
Impact de l'anisotropie du milieu dans les études de profilage géographique

Marie Trotta¹, Cécile Deprez² & Jean-Paul Donnay³

Unité de Géomatique, Université de Liège
17 allée du 6 Août (B5), 4000 Liège, Belgique

1. Marie.Trotta@ulg.ac.be
2. cecile.deprez@student.ulg.ac.be
3. jp.donnay@ulg.ac.be

RÉSUMÉ. Cette communication a pour objectif d'évaluer le biais introduit par l'usage de la distance euclidienne dans le cadre d'analyses de profilage géographique. Elle procède par une comparaison des distances routières et euclidiennes calculées sur près de 1 000 couples d'origines-destinations (sites de crimes – point d'ancrage du criminel) issus de la base de données de la police fédérale belge. Les comparaisons sont différenciées selon la densité du réseau de voirie (milieux rural et urbanisé) et les paramètres statistiques des distributions résultantes sont exploités dans l'application du modèle de décroissance avec la distance, principal outil de profilage géographique.

ABSTRACT. This communication aims at assessing the bias introduced by the use of the Euclidean distance in the context of geographic profiling analyzes. It proceeds by a comparison of road and Euclidean distances of about 1 000 pairs of origin-destination (crime sites – criminal's anchor point) from the database of the Belgian Federal Police. Comparisons are differentiated by the density of road network (rural and urban areas) and the statistical parameters of the resulting distributions are exploited in the application of the most common geographic profiling model called the distance decay model.

MOTS-CLES : science forensique, impact des distances routières, distinction urbain-rural, modèle de décroissance avec la distance.

KEYWORDS: forensic science, impact of road distances, urban-rural discrimination, distance decay model.

1. Introduction

Le caractère isotrope de l'espace est une hypothèse à laquelle ont recours de nombreux modèles géographiques. Les propriétés en un point ne variant pas selon la direction, la distance euclidienne convient pour mesurer la distance entre deux

points quelconques de l'espace. Plusieurs modèles géographiques sont utilisés en criminologie environnementale (Brantingham & Brantingham, 1990). En particulier le modèle de décroissance avec la distance appliqué lors du profilage géographique¹ utilise d'habitude cette hypothèse d'isotropie (Kent *et al.*, 2006 ; Goodwill *et al.*, 2014). Selon les analyses, plusieurs raisons peuvent justifier ce choix : analyse macroscopique, déplacements pédestres sans contraintes naturelles ou anthropiques, déplacements de véhicules sur un réseau routier dense et régulier (tout particulièrement en milieu urbain nord-américain), etc. (Rossmo, 2000). Certains auteurs ont cependant soulevé la question de l'impact de la sous-estimation systématique des distances euclidiennes par rapport aux distances réelles dans les modèles et les applications (Canter *et al.*, 2000 ; Rossmo, 2000 ; Kent, 2003 ; Kent *et al.*, 2006). D'autre part, il a aussi été démontré que la prise en compte de la topologie du réseau de déplacements (la voirie en première analyse), soit l'anisotropie du milieu, influence très sensiblement l'applicabilité du modèle de décroissance avec la distance et peut réclamer une autre forme de modélisation (Trotta, 2013).

Cette communication a pour objectif d'évaluer le biais introduit par l'isotropie, soit l'usage de la distance euclidienne partout sur le territoire, dans le cadre d'analyses de profilage géographique. Elle procède par une comparaison des distances le long de la route et euclidiennes calculées sur près de 1 000 couples d'origines-destinations (sites de crimes – point d'ancrage de l'agresseur) issus de la base de données ViCLAS de la police fédérale belge (Dupuis, 2005). Les comparaisons sont différenciées selon la densité du réseau de voirie (milieux rural et urbanisé) et les paramètres statistiques des distributions résultantes sont exploités dans l'application, selon les différents milieux, du modèle de profilage géographique le plus fréquemment utilisé, soit le modèle de décroissance avec la distance.

2. Sélection des données

La base de données ViCLAS rassemble de multiples données relatives aux faits criminels violents envers les personnes, et résolus, qui se sont déroulés sur le territoire belge depuis 2004. Parmi ces données, figurent la localisation des sites de crimes et délits, et celle du point d'ancrage (généralement la résidence) de l'agresseur, ainsi que, notamment, le type de crime ou délit commis. Les délits de nature sexuelle (agressions et viols) sont parmi les mieux documentés. Ce sont aussi ceux pour lesquels les services de police opérationnelle appliquent les méthodes de profilage géographique dès le moment où ils apparaissent en série.

