_n \\ j: \|\mathbf{i}-\mathbf{j}\| \leq b_{(n,T)}}} \sum_{\substack{\tau=1 \\ \tau: |t-\tau| \leq a_{(n,T)}}}^T \left(\frac{h_{(\hat{\mathbf{i}}, t)}}{h_{(n,T)}}\right)^{d(1+\zeta'')} \\ &\leq \frac{C b_{(n,T)}^N a_{(n,T)} h_{(n,T)}^{d\zeta''}}{B_{n,T,d(\zeta''+1)}^{-1} B_{n,T,d}^2} = O\left(b_{(n,T)}^N a_{(n,T)} h_{(n,T)}^{d\zeta''}\right) = o(1). \end{aligned}$$

Ceci met fin à la preuve de (3.4).

Etudions maintenant le terme de biais. Pour cela, notons que

$$\begin{aligned} \mathbb{E}(f_{n,T}(x)) - f(x) &= \frac{1}{S_{n,T}} \sum_{\mathbf{i} \in \mathcal{I}_n} \sum_{t=1}^T \left\{ \int_{\mathbb{R}^d} K\left(\frac{z-x}{h_{(\hat{\mathbf{i}}, t)}}\right) f(z) dz \right. \\ &\quad \left. - \int_{\mathbb{R}^d} h_{(\hat{\mathbf{i}}, t)}^d K(u) f(x) du \right\} \\ &= \frac{1}{S_{n,T}} \sum_{\mathbf{i} \in \mathcal{I}_n} \sum_{t=1}^T \int_{\mathbb{R}^d} h_{(\hat{\mathbf{i}}, t)}^d K(u) \left[f(x + h_{(\hat{\mathbf{i}}, t)} u) - f(x) \right] du. \end{aligned}$$

La formule de Taylor et l'hypothèse **H. 3.3.5** permettent d'écrire :

$$\begin{aligned} \mathbb{E}(f_{n,T}(x)) - f(x) &= \frac{1}{h_{(n,T)}^d B_{n,T,d} nT} \sum_{\mathbf{i} \in \mathcal{I}_n} \sum_{t=1}^T h_{(\hat{\mathbf{i}}, t)}^{d+2} \\ &\quad \times \int_{\mathbb{R}^d} \frac{K(u)}{2} \sum_{j_1, j_2=1}^d u_{j_1} u_{j_2} \frac{\partial^2 f(x + \theta u h_{(\hat{\mathbf{i}}, t)})}{\partial x_{j_1} \partial x_{j_2}} du \\ &= \frac{h_{(n,T)}^2}{B_{n,T,d} nT} \sum_{i=1}^n \sum_{t=1}^T \left(\frac{h_{(i,t)}}{h_{(n,T)}}\right)^{d+2} \\ &\quad \times \int_{\mathbb{R}^d} \frac{K(u)}{2} \sum_{j_1, j_2=1}^d u_{j_1} u_{j_2} \frac{\partial^2 f(x + \theta u h_{(i,t)})}{\partial x_{j_1} \partial x_{j_2}} du, \end{aligned}$$

avec $\theta \in]0, 1[$. Le théorème de convergence dominée et le Lemme 3.5.1 donnent ainsi le résultat (3.5). ■

3.5.2 Preuve du Théorème 3.3.2

Puisque D est un ensemble compact, il peut être couvert par un nombre fini v d'hypercubes I_k centrés en x_k et de rayon r . Ainsi

$$\begin{aligned} \sup_{x \in D} \left| f_{n,T}(x) - \mathbb{E}f_{n,T}(x) \right| &\leq \max_{1 \leq k \leq v} \sup_{x \in I_k} \left| f_{n,T}(x) - f_{n,T}(x_k) \right| \\ &\quad + \max_{1 \leq k \leq v} \sup_{x \in I_k} \left| \mathbb{E}f_{n,T}(x_k) - \mathbb{E}f_{n,T}(x) \right| \\ &\quad + \max_{1 \leq k \leq v} \left| f_{n,T}(x_k) - \mathbb{E}f_{n,T}(x_k) \right| = Q_1 + Q_2 + Q_3 \quad (3.18) \end{aligned}$$

L'hypothèse **H. 3.3.2** permet d'avoir

$$\begin{aligned} \left| f_{n,T}(x) - f_{n,T}(x_k) \right| &= \left| \frac{1}{S_{n,T}} \sum_{\mathbf{i} \in \mathcal{I}_n} \sum_{t=1}^T K \left(\frac{X_{(\mathbf{i},t)} - x}{h_{(\mathbf{i},t)}} \right) - K \left(\frac{X_{(\mathbf{i},t)} - x_k}{h_{(\mathbf{i},t)}} \right) \right| \\ &\leq \frac{L \|x - x_k\|}{S_{n,T}} \sum_{\mathbf{i} \in \mathcal{I}_n} \sum_{t=1}^T \frac{1}{h_{(\mathbf{i},t)}} \\ &\leq \frac{L \|x - x_k\|}{nT h_{n,T}^d B_{n,T,d}} \sum_{\mathbf{i} \in \mathcal{I}_n} \sum_{t=1}^T \frac{1}{h_{(\mathbf{i},t)}} \\ &\leq \frac{L \|x - x_k\|}{B_{n,T,d}} \frac{1}{h_{(n,T)}^{d+1}} \end{aligned}$$

ainsi le choix

$$r = h_{(n,T)}^{(d+1)} \sqrt{\frac{\log(nT)}{T n h_{(n,T)}^d}},$$

implique

$$\left| f_{n,T}(x) - f_{n,T}(x_k) \right| \leq L B_{n,T,d}^{-1} \sqrt{\frac{\log(nT)}{T n h_{(n,T)}^d}} \rightarrow 0 \text{ quand } n, T \rightarrow \infty.$$

Nous obtenons donc

$$\lim_{n,T \rightarrow \infty} Q_1 = 0, \text{ et } \lim_{n,T \rightarrow \infty} Q_2 = 0 \quad (3.19)$$

Pour étudier Q_3 , posons

$$G_{n,T}(x) = f_{n,T}(x) - \mathbb{E}f_{n,T}(x) = \sum_{\mathbf{i} \in \mathcal{I}_n} \sum_{t=1}^T \Delta_{\mathbf{i},t}$$

3.5. PREUVES

avec

$$\Delta_{\mathbf{i},t} := \frac{1}{S_{n,T}} \left(K_{(\mathbf{i},t)} - \mathbb{E}K_{(\mathbf{i},t)} \right),$$

considérons les suites

$$\lambda_{n,T} = \left\lfloor (Tnh_{(n,T)}^d \log nT)^{1/2} \right\rfloor,$$

$$p_n = \left\lfloor \left(\frac{nh_{(n,T)}^{d/2}}{\log n} \right)^{1/2N} \right\rfloor, \quad q_n = \left\lfloor (\log n)^{1/N} \right\rfloor, \quad p_T = \left\lfloor \left(\frac{Th_{(n,T)}^{d/2}}{\log T} \right)^{1/2} \right\rfloor, \quad q_T = \lfloor \log T \rfloor. \quad (3.20)$$

