

Interstats Méthode N° 15

Les cartes de chaleur appliquées aux taux de cambriolages

Septembre 2019

Auteur : JULIEN PRAMIL

Rédacteur en chef : OLIVIER FILATRIAU

Le lissage des infractions comme outil d'analyses territoriales fines

Le lissage spatial des faits de délinquance géolocalisés permet de représenter la répartition territoriale de la délinquance enregistrée de façon structurelle, à des niveaux géographiques fins et indépendamment des frontières administratives classiques (départements, zones de compétences des services, communes, etc.). De plus, le lissage de l'information limite les problèmes de secret statistique malgré la finesse géographique des représentations produites.

Des cartes de densité spatiale des cambriolages à Paris, Lyon et Marseille ont été précédemment publiées par le SSMSI, ainsi qu'un support méthodologique d'aide à l'interprétation (voir [Interstats Méthode n°11](#) et [Interstats Info rapide n°10](#)). Comme indiqué dans ces deux publications, les cartes issues de ces opérations de lissage (appelées aussi « cartes de chaleur » ou cartes « lissées ») représentent le nombre de cambriolages constatés par kilomètre-carré, sans rapporter à la densité de population ou à la densité de logements au kilomètre-carré. Autrement dit, ces cartes permettent partiellement de répondre à la question « où sont perpétrés les cambriolages ? », mais nullement à la question « où risque-t-on le plus d'être cambriolé ? ». En effet, les zones de forte concentration des cambriolages correspondent quasi-systématiquement à des zones densément peuplées.

Des taux lissés pour approcher une mesure du risque

Pour autant, les taux de cambriolages (nombre de faits constatés pour 1 000 logements) à niveaux géographiques fins présentent un intérêt certain pour s'approcher d'une représentation des risques (voir *encadré 1*).

Une technique relativement simple existe pour calculer des taux lissés : elle consiste à réaliser un premier lissage spatial du numérateur (les cambriolages), un second avec les mêmes paramètres de lissage¹ pour le dénominateur (le nombre de logements ou la population résidente), puis à rapporter, pour chaque pixel de l'espace², la valeur lissée du numérateur sur la valeur lissée du dénominateur.

Cette technique doit néanmoins être réalisée sur des espaces suffisamment denses (en logement ou en population selon le type de taux voulu) pour aboutir à des représentations fiables et stables (voir *infra*).

Encadré 1 : Particularités des cambriolages parmi les formes de délinquance

Les cambriolages de logements présentent plusieurs particularités favorables pour la représentation sous formes de cartes de chaleur. Comme évoqué dans [l'Interstats Méthode n°11](#), le taux moyen de dépôt de plainte et la qualité moyenne des adresses de commission renseignées sont élevés relativement aux autres formes de délinquance. Ces deux propriétés des cambriolages permettent aux données de la police et de la gendarmerie de former des indicateurs plus proches de la réalité effective pour ce type de délinquance. D'autre part, les cambriolages sont commis, par définition, dans des zones habitées (sauf éventuellement dans les zones composées exclusivement de résidences secondaires). Ceci facilite et légitime la représentation des cambriolages sous forme de taux sur un territoire donné. Par exemple, il n'est actuellement pas possible de représenter des taux de vols violents dans un bois ou dans une zone industrielle faute de disposer de données de fréquentations de ces lieux dépourvus de logements et d'habitants. Pour les mêmes raisons, dans une zone habitée, rapporter le nombre de vols violents à la population résidente peut présenter des limites

¹ Même zone d'étude, même quadrillage de cette zone et même rayon de lissage.

² Plus précisément, pour chaque carreau du quadrillage de la zone d'étude.

dans les quartiers où la population effectivement présente au cours de la journée est très différente de la population résidente (par exemple, les quartiers très fréquentés des centres-villes où le nombre de résidents est très en-deçà du nombre de passants et où les taux de délinquance sont donc parfois artificiellement élevés). Contrairement à ces deux exemples, les taux de cambriolages ne sont pas concernés par ces limites pratiques et théoriques.

Délimitation des zones suffisamment denses pour calculer des ratios lissés

Le calcul de taux lissés de cambriolages consiste à rapporter, pour chaque pixel de l'espace, la valeur de lissage des cambriolages par la valeur de lissage du nombre de logements (voir *encadré 2* pour les données carroyées de logements).

Mathématiquement, ce calcul ne peut donc être réalisé que sur les zones dont la valeur de lissage du nombre de logements est strictement supérieure à zéro, c'est-à-dire, tous les pixels de l'espace se trouvant à une distance inférieure au rayon de lissage (R) d'un logement.