Après filtrage des cas erratiques (origine ou destination hors du territoire national, impossibilité de rattachement au réseau de voirie, multiples résidences d'un criminel en série, etc.), 1 673 couples origines-destinations liés à des délits de nature sexuelle non intrafamiliaux ont été retenus.

1. "Geographic profiling [...] was originally defined as a criminal investigative methodology for analysing the locations of a connected series of crimes in order to determine the most probable area of offender residence" (Rossmo, 2000)

Afin d'étudier la variabilité spatiale des comparaisons entre distances routières et euclidiennes, le territoire national a été subdivisé en deux catégories, le milieu rural d'une part, et le milieu urbanisé d'autre part, censées traduire des densités et des topologies distinctes du réseau de voirie. Selon O'Leary (2011), cette distinction peut influencer l'allure des courbes des fonctions de décroissance. Pour ce faire, la classification des communes de Belgique a été utilisée (Van Hecke *et al.*, 2009), où les catégories originales « agglomérations » et « banlieues » ont été regroupées pour former l'espace urbain. L'espace rural, regroupant les catégories « rural » et « zone de migrations alternantes » (déplacements quotidiens domicile-travail), a été limité au sud du pays, seul territoire où un nombre suffisant de trajets origine-destination « rurale–rurale » peut être identifié sans traverser les zones urbaines (figure 1).

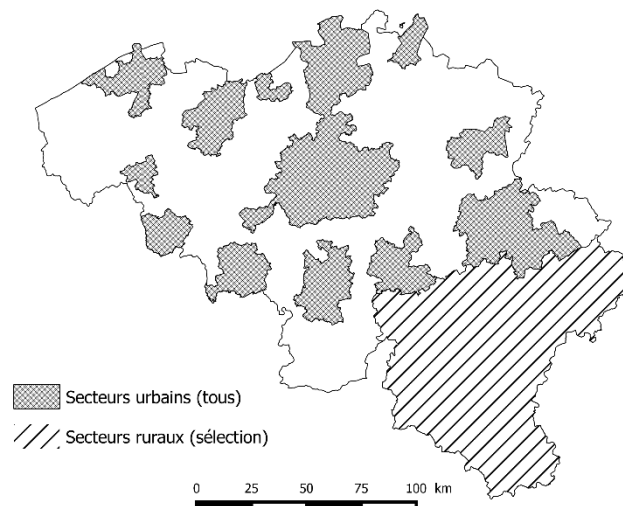


Figure 1. Répartition des zones urbaines et rurales en Belgique, retenues pour l'analyse de la comparaison des distances

3. Comparaison des distances

3.1. Analyse des distributions

Les deux types de distances – le long de la route et euclidiennes – et les deux milieux – rural et urbain – conduisent à l'analyse de quatre distributions des longueurs des trajets entre origines et destinations (figure 2). Les distances euclidiennes sont directement obtenues au départ des coordonnées métriques (Lambert belge 72) fixées lors de la géolocalisation des origines et des destinations. Les distances le long de la route correspondent aux plus courts chemins sur le réseau de voirie paramétré (type de voirie, sens interdit, etc.) selon l'algorithme de Dijkstra

(logiciel ArcGIS 10 – ESRI, 2012). Il faut mentionner que le plan de circulation utilisé (NAVTECH, 2000) n'est pas synchrone avec les faits criminels qui s'étalent sur une décennie. Ceci peut entraîner un biais, sans doute léger mais difficile à apprécier, sur les distances calculées le long de la voirie