Nous pouvons alors représenter les variables aléatoires $\Delta_{\mathbf{i},t}$ avec les blocs suivants :

$$\begin{aligned} U(1, n, \mathbf{j}, T, u, x) &= \sum_{\substack{i_k = j_k(p_n + q_n) + 1 \\ k=1, \dots, N}}^{j_k(p_n + q_n) + p_n} \sum_{t=u(p_T + q_T) + 1}^{u(p_T + q_T) + p_T} \Delta_{\mathbf{i},t}; \\ U(2, n, \mathbf{j}, T, u, x) &= \sum_{\substack{i_k = j_k(p_n + q_n) + 1 \\ k=1, \dots, N}}^{j_k(p_n + q_n) + p_n} \sum_{t=u(p_T + q_T) + p_T + 1}^{(u+1)(p_T + q_T)} \Delta_{\mathbf{i},t}; \\ U(3, n, \mathbf{j}, T, u, x) &= \sum_{\substack{i_k = j_k(p_n + q_n) + 1 \\ k=1, \dots, N-1}}^{j_k(p_n + q_n) + p_n} \sum_{i_N = j_N(p_n + q_n) + p_n + 1}^{(j_k+1)(p_n + q_n)} \sum_{t=u(p_T + q_T) + 1}^{u(p_T + q_T) + p_T} \Delta_{\mathbf{i},t}; \\ U(4, n, \mathbf{j}, T, u, x) &= \sum_{\substack{i_k = j_k(p_n + q_n) + 1 \\ k=1, \dots, N-1}}^{j_k(p_n + q_n) + p_n} \sum_{i_N = j_N(p_n + q_n) + p_n + 1}^{(j_k+1)(p_n + q_n)} \sum_{t=u(p_T + q_T) + p_T + 1}^{(u+1)(p_T + q_T)} \Delta_{\mathbf{i},t} \quad (3.21) \end{aligned}$$

et ainsi de suite. On a :

$$\begin{aligned} U(2^N, n, \mathbf{j}, T, u, x) &= \sum_{\substack{i_k = j_k(p_n + q_n) + p_n + 1 \\ k=1, \dots, N}}^{(j_k+1)(p_n + q_n)} \sum_{t=u(p_T + q_T) + 1}^{u(p_T + q_T) + p_T} \Delta_{\mathbf{i},t}; \\ U(2^{N+1}, n, \mathbf{j}, T, u, x) &= \sum_{\substack{i_k = j_k(p_n + q_n) + p_n + 1 \\ k=1, \dots, N}}^{(j_k+1)(p_n + q_n)} \sum_{t=u(p_T + q_T) + p_T + 1}^{(u+1)(p_T + q_T)} \Delta_{\mathbf{i},t}. \end{aligned}$$

CHAPITRE 3. ESTIMATEURS RÉCURSIFS DE LA DENSITÉ D'UN FLUX DE DONNÉES SPATIO-TEMPORELLES

On pose :

$$r_k = \frac{n_k}{(p_n + q_n)}, \quad k = 1, \dots, N; \quad r_{N+1} = \frac{T}{(p_T + q_T)}$$

et pour tout entier $1 \leq \ell' \leq 2^N$

$$E(n, \ell', T, x) = \sum_{k=1, \dots, N}^{r_k-1} \sum_{u=0}^{r_{N+1}-1} U(\ell', n, \mathbf{j}, T, u, x),$$

On peut donc écrire

$$G_{n,T}(x) = \sum_{\ell'=1}^{2^{N+1}} E(n, \ell', T, x).$$

Sans perte de généralité nous traitons seulement le terme $E(n, 1, T, x)$. Puisque K est bornée on obtient :

$$U(1, n, \mathbf{j}, T, u, x) \leq \frac{p_n^N p_T \|K\|_\infty}{nT h_{(n,T)}^d}.$$

pour (n, T) suffisamment grands , on en déduit alors que

$$\lambda_{n,T} |U(1, n, \mathbf{j}, T, u, x)| = O\left(\sqrt{\frac{\log nT}{\log n \log T}}\right) = o(1). \quad (3.22)$$

Enumérons les variables $U(1, n, \mathbf{j}, T, u, x)$ dans un ordre arbitraire et nommons les $\hat{U}_1, \dots, \hat{U}_M$, $M = r_1 \dots r_{N+1}$.

Puisque $\hat{U}_1, \dots, \hat{U}_M$ sont τ -dépendants, alors par le Lemme 5 de [Dedecker et Prieur \(2004\)](#), il existe des variables aléatoires $\tilde{U}_1, \dots, \tilde{U}_M$ indépendantes, de mêmes distributions que $\hat{U}_1, \dots, \hat{U}_M$ telles que

$$\mathbb{E} \left| \hat{U}_i - \tilde{U}_i \right| \leq \frac{p_n^N p_T \|K\|_\infty}{nT h_{(n,T)}^d} \tau_1(q_n) \tau_2(q_T)$$

Par l'inégalité de Markov, nous avons

$$\mathbb{P} \left(\sum_{i=1}^M \left| \hat{U}_i - \tilde{U}_i \right| > \epsilon \right) \leq \frac{M p_n^N p_T \|K\|_\infty}{\epsilon nT h_{(n,T)}^d} \tau_1(q_n) \tau_2(q_T) \quad (3.23)$$

3.5. PREUVES

Soit

$$\epsilon_{n,T} = C^* \sqrt{\frac{\log(nT)}{nTh_{(n,T)}^d}},$$

où C^* un nombre positif dont la valeur sera spécifiée plus tard. En conséquence,

$$\begin{aligned} \mathbb{P}(|E(n, 1, T, x)| > \epsilon_{n,T}) &\leq \mathbb{P}\left(\left|\sum_{j=1}^M \tilde{U}_j\right| > \frac{\epsilon_{n,T}}{2}\right) \\ &+ \mathbb{P}\left(\left|\sum_{j=1}^M \hat{U}_j - \tilde{U}_j\right| > \frac{\epsilon_{n,T}}{2}\right). \end{aligned} \quad (3.24)$$

Pour le premier terme du second membre de (3.24), grâce à l'indépendance des \tilde{U}_i et l'application de l'inégalité de Bernstein nous obtenons

$$\begin{aligned} \mathbb{P}\left(\left|\sum_{j=1}^M \tilde{U}_j\right| > \frac{\epsilon_{n,T}}{2}\right) &\leq 2 \exp\left(-\frac{\lambda_{n,T}\epsilon_{n,T}}{2} + \frac{\lambda_{n,T}^2}{4} \sum_{i=1}^M \mathbb{E}(\hat{U}_i^2)\right) \\ &\leq 2 \exp\left(-\frac{C^* \log nT}{2} + \frac{\lambda_{n,T}^2}{4} \sum_{i=1}^M \mathbb{E}(\hat{U}_i^2)\right). \end{aligned}$$

D'une part :

$$\frac{\lambda_{n,T}^2}{4} \sum_{i=1}^M \mathbb{E}(\hat{U}_i^2) \leq \frac{\lambda_{n,T}^2}{4} \mathbb{V}(f_{n,T}(x)) \leq \frac{\log(nT)}{4} nTh_{(n,T)}^d \mathbb{V}(f_{n,T}(x))$$

Pour contrôler le terme $\mathbb{V}(f_{n,T}(x))$, observons que grâce à (3.7), on obtient

$$m^\nu \sum_{r>m} \tau_1(r) = O(\exp((\nu - \rho_1)m)) = o(1) \text{ pour tout } \nu < \rho_2$$

et

$$m^{\mu-N} \sum_{r>m} r^N \tau_2(r) = O(\exp((\mu - \rho_2)m)) = o(1) \text{ pour tout } \mu < \rho_2.$$

En particulier, quand

$$\frac{d+1}{2} < \mu < \rho_1 \text{ et } Nd < \mu < \rho_2,$$

l'hypothèse **H. 3.3.1** est vérifiée. De plus, puisqu'un processus τ -faiblement dépendant est η -faiblement dépendant, nous pouvons alors utiliser le résultat concernant

CHAPITRE 3. ESTIMATEURS RÉCURSIFS DE LA DENSITÉ D'UN FLUX DE DONNÉES SPATIO-TEMPORELLES

la variance dans le cadre d'un processus η -faiblement dépendant. Il vient alors, à partir de (3.4), que pour n, T assez grands,

$$nTh_{(n,T)}^d \mathbb{V}(f_{n,T}(x)) \leq f(x)B_d^{-1} \int_{\mathbb{R}^d} K^2(u)du \leq \sup_{x \in D} f(x)B_d^{-1} \int_{\mathbb{R}^d} K^2(u)du.$$

Ainsi :

$$\frac{\lambda_{n,T}^2}{4} \sum_{i=1}^M \mathbb{E}(\hat{U}_i^2) \leq \log(nT) \sup_{x \in D} f(x) \frac{B_d^{-1}}{4} \int_{\mathbb{R}^d} K^2(u)du.$$