Une fois que les zones ayant une valeur de lissage nulle sont retirées de l'analyse, le calcul des taux lissés de cambriolages devient possible. Pour autant, le ratio obtenu n'est pas systématiquement pertinent : par exemple, certains pixels possèdent une valeur lissée du nombre de logements infime (ceux relativement éloignés des habitations par rapport à R) mais une valeur de lissage des cambriolages non nulle, aboutissant à un taux artificiellement élevé.

Face au problème causé par les trop petites valeurs de lissage du dénominateur, limiter l'analyse aux pixels ayant une valeur de lissage du dénominateur suffisante pourrait constituer une ébauche de solution. Pour cela, un seuil minimal (S) peut être défini : seuls les pixels dont la valeur de lissage du nombre de logements est supérieure à S sont représentés sur la carte de chaleur.

Néanmoins, cette alternative génère aussi des taux dont l'interprétation pose problème au niveau local. En effet, les zones non habitées (parcs, bois, zones industrielles ou commerciales, etc.) influencent à la baisse la valeur de lissage du dénominateur des zones attenantes, pourtant potentiellement denses en réalité, dans un rayon R. Elles influencent aussi à la baisse la valeur du numérateur des zones habitées attenantes (car les cambriolages de logement ne s'y produisent pas théoriquement). Dans l'autre sens, les zones habitées influencent à la hausse la valeur de lissage du numérateur et du dénominateur des zones vides, toujours dans un rayon R. Ainsi, aux abords de zones non-habitées, les taux de cambriolages calculés localement peuvent être difficiles à interpréter, voire instables et non-représentatifs de la réalité locale (voir *encadré 3*).

La solution proposée consiste à définir les zones suffisamment habitées (zSh), à les détecter et à supprimer les autres zones de l'analyse *a priori*. Les zSh peuvent être définies à l'aide d'un seuil minimal S de densité de logements par km² (par exemple 500 logements³ par km²).

Pour les détecter, une première opération de lissage de la population est réalisée avec un rayon de lissage minimal r (ici 200 mètres de côté, correspondant à la granularité de diffusion des logements) : une représentation de la densité de logements à niveau très fin est obtenue. Avec ce lissage, une grande zone non habitée (comme le Parc de Sceaux) n'influencera pas la valeur de lissage de la population des quartiers

³ Le choix d'un seuil se fait en fonction de la zone d'étude, mais aussi en vue de respecter le secret statistique en cas de diffusion d'une carte de chaleur. En effet, 500 logements par km² équivaut à 20 logements par pixel de 200 mètres de côté (maillage utilisé dans l'*encadré 3*). Le seuil minimal de 20 logements dans chaque pixel assure un bon compromis entre l'objectif de diffuser de l'information à fine échelle et limitation des risques d'identifications d'informations individuelles (d'autant que le taux diffusé est « brouillé » par le lissage et la classification).

atendants au-delà de r mètres. Ensuite, les carreaux ayant une valeur de lissage du nombre de logements supérieure à S peuvent être détectés et définis comme zSh . Ceci permet d'exclure toutes les autres zones par la suite (celles pas suffisamment denses en logements, voir *Annexe*).

Une fois les zSh déterminées à la suite de cette première opération de lissage, une seconde opération de lissage est réalisée uniquement à l'intérieur de ces zSh (pas de diffusion du lissage au-delà) et uniquement à partir des informations contenues⁴ dans ces zSh (cambriolages et nombre de logements). Ainsi, les logements sont lissés uniquement à l'intérieure des zSh , seulement à partir des pixels de logements fournis par l'Insee appartenant à ces zones. De même, les cambriolages sont uniquement lissés dans les zSh , et seulement à partir des cambriolages effectivement commis dans ces zones.

Le lissage des taux de cambriolages permet de représenter, sous forme de cartes de chaleur, le nombre de cambriolages enregistrés par les forces de sécurité et de les rapporter au nombre de logements en chaque pixel d'une zone d'étude. Par rapport au lissage simple (nombre de cambriolages au kilomètre-carré), le lissage des taux offre une représentation de la criminalité enregistrée ramenée à la densité de logements et permet une approche, même imparfaite, du risque de commission des cambriolages en chaque point de l'espace sur une période donnée. La détermination et l'exclusion préalable des zones non-habitées permettent d'assurer la pertinence et la stabilité des représentations fournies. Enfin, au-delà des cambriolages de logements, la représentation d'autres formes de délinquance nécessitera, dans la mesure du possible, une adaptation de cette méthode.