Par conséquent :

$$\begin{aligned} \mathbb{P} \left(\left| \sum_{j=1}^M \tilde{U}_j \right| > \frac{\epsilon_{n,T}}{2} \right) &\leq 2 \exp \left[\left(-\frac{C^*}{2} + \sup_{x \in D} f(x) \frac{B_d^{-1}}{4} \int_{\mathbb{R}^d} K^2(u)du \right) \log(nT) \right] \\ &\leq C(nT)^{-\delta} \end{aligned}$$

où

$$\delta = \frac{C^*}{2} - \sup_{x \in D} f(x) \frac{B_d^{-1}}{4} \int_{\mathbb{R}^d} K^2(u)du.$$

Etudions maintenant le deuxième terme du second membre de (3.24). À partir de (3.23), on a :

$$\mathbb{P} \left(\left| \sum_{i=1}^M \hat{U}_i - \tilde{U}_i \right| > \frac{\epsilon_{n,T}}{2} \right) \leq \frac{n}{2^N} \frac{\tau_1(q_n)\tau_2(q_T)}{nh_{(n,T)}^d \epsilon_{n,T}} \quad (3.25)$$

Par conséquent :

$$\mathbb{P} \left(\left| \sum_{i=1}^M \hat{U}_i - \tilde{U}_i \right| > \frac{\epsilon_{n,T}}{2} \right) \leq C \sqrt{\frac{Tn}{h_{(n,T)}^d \log nT}} \tau_1(q_n)\tau_2(q_T). \quad (3.26)$$

Ainsi l'hypothèse (3.7), permet d'écrire :

$$\mathbb{P} \left(\left| \sum_{i=1}^M \hat{U}_i - \tilde{U}_i \right| > \frac{\epsilon_{n,T}}{2} \right) \leq C \left(\frac{1}{n^{\alpha_1} T^{\alpha_2}} \right) \frac{1}{(\log nT)^{1/2}} := \beta_{n,T}, \quad (3.27)$$

avec

$$\alpha_1 = \frac{2\rho_1 - 1 - dq'}{2} \text{ and } \alpha_2 = \frac{2\rho_2 - 1 - dq'}{2}.$$

3.5. PREUVES

Finalement,

$$\mathbb{P} \left(\sup_{x \in D} |E(n, 1, T, x)| > \epsilon_{n,T} \right) \leq C ((nT)^{-\delta} + \beta_{n,T}).$$

Puisque

$$\rho_1 > \frac{dq' + 3}{2} \text{ et } \rho_2 > \frac{dq' + 3}{2},$$

nous avons $\alpha_1 > 1, \alpha_2 > 1$ et la série de terme général $\beta_{n,T}$ est convergente. Le Théorème 3.3.2 est alors une conséquence du lemme de Borel-Cantelli en choisissant C^* tel que

$$\delta > 1 \text{ et } C^* > 2 \left(1 + \sup_{x \in D} f(x) \frac{B_d^{-1}}{2} \int_{\mathbb{R}^d} K^2(u) du \right).$$

■

Chapitre 4

Estimation de la régression spatiale par une approche récursive

Sommaire

4.1 Cadre du travail	66
4.1.1 Notations et hypothèses	70
4.1.2 Preuve du Théorème 4.1.1	71

Une des applications de la statistique spatiale est de produire des prévisions pour des processus spatiaux. On parle donc de prévision spatiale, lorsque les données étudiées sont collectées à différents sites et l'influence de ces derniers sur le processus générant les données est manifeste. Un des modèles les plus utilisés pour la prévision est la régression. Le modèle de régression considéré ici est une relation statistique qui décrit le lien entre deux variables aléatoires : une variable réponse $Y \in \mathbb{R}$ et un régresseur $X \in \mathbb{R}^d$ ($d \geq 1$). Mathématiquement, on écrit

$$Y = r(X) + \varepsilon$$

où r est une fonction de \mathbb{R}^d à valeurs dans \mathbb{R} et ε une variable d'erreur. L'estimation de la régression est définie comme la construction d'un estimateur de la fonction de régression $r(\cdot)$ à partir d'un échantillon de données tiré selon la loi de (X, Y) .

La communauté statistique a consacré beaucoup de travaux pour une estimation de la fonction de régression à l'aide des méthodes paramétriques, semi-paramétriques et non paramétriques pour des données univariées et multivariées. Une

estimation paramétrique de la régression est réalisée sous l'hypothèse que la fonction de régression admet une forme paramétrique connue, telle qu'un polynôme. Cependant, dans le cas d'une estimation non paramétrique, aucune hypothèse de forme n'est nécessaire. Du fait que le lien statistique entre deux variables aléatoires est totalement inconnu dans la plupart des applications, l'estimation non paramétrique de la régression trouve un nombre considérable d'applications dans de nombreux domaines. L'estimation de la fonction de régression par la méthode des noyaux est une technique non paramétrique classique pour étudier le modèle de régression. Elle a récemment attiré une attention considérable parmi les statisticiens et les économètres.

Dans le cadre spatial traité dans ce chapitre, nous considérons que statisticien dispose des réalisations du processus (X, Y) à différents lieux de l'espace et que chaque localisation d'une observation a un impact sur la valeur de celle-ci. Le modèle de régression s'écrit alors

$$Y_{\mathbf{s}} = r(X_{\mathbf{s}}) + \varepsilon_{\mathbf{s}}, \quad \mathbf{s} \in \mathcal{I}_n$$

où $\mathcal{I}_n \subset \mathbb{R}^N$ ($N \geq 1$) est un domaine spatial de cardinal n .

4.1 Cadre du travail

Soit $(X_{\mathbf{s}}, Y_{\mathbf{s}})$, $\mathbf{s} \in \mathcal{I}_n$ un processus stochastique spatial, bivarié et défini sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans \mathbb{R}^{d+1} . On suppose que pour tout $\mathbf{s} \in \mathcal{I}_n$, le couple $(X_{\mathbf{s}}, Y_{\mathbf{s}})$ admet une densité jointe $g(\cdot, \cdot)$ relativement à la mesure de Lebesgue. La fonction de régression de Y en X est définie par :

$$r(x) = \mathbb{E}(Y|X = x).$$

Notons que la connaissance du paramètre de régression r est très utile pour prédire des valeurs futures de Y sachant $X = x$.

CHAPITRE 4. ESTIMATION DE LA RÉGRESSION SPATIALE PAR UNE APPROCHE RÉCURSIVE

On peut définir une version de la fonction $r(\cdot)$ par

$$r(x) = \begin{cases} \frac{\int_{\mathbb{R}} yg(x, y)dy}{\int_{\mathbb{R}} g(x, y)dy} & \text{si } \int_{\mathbb{R}} g(x, y)dy \neq 0 \\ \mathbb{E}(Y) & \text{si } \int_{\mathbb{R}} g(x, y)dy = 0 \end{cases}. \quad (4.1)$$

Pour estimer $r(\cdot)$, on se base sur le Chapitre 2. On considère ainsi une version spatiale de l'estimateur récursif de la densité jointe $g(\cdot, \cdot)$ définie par :

$$\widehat{g}_n(x, y) = \frac{1}{\widetilde{S}_{n,\ell}} \sum_{\mathbf{s} \in \mathcal{I}_n} \frac{1}{(h_{\mathbf{s}}^d h'_{\mathbf{s}})^{\ell}} L\left(\frac{X_{\mathbf{s}} - x}{h_{\mathbf{s}}}, \frac{Y_{\mathbf{s}} - y}{h'_{\mathbf{s}}}\right), \quad (x, y) \in \mathbb{R}^{d+1} \quad (\ell \in [0, 1]), \quad (4.2)$$

où h et h' sont deux suites de fenêtres choisies dans les directions de x et y respectivement, $\widetilde{S}_{n,\ell}$ est un paramètre de normalisation défini par

$$\widetilde{S}_{n,\ell} := \sum_{\mathbf{s} \in \mathcal{I}_n} (h_{\mathbf{s}}^d h'_{\mathbf{s}})^{1-\ell}$$

et L est une fonction de densité définie dans \mathbb{R}^{d+1} à valeurs dans \mathbb{R} telle que

$$\int_{\mathbb{R}^d} L(u, v)du = H(v); \quad \int_{\mathbb{R}} L(u, v)dv = K(u) \quad \text{et} \quad \int_{\mathbb{R}^d} uL(u, v)du = \int_{\mathbb{R}} vL(u, v)dv = 0,$$

avec H et K deux noyaux vérifiant les conditions du Chapitre 2.

Ensuite, on intègre par rapport à la variable y et moyennant le changement de variables $v = \frac{Y_{\mathbf{s}} - y}{h'_{\mathbf{s}}}$ pour obtenir une estimation de (4.1), à savoir pour le numérateur :

$$\begin{aligned} \int_{\mathbb{R}} \widehat{g}_n(x, y)dy &= \frac{1}{\widetilde{S}_{n,\ell}} \sum_{\mathbf{s} \in \mathcal{I}_n} \frac{1}{(h_{\mathbf{s}}^d h'_{\mathbf{s}})^{\ell}} \int_{\mathbb{R}} L\left(\frac{X_{\mathbf{s}} - x}{h_{\mathbf{s}}}, \frac{Y_{\mathbf{s}} - y}{h'_{\mathbf{s}}}\right) dy \\ &= \frac{1}{\widetilde{S}_{n,\ell}} \sum_{\mathbf{s} \in \mathcal{I}_n} \frac{h_{\mathbf{s}}'^{1-\ell}}{h_{\mathbf{s}}^{d\ell}} K\left(\frac{X_{\mathbf{s}} - x}{h_{\mathbf{s}}}\right). \end{aligned} \quad (4.3)$$

Pour ce qui est du numérateur dans (4.1), on a :

$$\begin{aligned}
\int_{\mathbb{R}} y \widehat{g}_n(x, y) dy &= \frac{1}{\widetilde{S}_{n,\ell}} \sum_{\mathbf{s} \in \mathcal{I}_n} \frac{1}{(h_{\widehat{\mathbf{s}}}^d h'_{\widehat{\mathbf{s}}})^\ell} \int_{\mathbb{R}} y L\left(\frac{X_{\mathbf{s}} - x}{h_{\widehat{\mathbf{s}}}}, \frac{Y_{\mathbf{s}} - y}{h'_{\widehat{\mathbf{s}}}}\right) dy \\
&= \frac{1}{\widetilde{S}_{n,\ell}} \sum_{\mathbf{s} \in \mathcal{I}_n} \frac{h_{\widehat{\mathbf{s}}}^{1-\ell}}{h_{\widehat{\mathbf{s}}}^{d\ell}} \int_{\mathbb{R}} (Y_{\mathbf{s}} - h'_{\widehat{\mathbf{s}}} z) L\left(\frac{X_{\mathbf{s}} - x}{h_{\widehat{\mathbf{s}}}}, z\right) dz \\
&= \frac{1}{\widetilde{S}_{n,\ell}} \sum_{\mathbf{s} \in \mathcal{I}_n} \frac{h_{\widehat{\mathbf{s}}}^{1-\ell}}{h_{\widehat{\mathbf{s}}}^{d\ell}} Y_{\mathbf{s}} K\left(\frac{X_{\mathbf{s}} - x}{h_{\widehat{\mathbf{s}}}}\right). \tag{4.4}
\end{aligned}$$

Pour $x \in \mathbb{R}^d$ fixé, notre objectif est d'estimer $r(x)$, une quantité qui ne dépend pas de y . Il est donc tout à fait naturel de choisir dans l'équation (4.2) la même fenêtre dans la direction de y . Autrement dit, l'attribution des poids aux observations, dans l'estimation de la fonction de régression, dépend uniquement de la proximité ou non à x . Ainsi, en posant $h'_{\widehat{\mathbf{s}}} = h_0 > 0$ pour tout $\mathbf{s} \in \mathcal{I}_n$, on déduit de (4.3) et (4.4) un estimateur de la fonction $r(\cdot)$ défini par

$$r_n^\ell(x) = \begin{cases} \frac{\phi_n^\ell(x)}{p_n^\ell(x)} & \text{si } p_n^\ell(x) > 0 \\ \frac{\phi_n^\ell(x)}{\bar{Y}} & \text{si } p_n^\ell(x) = 0 \end{cases}, \quad x \in \mathbb{R}^d \quad (\ell \in [0, 1]) \tag{4.5}$$

avec

$$\phi_n^\ell(x) = \frac{1}{S_{n,\ell}} \sum_{\mathbf{s} \in \mathcal{I}_n} \frac{Y_{\mathbf{s}}}{h_{\widehat{\mathbf{s}}}^{d\ell}} K\left(\frac{X_{\mathbf{s}} - x}{h_{\widehat{\mathbf{s}}}}\right),$$

$S_{n,\ell}$ et $p_n^\ell(\cdot)$ définis en (2.1) et (2.2).

Comme pour l'estimateur de la densité étudié au Chapitre 2, l'estimateur de la régression r_n^ℓ peut être déduit d'un algorithme stochastique de type Robbins et Monro (1951). En effet, soit \mathbf{s}_{n+1} un nouveau site d'observation qui vient s'ajouter à \mathcal{I}_n , à savoir :

$$\mathcal{I}_{n+1} = \mathcal{I}_n \cup \{\mathbf{s}_{n+1}\}.$$

Si l'on note $(X_{\mathbf{s}_{n+1}}, Y_{\mathbf{s}_{n+1}})$ la nouvelle observation du processus collectée au site \mathbf{s}_{n+1} , alors l'estimateur r_n^ℓ peut s'exprimer de manière autorégressive par le biais de la formule suivante :

$$r_{n+1}^\ell(x) = r_n^\ell(x) + \rho_{n+1} \left[Y_{\mathbf{s}_{n+1}} - r_n^\ell(x) \right] \tag{4.6}$$

CHAPITRE 4. ESTIMATION DE LA RÉGRESSION SPATIALE PAR UNE APPROCHE RÉCURSIVE

où

$$\rho_{n+1} = \Delta_{n+1}(x) \left[h_{\mathbf{s}_{n+1}}^{dl} \sum_{i=1}^{n+1} \frac{\Delta_i(x)}{h_{\mathbf{s}_i}^{dl}} \right]^{-1} \text{ avec } \Delta_i(x) = K \left(\frac{X_{\mathbf{s}_i} - x}{h_{\mathbf{s}_i}} \right), \quad i = 1, \dots, n.$$

En effet, on peut exprimer de manière récursive l'estimateur ϕ_n^ℓ par :

$$\phi_{n+1}^\ell(x) = \frac{S_{n,\ell}}{S_{n+1,\ell}} \phi_n^\ell(x) + \frac{Y_{\mathbf{s}_{n+1}}}{S_{n+1,\ell} h_{\hat{\mathbf{s}}_{n+1}}^{dl}} K \left(\frac{X_{\mathbf{s}_{n+1}} - x}{h_{\hat{\mathbf{s}}_{n+1}}} \right).$$

Or pour tout $x \in \mathbb{R}^d$ tel que $p_n^\ell(x) > 0$, on a

$$\phi_n^\ell(x) = p_n^\ell(x) r_n^\ell(x),$$

ce qui nous donne :

$$r_{n+1}^\ell(x) = \frac{S_{n,\ell} p_n^\ell(x)}{S_{n+1,\ell} p_{n+1}^\ell(x)} r_n^\ell(x) + \frac{Y_{\mathbf{s}_{n+1}}}{S_{n+1,\ell} p_{n+1}^\ell(x) h_{\hat{\mathbf{s}}_{n+1}}^{dl}} K \left(\frac{X_{\mathbf{s}_{n+1}} - x}{h_{\hat{\mathbf{s}}_{n+1}}} \right).$$

Posons

$$\frac{S_{n,\ell} p_n^\ell(x)}{S_{n+1,\ell} p_{n+1}^\ell(x)} := 1 - \rho_{n+1}.$$

Il s'en suit que :

$$\begin{aligned} \rho_{n+1} &= 1 - \frac{S_{n,\ell} p_n^\ell(x)}{S_{n+1,\ell} p_{n+1}^\ell(x)} \\ &= \frac{S_{n+1,\ell} p_{n+1}^\ell(x) - S_{n,\ell} p_n^\ell(x)}{S_{n+1,\ell} p_{n+1}^\ell(x)}. \end{aligned}$$