Encadré 2 : les données de logements carroyées

L'Insee fournit des données carroyées issues du dispositif *Filosofi* 2015⁵. Ces carreaux font 200 mètres de côté et ne se superposent pas nécessairement aux pixels découpés dans la grille de lissage utilisée.

Pour lisser les informations contenues dans cette base (population, logements, etc.), chaque carreau de l'Insee est assimilé à son centroïde. Les centroïdes sont au minimum espacés de 200 mètres (pour deux carreaux contigus). Ainsi, le lissage de la population ne peut être réalisé avec un rayon de lissage inférieur à 200 mètres car la granularité des données de base ne le permet pas, sauf à augmenter artificiellement la densité de logements au niveau des centroïdes des carreaux de base.

Encadré 3 : Exemple de l'influence du contrôle des zones peu habitées sur les cartes de chaleur

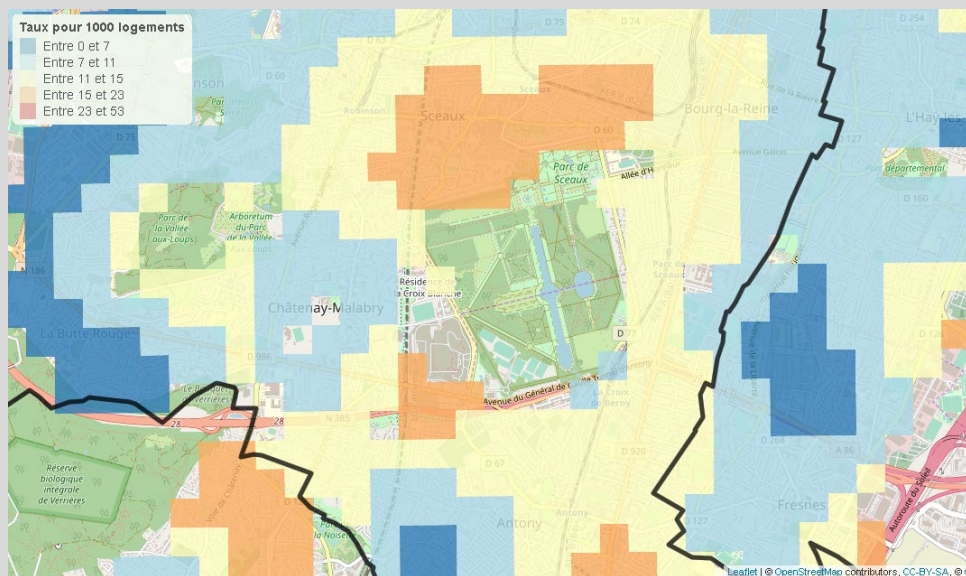
Les deux cartes ci-après permettent, sur un même territoire, de comparer la méthode avec contrôle des zSh , et la méthode sans contrôle *a priori* (simple suppression *a posteriori* des zones dont la valeur de lissage du nombre de logements est inférieure à un seuil S). Tous les paramètres de lissage et de représentation sont identiques malgré la différence d'approche (rayon de lissage de 1 000 mètres, pixels de 200 mètres de côté, mêmes données de logements et de cambriolages, seuil S de 500 logements par km², même classification en 5 groupes par la méthode des moyennes mobiles). À noter que les groupes vont du bleu (faibles taux) au rouge (forts taux).

⁴ Dans un pixel contenant des logements et des cambriolages, mais dont la densité de logements est trop faible, ces informations sont exclues du lissage.

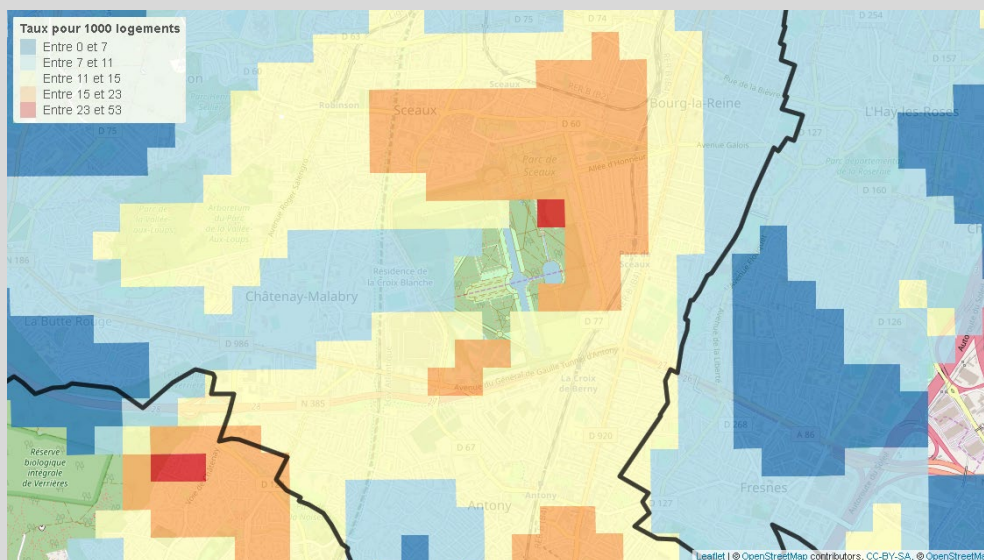
⁵ <https://www.insee.fr/fr/statistiques/4176290?sommaire=4176305#titre-bloc-1>