Grâce à la forme récursive de p_n^ℓ , on obtient :

$$\begin{aligned} \rho_{n+1} &= \frac{1}{p_{n+1}^\ell(x) S_{n+1,\ell} h_{\hat{\mathbf{s}}_{n+1}}^{dl}} K \left(\frac{X_{\mathbf{s}_{n+1}} - x}{h_{\hat{\mathbf{s}}_{n+1}}} \right) \\ &= \Delta_{n+1}(x) \left[h_{\mathbf{s}_{n+1}}^{dl} \sum_{i=1}^{n+1} \frac{\Delta_i(x)}{h_{\mathbf{s}_i}^{dl}} \right]^{-1} \end{aligned}$$

Ainsi, on peut exprimer de manière autorégressive r_n^ℓ par :

$$\begin{aligned} r_{n+1}^\ell(x) &= (1 - \rho_{n+1}) r_n^\ell(x) + \rho_{n+1} Y_{\mathbf{s}_{n+1}} \\ &= r_n^\ell(x) + \rho_{n+1} (Y_{\mathbf{s}_{n+1}} - r_n^\ell(x)). \end{aligned}$$

4.1.1 Notations et hypothèses

Soient

$$p(x) = \int_{\mathbb{R}} g(x, y) dy \text{ et } \varphi(x) = \int_{\mathbb{R}} yg(x, y) dy.$$

Dans la suite de ce chapitre, on considérera seulement les $x \in \mathbb{R}^d$ tels que $f(x) > 0$. Dans ce cadre, la fonction $r(x)$ est estimée par la famille d'estimateurs (4.5). En plus des hypothèses faites au Chapitre 2, nous introduisons des hypothèses supplémentaires pour obtenir le comportement asymptotique de notre famille d'estimateurs de la régression.

Hypothèse H. 4.1.1 *Le processus spatial $(X_{\mathbf{s}}, Y_{\mathbf{s}})_{\mathbf{s} \in \mathcal{I}_n}$ est strictement stationnaire et α -mélangeant.*

Hypothèse H. 4.1.2 *Il existe un réel $q > 2$ tel que pour tous $\mathbf{s} \in \mathcal{I}_n$ et $k \in \mathbb{Z}$ on ait :*

$$\sum_{i=k}^{\infty} \alpha(i)^{1-2/q} = O(k^{-N})$$

et la fonction $u \mapsto \mathbb{E}(|Y|^q | X = u) f(u)$ est continue en x et intégrable sur \mathbb{R}^d .

Hypothèse H. 4.1.3 *Pour tous $y, y' \in \mathbb{R}^d$ il existe $C > 0$ tel que*

$$|r(y) - r(y')| \leq C \|y - y'\|.$$

Hypothèse H. 4.1.4 *Pour tous $\mathbf{s} \neq \mathbf{s}'$, la densité jointe de $(X_{\mathbf{s}}, X_{\mathbf{s}'})$ est uniformément bornée.*

Hypothèse H. 4.1.5 *La fonction noyau K est bornée et positive et il existe $\beta > 0$ tel que*

$$\int_{\mathbb{R}^d} \|z\|^\beta K(z) dz < \infty.$$

Théorème 4.1.1 *Sous les hypothèses H 4.1.1- H 4.1.5 et celles du Théorème 2.2.3, on a*

$$r_n^\ell(x) - r(x) = O_{\mathbb{P}} \left(\sqrt{\frac{1}{nh_n^d}} \right) + O_{\mathbb{P}} (h_n^\beta)$$

lorsque $n \rightarrow \infty$ pour tout x point de continuité de la fonction $u \mapsto \mathbb{E}(Y^2 | X = u)$ tel que $p(x) > 0$.

4.1.2 Preuve du Théorème 4.1.1

Pour tout $x \in \mathbb{R}^d$ on a :

$$r_n^\ell(x) - r(x) = \frac{\phi_n^\ell(x) - r(x)p_n^\ell(x)}{p_n^\ell(x)} = \frac{G_{1n}(x)}{p_n^\ell(x)} + \frac{G_{2n}(x)}{p_n^\ell(x)} \quad (4.7)$$

avec

$$G_{1n}(x) = \frac{\sum_{\mathbf{i} \in \mathcal{I}_n} \epsilon_{\mathbf{i}} K_{\mathbf{i}}}{S_{n,\ell}} \quad \text{et} \quad G_{2n}(x) = \frac{\sum_{\mathbf{i} \in \mathcal{I}_n} (r(X_{\mathbf{i}}) - r(x)) K_{\mathbf{i}}}{S_{n,\ell}},$$

où :

$$\epsilon_{\mathbf{i}} = Y_{\mathbf{i}} - r(X_{\mathbf{i}}), \quad K_{\mathbf{i}} = \frac{1}{h_{\mathbf{i}}^{d\ell}} K\left(\frac{X_{\mathbf{i}} - x}{h_{\mathbf{i}}}\right)$$

D'une part on a :

$$\begin{aligned} \mathbb{E}(G_{1n}^2(x)) &= S_{n,\ell}^{-2} \left(\sum_{\mathbf{i} \in \mathcal{I}_n} \mathbb{E}(\epsilon_{\mathbf{i}}^2 K_{\mathbf{i}}^2) + \sum_{\mathbf{i}, \mathbf{j} \in A_1} \mathbb{E}(\epsilon_{\mathbf{i}} K_{\mathbf{i}} \epsilon_{\mathbf{j}} K_{\mathbf{j}}) + \sum_{\mathbf{i}, \mathbf{j} \in A_2} \mathbb{E}(\epsilon_{\mathbf{i}} K_{\mathbf{i}} \epsilon_{\mathbf{j}} K_{\mathbf{j}}) \right) \\ &:= G'_{1n} + G''_{1n} + G'''_{1n}. \end{aligned}$$

avec :

$$A_1 = \{\mathbf{i}, \mathbf{j} \in \mathcal{I}_n : 0 < \|\mathbf{i} - \mathbf{j}\| \leq c_n\}$$

et

$$A_2 = \{\mathbf{i}, \mathbf{j} \in \mathcal{I}_n : \|\mathbf{i} - \mathbf{j}\| > c_n\}.$$

En utilisant l'hypothèse de stationnarité, et en appliquant le Lemme de Bochner et l'inégalité de Markov, on obtient :

$$\begin{aligned} G'_{1n} &= S_{n,\ell}^{-2} \sum_{\mathbf{i} \in \mathcal{I}_n} h_{\mathbf{i}}^{-2d\ell} \int_{\mathbb{R}^d} K^2\left(\frac{x-u}{h_{\mathbf{i}}}\right) (\mathbb{E}((Y_{\mathbf{i}} - r(X_{\mathbf{i}}))^2 | X = u)) p(u) du \\ &= S_{n,\ell}^{-2} \sum_{i=1}^n h_i^{-2d\ell} \int_{\mathbb{R}^d} K^2\left(\frac{X - u}{h_i}\right) \mathbb{V}(Y_i | X_i = u) p(u) du \\ &= O\left(\frac{1}{nh_n^d}\right) \end{aligned} \quad (4.8)$$

Maintenant grâce à l'inégalité de Holder et du fait que la fonction K est bornée, on en déduit que pour tous $\mathbf{i}, \mathbf{j} \in \mathcal{I}_n$ tels que $\mathbf{i} \neq \mathbf{j}$ on a :

$$\begin{aligned} |\mathbb{E}(\epsilon_i K_i \epsilon_j K_j)| &\leq \mathbb{E}^{1/q}(|\epsilon_i|^q K_i K_j) \mathbb{E}^{1/q}(|\epsilon_j|^q K_i K_j) \mathbb{E}^{1-2/q}(K_i K_j) \\ &\leq \|K\|_\infty^{2/q} \mathbb{E}^{1/q}(|\epsilon_i|^q K_i) \mathbb{E}^{1/q}(|\epsilon_j|^q K_j) \mathbb{E}^{1-2/q}(K_i K_j) \end{aligned} \quad (4.9)$$

On a par ailleurs grâce à l'inégalité c.r. :

$$\begin{aligned} \mathbb{E}(|\epsilon_i|^q K_i) &\leq C h_i^{-d\ell} \int_{\mathbb{R}^d} K \left(\frac{X-u}{h_i} \right) (\mathbb{E}(|Y|^q | X=u) + |r(x)|^q) f(u) du \\ &= O\left(h_i^{d(1-\ell)}\right) \end{aligned}$$