Les deux représentations diffèrent localement. En particulier, la méthode retenue (première carte) évite l'apparition de points chauds / *hot spots* (taches rouges) dans le voisinage de zones vides (notamment dans le Parc de Sceaux).

Carte du sud des Hauts-de-Seine avec contrôle *a priori* des zones très faiblement habitées :



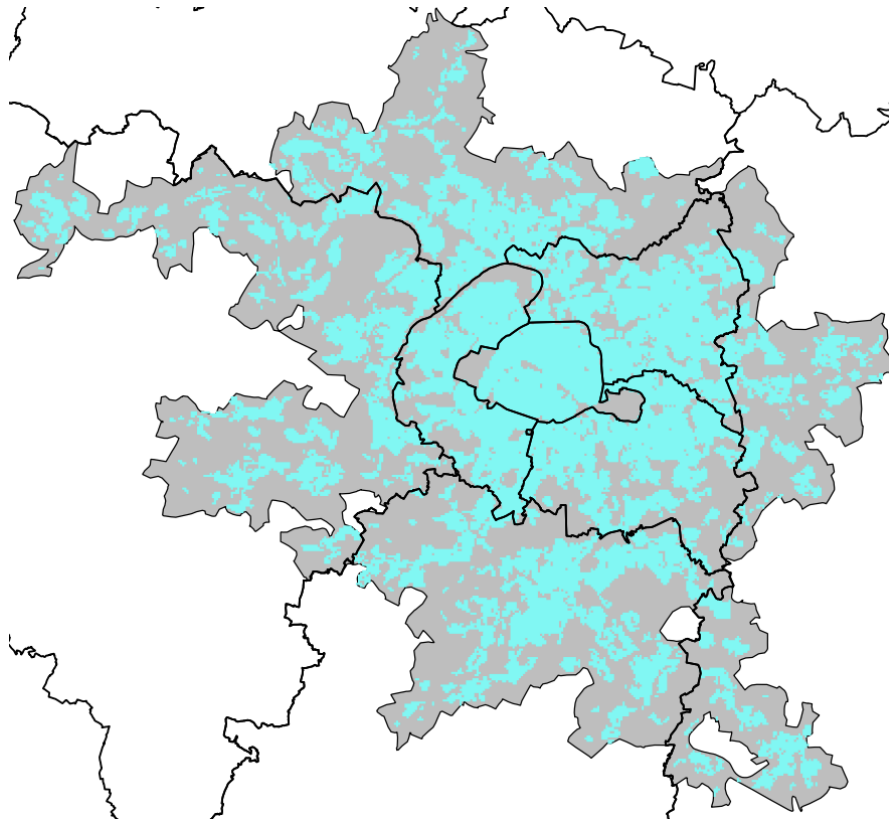
Carte du sud des Hauts-de-Seine sans contrôle *a priori* des zones très faiblement habitées :



Pour en savoir plus :

- « [Manuel d'analyse spatiale](#) », Chapitre 8 « Lissage spatial », Insee Méthodes n° 131, octobre 2018, Insee
- « [Le lissage spatial de la délinquance enregistrée et la représentation sous forme de cartes de chaleur](#) », Interstats Méthode n° 11, mai 2019, SSMSI
- [Les cambriolages de logements à Paris, Lyon et Marseille sont concentrés dans les coeurs d'agglomérations](#), Interstats Info rapide n° 10, mai 2019, SSMSI
- « Recensement agricole Mayotte 2010 - Traitement et interprétation des données géoréférencées », DAAF Mayotte, Ministère de l'Agriculture

Annexe : Exemple de carte des zones très faiblement habitées (gris) et suffisamment habitées (bleu) sur l'agglomération parisienne



Paramètres :

- grille de lissage composée de carreaux de 200 mètres de côté
- rayon de lissage de 1 000 mètres
- Seuil de densité de logements : 500/km²