Par conséquent :

$$|\mathbb{E}(\epsilon_i K_i \epsilon_j K_j)| = O\left(h_i^{d(1-\ell)(1-1/q)} h_j^{d(1-\ell)(1-1/q)}\right).$$

Ainsi :

$$\begin{aligned} G''_{1n} &\leq C S_{n,\ell}^{-2} \sum_{\mathbf{i} \in \mathcal{I}_n} \sum_{\mathbf{j} \in \mathcal{I}_n: \|\mathbf{i}-\mathbf{j}\| \leq c_n} h_i^{d(1-\ell)(1-1/q)} h_j^{d(1-\ell)(1-1/q)} \\ &\leq C S_{n,\ell}^{-2} c_n^N \sum_{i=1}^n h_i^{2d(1-\ell)(1-1/q)} \\ &= O\left(\frac{c_n^N}{n h_n^{2d/q}}\right). \end{aligned} \quad (4.10)$$

D'autre part, en utilisant le Lemme 2.1, p.39 dans [Tran \(1990\)](#), on a :

$$\begin{aligned} |\mathbb{E}(\epsilon_i K_i \epsilon_j K_j)| &\leq (\mathbb{E}|\epsilon_i K_i|^q)^{1/q} (\mathbb{E}|\epsilon_j K_j|^q)^{1/q} \alpha(\|\mathbf{i}-\mathbf{j}\|)^{1-2/q} \\ &= O\left(h_i^{d(1-\ell)/q} h_j^{d(1-\ell)/q} \alpha^{1-2/q}(\|\mathbf{i}-\mathbf{j}\|)\right). \end{aligned} \quad (4.11)$$

On en déduit que :

$$\begin{aligned} G'''_{1n} &= O\left(S_{n,\ell}^{-2} \sum_{i=1}^n h_i^{2d(1-\ell)/q} \sum_{c_n \leq k \leq n} \alpha^{1-2/q}(k)\right) \\ &\leq O\left(\frac{c_n^{-N}}{n h_n^{2d(1-\ell)(1-1/q)}}\right) \end{aligned} \quad (4.12)$$

CHAPITRE 4. ESTIMATION DE LA RÉGRESSION SPATIALE PAR UNE APPROCHE RÉCURSIVE

Avec le choix $c_n^N = O\left(h_n^{d((1-\ell)(1-1/q)-1/q)}\right)$ et les relations (4.10) et (4.12) on obtient :

$$G_{1n}'' + G_{1n}''' = O\left(\frac{1}{nh_n^{d(1-\ell(1-1/q))}}\right).$$

Cette dernière relation combinée avec (4.8) nous permet d'écrire :

$$\mathbb{E}(G_{1n}(x))^2 = O\left(\frac{1}{nh_n^{2d}}\right) \quad (4.13)$$

En utilisant les hypothèses H 4.1.3 et H 4.1.5, on obtient :

$$\begin{aligned} |G_{2n}| &\leq Ch_n^\beta S_{n,\ell}^{-1} \sum_{i=1}^n \left\| \frac{X_i - x}{h_n} \right\|^\beta \|K\|_\infty \\ &= O(h_n^\beta). \end{aligned} \quad (4.14)$$

Le résultat du Théorème 4.1.1 suit grâce à (4.1.2) et (4.14) et le Théorème 2.2.3.

■

Chapitre 5

Conclusion générale et perspectives

Sommaire

5.1 Conclusion générale	75
5.2 Perspectives	76

5.1 Conclusion générale

Dans ce mémoire de thèse, nous avons considéré l'estimation récursive non paramétrique dans le cas de flux de données spatiales. Plus particulièrement, nous avons considéré deux cas de figures, à savoir des flux spatiaux et des flux spatio-temporels. Nous avons estimé les fonctions de densité et de régression. Dans le premier chapitre de cette thèse, nous donnons une estimation de la fonction de densité de probabilité d'une variable aléatoire spatiale multivariée par noyaux récursifs qui généralise la méthode de [Parzen \(1962\)](#). Nous avons ainsi étendu les travaux de [Amiri \(2010\)](#) au cas de données spatiales dépendantes. Nous avons donné des résultats asymptotiques d'un tel estimateur ainsi qu'une application à l'estimation d'une fonction de risque.

Le chapitre 2 traite et analyse des flux des données, *Data Stream*. Nous nous sommes intéressés à l'estimation non-paramétrique de la densité de probabilité d'un "Data Stream" de nature spatio-temporelle. Nous avons proposé et étudié un estimateur à noyau de la fonction de densité de probabilité d'un flux de données

spatio-temporelles. Nous avons combiné la méthodologie sur l'estimateur récursif de la densité et celle d'une distribution de nature spatiale (Tran (1990)).

Le dernier chapitre vient contribuer à la régression spatiale. Dans ce chapitre, nous nous sommes intéressés à l'estimation de la fonction de régression spatiale par le biais des méthodes non paramétriques. Nous avons ainsi étendu les travaux effectués dans le chapitre 2 au cadre de la régression et donné des résultats asymptotiques de l'estimateur proposé.

Les illustrations numériques montrent que les estimateurs proposés sont performants et sont comme dans le cadre non-spatial des alternatives aux méthodes d'estimation à noyau non-récursifs.

À travers ce travail de thèse, nous venons contribuer à la statistique spatiale non-paramétrique aussi bien dans un cadre théorique mais aussi appliqué. Tout au long de ce travail, des questions et remarques sont apparues laissant place à quelques perspectives de recherche que nous développons ci-après.

5.2 Perspectives

Tout d'abord, une étude de simulations et une application à la prévision et à des données réelles seront faites pour étudier les performances numériques de l'estimateur de la régression étudié au Chapitre 4. D'autres types de convergences, notamment la normalité asymptotique seront considérées.

Nous nous sommes limités dans ce travail au cadre d'un design spatial régulier, il serait judicieux de considérer le cadre irrégulier.

Deux types d'asymptotique ont été considérés dans le Chapitre 3. Il est naturel de regarder le cas où l'un des domaines est fixé (par exemple la région est fixée) et l'autre évolue. Le cadre localement stationnaire est aussi à envisager.

Nous avons considéré dans cette contribution uniquement le cadre géostatistique où les sites sont dans \mathbb{N}^N , une extension possible serait de considérer un champs aléatoire indexé de manière continue, des données latentes ou les processus ponctuels spatiaux dont le cadre d'étude serait intéressant pour l'estimation de la fonction de risque étudiée au Chapitre 2.

Les travaux effectués dans ce mémoire peuvent également être étendus aux flux de données spatiotemporels avec des tailles de fenêtres aléatoires. De même il sera

CHAPITRE 5. CONCLUSION GÉNÉRALE ET PERSPECTIVES

plus intéressant de considérer le cas où la distribution sous-jacente évolue avec le temps et l'espace qui constitue un cadre la plus réaliste dans le contexte des *data stream*.

Bibliographie

- Abramson, I., (1982). On bandwidth variation in kernel estimates. *The Annals of Statistics*, 10 : 1217–1223.
- Ahmad, I. , & Lin, P.E. (1976). Nonparametric sequential estimation of a multiple regression function. *Bulletin of Mathematics and Statistics*, 17 : 63–75.
- Amiri, A., (2010). Sur une famille paramétrique d’estimateurs séquentiels de la densité pour un processus fortement mélangeant. *Comptes Rendus de l’Académie des Sciences de Paris, Serie I*, 347 : 309–314.
- Amiri, A., (2012). Recursive regression estimators with application to nonparametric prediction. *Journal of Nonparametric Statistics*, 24 : 169–186.
- Amiri, A., (2013). Asymptotic normality of recursive estimators under strong mixing conditions. *Statistical Inference for Stochastic Processes*, 16 (2) : 181–96.
- Amiri, A., Dabo-Niang, S., & Yahaya, M., (2015). Nonparametric recursive density estimation for spatial data. *Comptes Rendus de l’Académie des Sciences de Paris, Serie I*, 354 : 205–210.
- Anselin, L., & Florax, R J M., (1995). New directions in spatial econometrics. *Springer, Berlin*.
- Biau, G., (2003). Spatial Kernel Density Estimation. *Mathematical Methods of Statistics*, 12 : 371–390.

BIBLIOGRAPHIE

- Biau, G., & Cadre, B., (2004). Nonparametric Spatial Prediction. *Statistical Inference for Stochastic Processes*, 7 : 327–349.
- Bithell JF., (1990). An application of density estimation to geographical epidemiology. *Statistics in Medicine*, 9 : 691-701.
- Blanke, D., & Bosq, D., (2007). Inference and prediction in large dimensions. *Wiley series in probability and statistics*, Dunod.
- Bosq, D., & Cheze-Payaud, N., (1999). Optimal asymptotic quadratic error of non-parametric regression function estimates for a continuous-time process from sampled data. *Statistics* , 32 (3) : 229–247.
- Carbon, M., Tran, L.-T, & Wu, B., (1997). Kernel density estimation for random fields : the L_1 theory. *Nonparametric statistics*, 6 : 157–170.
- Carbon, M., Tran, L T., & Wu, B., (1997). Kernel density estimation for random fields. *Statistics and Probability Letters*, 36 : 115–125.
- Carbon, M., Francq, C., & Tran, L. T., (2007). Kernel regression estimation for random fields. *Journal of Statistical Planning and Inference*, 137 : 778–798.
- Carbon, M., Francq, C., & Tran, L T., (2007). Kernel regression estimation for random fields. *Journal of statistical planning and inference*, 137 : 778–798.
- Cao, Y., He, H., & Man, H., (2012). “SOMKE : Kernel density estimation over data streams by sequences of selforganizing maps”. *IEEE TNNLS*, 23 (8) : 1254–1268.
- Chiles, J.-P, & Delfiner, P., (1999). Statistics for spatial data, *Wiley. New York*.
- Christakos, G., (2000). Modern Spatiotemporal Geostatistics, .
- Conley, T.,(1999). GMM Estimation with Cross Sectional Dependence. *Journal of Econometrics*, 92 : 1–45.
- Crambes, C., Kneip, A., & Sarda, P., (2009). Smoothing splines estimators for functional linear regression, *Ann. Statist*, 37 : 35-72.

BIBLIOGRAPHIE

- Cressie, N., (1991). Statistics for spatial Data, *Wiley Series in Probability and Mathematical Statistics, New-York*.
- Cressie, N., (1993). Statistics for spatial data. *John Wiley and Sons, New York, 1–900*.
- Cressie, N., & Wikle, C K., (2011). Statistics for Spatio-Temporal Data, *John Wiley & Sons*.
- Dabo-Niang, S., & Yao, A.-F. (2007). Kernel regression estimation for continuous spatial processes. *Mathematical Methods of Statistics, 16 : 298–317*.
- Dabo-Niang, S., & Yao, A.-F., (2007). Kernel regression estimation for continuous spatial processes. *Mathematical Methods of Statistics, 16 : 298–317*.
- Dabo-Niang, S., & Thiam, B. (2010). Robust quantile estimation and prediction for spatial processes. *Statistics and Probability Letters, 80 : 1447–1458*.
- Dabo-Niang, S., Rachdi, M., & Yao, A F., (2011). Kernel regression estimation for spatial functional random variables. *Far East Journal of Theoretical Statistics, 37 : 77–113*.
- Dabo-Niang, S., & Yao, A.-F. (2013). Kernel spatial density estimation in infinite dimension space. *Metrika, 76 : 19–52*.
- Dabo-Niang, S., & Hamdad, L., & Ternynck, C., & Yao, A F.,(2014). Kernel spatial density estimation with applications to spatial clustering and Monsoon Asia Drought Atlas analysis, *Stochastic Environmental Research and Risk Assessment, Springer Verlag (Germany) : 2075–2099*.
- Dabo-Niang, S., Hamdad, L., & Ternynck, C.,& Yao, A.-F. A kernel spatial density estimation allowing for the analysis of spatial clustering : application to Mousoon Asia Drought Atlas data. *Stochastic Environmental Research and Risk Assessment, 28 : 2075–2099*.
- Dabo-Niang, S., Ould-Abdi, S. A., Ould-Abdi, A., & Diop, A. (2014). Consistency of a nonparametric conditional mode estimator for random fields. *Statistical Methods & Applications, 23 : 1–39*.

BIBLIOGRAPHIE

- Davies, H. I., (1973). Strong consistency of a sequential estimator of a probability density function. *Bulletin of Mathematical Statistics*, 15 : 49–54.
- Davies, T M., & Hazelton, M L., (2010). Adaptive kernel estimation of spatial relative risk. *Statistics in Medicine*, 29 : 2423–2437.
- Davies, H I., & Wegman, E J., (1979). Remarks on Some Recursive Estimators of a Probability Density, *Annal of Statistics*, 7 : 316–327.
- Dedecker, J., Doukhan, P., & Gabriel, L., & Léon, J-R., & Louichi, S., & Prieur, C., (2007). “Mixing : Properties and examples”. *Lecture notes in statistics*, Springer-Verlag, New York.
- Dedecker, J., & Prieur, C., (2004). Coupling fort τ -Dependent Sequences and Applications. *Journal of Theoretical Probability*, 17 : 861-885.
- Deheuvels, P., (1974). Sur une famille d’estimateurs de la densité d’une variable aléatoire. *Comptes Rendus de l’Académie des Sciences de Paris*, 276 : 1013–1015.
- Devroye, L., & Wagner, T. J., (1980). On the L^1 convergence of kernel estimators of regression functions with application in discrimination. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 51 : 15–25.
- Delsol, L., (2009). Advances on asymptotic normality in nonparametric functional time series analysis. *Statistics*, 43(1) : 13–33.
- Doukhan, P., & Louhichi, S., (1999). A new weak dependance condition and application moment inequalities. *Stochastic Processes and their Applications*, 84 : 313–342.
- Doukhan, P., & Louhichi, S., (2001). Functional estimation of a density under a new weak dependence condition *Scandinavian Journal of Statistics*, 28 : 325–341.
- El Machkouri, M., (2007). Nonparametric regression estimation for random fields in a fixed-design. *Statistical Inference for Stochastic Process*, 10 : 29–47.

BIBLIOGRAPHIE

- El Machkouri, M., & Stoica, R., (2010). Asymptotic normality of kernel estimates in a regression model for random fields. *Journal of Nonparametric Statistics*, 22 : 955–971.
- El Machkouri, M.,(2014). Kernel Density Estimation For Stationary Random Rields the L_1 theory. *hal-00622861*, 5 : 1–25.
- Fazekas, I., & Chuprunov, A., (2006). Asymptotic normality of kernel type density estimators for random fields. *Statistical Inference for Stochastic Processes*, 9 : 161–178.
- Ferraty, F., & Vieu, P., (2000). Dimension fractale et estimation de la régression dans desespaces vectoriels semi-normés. *Comptes Rendus Mathématique, Académie des Sciences de Paris*, 330 : 139—142.
- Ferraty, F., & Vieu, P., (2006). Nonparametric modelling for functional data. methods, theory, applications and implementations. *Springer-Verlag, London*.
- Ferraty, F., Mas, A., & Vieu, P., (2007). Nonparametric regression on functional data : Inference and practical aspects. *Australian & New Zealand Journal of Statistics*, 49(3) : 267—286.
- Francisco-Fernandez, M., & Opsomer, J.D., (2005). Smoothing parameter selection methods for nonparametric with spatially correlated errors. *The Canadian Journal of Statistics*, 33 : 279–295.
- Gaetan, C., & Guyon, X., (2008). *Modélisation et statistique spatiales*. New York : Springer.
- García-Soidán, P., & Menezes, R. (2012). Estimation of the spatial distribution through the kernel indicator variogram. *Environmetrics*, 23 : 535–548.
- Gheriballah, A., Laksaci, A., & Rouane, R., (2010). Robust nonparametric estimation for spatial regression. *Journal of Statistical Planning and Inference*, 140 : 1656–1670.
- Györfi, L., Khöler, M., & Krzyżak, A.,& Walk, H., (2002). A distribution-free theory of nonparametric regression, *Springer-Verlag, New york*.

BIBLIOGRAPHIE

- Guyon, X., (1995). Random fields on a network : modeling, Statistics and applications. *Springer, New York*.
- Hall, P., & Marron, JS., (1988). Variable window width kernel density estimates of probability densities, *Prob. Th. Rel. Fields*, 80, 37–49.
- Hall, P., Müller, H., & Wu, P., (2005). Time-dynamic density and mode estimation with applications to fast mode tracking. *Journal of Computational and Graphical Statistics*, 15 : 82–100.
- Hallin, M., Lu, Z., & Tran, L. T., (2001). Density estimation for spatial linear processes. *Bernoulli*, 7 : 657–668.
- Hallin, M., Lu, Z., & Tran, L. T., (2004). Local linear spatial regression. *The Annals of Statistics*, 32, 2469–2500.
- Hallin, M., Lu, Z., & Tran, L. T., (2004). Local linear spatial regression. *Annals of Statistics*, 32 : 2469–2500.
- Hallin, M., Lu, Z., & Yu, K., (2009). Local linear spatial quantile regression. *Bernoulli*, 15 : 659–686.
- Hwang, E., & Shin, D W., (2014). Kernel estimators of mode under ψ -weak dependence. *Annals of the Institute of Statist.*
- Isogai, E., (1994). A Berry-Esseen-type bound for recursive estimators of a density and its derivatives. *Journal of Statistical Planning and Inference*, 40 : 1-4.
- Karácsony, Z., & Filzmoser, P. (2010). Asymptotic normality of kernel type regression estimators for random fields. *Journal of Statistical Planning and Inference*, 140 : 872–886.
- Kelsall, JE., & Diggle, PJ., (1995). Kernel estimation of relative risk, *Bernoulli*, 1 : 3–16.
- Krzyżak, A., (1992). “Global convergence of the recursive kernel regression estimates with applications in classification and nonlinear system estimation”. *IEEE Transactions on Information Theory*, 38 : 1323–1338.

BIBLIOGRAPHIE

- Li, J., & Tran, L. T. (2009). Nonparametric estimation of conditional expectation. *Journal of Statistical Planning and Inference*, 139 : 164–175.
- Liang, H Y., & Baek, J., (2004). Asymptotic normality of recursive density estimates under some dependence assumptions. *Metrika*, 60 : 155-166.
- Ling, N., & Wu, Y., (2012). Consistency of modified kernel regression estimation for functional data. *Statistics*, 46(2) : 149—158.
- Lu, Z.-d., & Chen, X. (2002). Spatial nonparametric regression estimation : Non-isotropic case. *Acta Mathematicae Applicatae Sinica (English Series)*, 18 : 641–656.
- Masry, E., (1986). Recursive probability density estimation for weakly dependent stationary processes. *IEEE Trans. Inform. Theory* , 32 : 254–267.
- Masry, E., (1987). Almost sure convergence of recursive density estimators for stationary mixing processes. *Statistics & Probability Letters* , 5 : 249-254.
- Masry, E., & Györfi, L., (1987). Strong consistency and rates for recursive probability density estimators of stationary processes. *Journal of Multivariate Analysis*, 22 : 79– 93.
- Menezes, R., Garcia-Soidan, P., & Ferreira, C., (2010). Nonparametric spatial prediction under stochastic sampling design, *Journal of Nonparametric Statistics*, 22 : 363–377.
- Mezhoud, K A *et al.*, (2014). Recursive Kernel estimation of the density under η -dependence, *Journal of the Korean Statistical Society* : 1-12.
- Mugisha, R.-X, & Samanta, M., (1981). On the class of estimates of the probability function and mode based on random number of observations. *Calcutta Statistical Association Bulletin*, 117 : 23–40.
- Nadraya, E., (2010). On estimating regression. *Theory of Probability and Its Applications*, 9 : 141–142.
- Neveu, J., (1970). Bases mathématiques du calcul des probabilités. *2nd edition*, Masson et Cie, Editeurs, Paris.

BIBLIOGRAPHIE

- Ould-Abdi, A. O., Dabo-Niang, S., & Diop, A., (2010). Consistency of a nonparametric conditional quantile estimator for random fields. *Mathematical Methods of Statistics*, 19 : 1–21.
- Parzen, E., (1962). On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33 (3) : 1065–1076.
- Qin, Y.S., (1995). Asymptotic distribution of a recursive kernel estimator for a non-parametric regression function under dependent sampling. *Mathematica Applicata*, 8(1) : 7–13.
- Ripley, B., (1981). Spatial statistics, *Wiley, New York*.
- Robbins, R., & Monro, S. A., (1951). A stochastic approximation method. *Ann. Statist*, 22(3) : 400–407.
- Robison, G M., (1926). Divergent double sequences and series. *American Mathematical Society*, 28 : 50–73.
- Robinson, P.-M., (2011). Asymptotic theory for nonparametric regression with spatial data. *Journal of Econometrics*, 165 : 5–19.
- Robinson, P. M., (2011). Asymptotic theory for nonparametric regression with spatial data. *Journal of Econometrics*, 165 : 5–19.
- Rodriguez-Corté, F J., & Ghorbani M., Mateu, J., & Stoyan, D., (2014). On the expected value and variance for an estimator of the spatio-temporal product density function. *Centre For Stochastic Geometry And Advanced BioImaging*, 6 : 1–21.
- Rosenblatt, M., (1985). Stationary sequences and random fields, *Birkhauser, Boston*.
- Roussas, G.G., (1990). Nonparametric regression estimation under mixing conditions. *Stochastic Processes and their Applications*, 36 (1) : 107–116.
- Roussas, G.G., & Tran, L.T., (1992). Asymptotic normality of the recursive kernel regression estimate under dependence conditions. *Annals of Statist*, 20(1) : 98–120.

BIBLIOGRAPHIE

- Sigrist, F., Künsch, H R., & Stahel, W A., (2015). spate : An R Package for Spatio-Temporal Modeling with a Stochastic Advection-Diffusion Process. *Journal of Statistical Software*, 63 : 1–23.
- Stein, M. L., (1999). Interpolation of spatial data : some theory for kriging. *Springer*.
- Stout, W.-F., (1974). Almost sure convergence. *Academic Press, New york* : 120–121.
- Takahata, H., (1980). Almost sure convergence of density estimators for weakly dependent stationary processes. *Bulletin of Tokyo Gakugei University. Seies IV, Mathematics and Natural Sciences* : 11–32.
- Ternynck, C. (2014). Spatial regression estimation for functional data with spatial dependency. *Journal de la Société Française de Statistique*, 155 : 138–160.
- Tjøstheim, D. (1987). *Spatial series and time series : similarities and differences*.
- Tran, L T., (1990). Kernel density estimation on random fields. *Journal of Multivariate Analysis*, 34 : 37–53.
- Tran, L. T., & Yakowitz, S. (1993). Nearest neighbor estimators for random fields. *Journal of Multivariate Analysis*, 44 : 23–46.
- Wagner, T., & Wolverson, C., (1969). Recursive estimates of probability density. *IEEE Transactions on Systems Science and Cybernetics*, 5 : 307–308.
- Wang, L., & Liang, H.Y., (2004). Strong uniform convergence of the recursive regression estimator under φ -Mixing Conditions. *Metrika*, 59(3) : 245–261.
- Wang, H., & Wang, J., (2009). Estimation of the trend function for spatio-temporal models. *Journal of Nonparametric Statistics*, 21 : 567–588.
- Wang, H., Wang, J., & Huang, B., (2012). Prediction for spatio-temporal models with autoregression in errors. *Journal of Nonparametric Statistics*, 24 : 217–244.
- Watson G.S., (1964). Smooth regression analysis. *Sankhya*, 26, 359–372.

BIBLIOGRAPHIE

Wertz, W., (1985). Sequential and recursive estimators of the probability density. *Statistics, 16* : 277-295.

Yamato, H., (1972). Sequential estimation of a continuous propability density function and mode. *Bulletin of Mathematics and Statistics, 14* : 1-12.

BIBLIOGRAPHIE