Pour le type de crime considéré, deux comportements criminels doivent être différenciés : maraudeurs (agissant à proximité de leur domicile) et navetteurs (agissant loin de leur domicile). Seul le comportement de maraudeur est susceptible de faire l'objet d'un profilage géographique (Trotta *et al.*, 2013) et les distances attendues ne devraient pas dépasser, pour cette catégorie, quelques kilomètres. Les distances supérieures traduisent soit un comportement de navetteur, soit des confusions de localisation des points d'ancrage des criminels (par exemple entre résidence et lieu de travail). Les déplacements étant limités ici à un seul type de milieu, nous avons retenu le rayon moyen des aires urbaines, soit 13 300 m, comme limite supérieure de distance. La même distance maximale sera considérée en milieu rural pour permettre les comparaisons. D'autre part, les distances proches de 0 correspondent à des crimes localisés sur le site d'ancrage du criminel (aux erreurs de localisation près) qui, comme les crimes intrafamiliaux, doivent être exclus de l'analyse géographique. Les analyses réalisées dans cette étude sont ainsi réduites à l'intervalle de distances euclidiennes [50-13 300] m. Cette sélection maintient 770 couples origines-destinations en milieu urbanisé, et 167 couples en milieu rural.

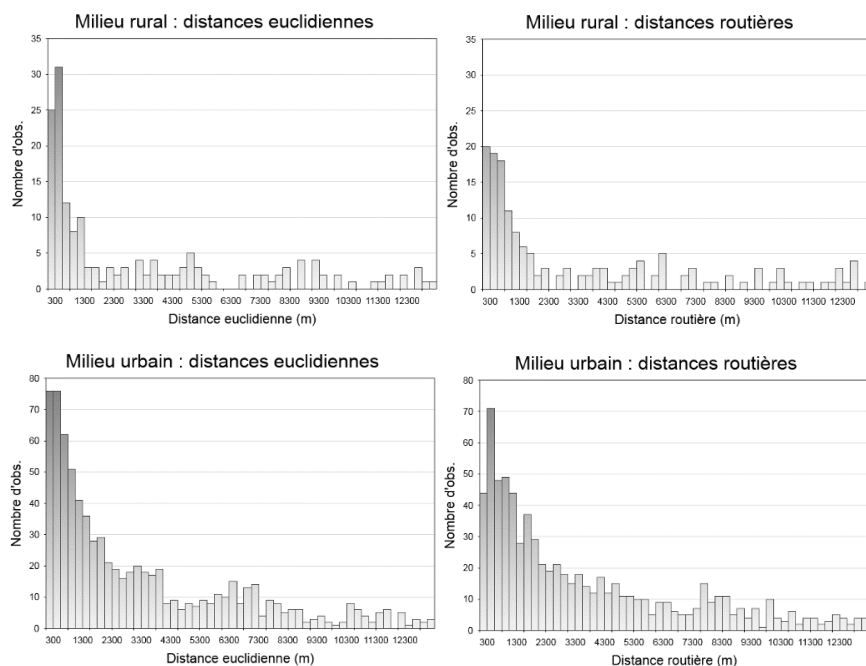


Figure 2. Histogrammes des distances entre le site de crime et le point d'ancrage du criminel

Les distributions ont la même allure deux à deux, tant en milieu rural qu'en milieu urbain. On constate une corrélation manifeste entre les distributions (Kendall sur données appariées – tableau 2), influencée par le fait que les distances routières sont toujours plus grandes que les distances euclidiennes. Cependant, les distributions des deux types de distances sont indépendantes tant en milieu urbain qu'en milieu rural même si la significativité est moindre dans ce dernier milieu, ce qui pourrait, en première analyse, s'expliquer par son effectif sensiblement plus faible (tableau 2). D'autre part, on notera que la distinction entre milieu rural et milieu urbain semble justifiée par l'indépendance significative entre les distributions des distances routières entre elles, d'une part, et les distances euclidiennes entre elles, d'autre part (dernière ligne du tableau 2).

Tableau 1. Paramètres statistiques des distributions (distances en m)

	Milieu rural		Milieu urbain	
	Distance routière	Distance euclidienne	Distance routière	Distance euclidienne
Effectif	167		770	
Minimum	64	56	59	57
Maximum	16 628	13 199	17 319	13 253
Moyenne	4 183	3 305	3 931	3 213
Écart-type	4 633	3 736	3 820	3 217
Médiane	1 562	1 291	2 490	1 942
1 ^{er} quartile	595	452	947	697
3 ^e quartile	6 941	5 114	5 915	5 027

Tableau 2. Tests de comparaisons statistiques des distributions

	Milieu rural		Milieu urbain	
	$\tau = 0.8914$	$p < 0.05$	$\tau = 0.9228$	$p < 0.05$
K-S	$D_{\text{Max}} = 0.1078$	$p < 0.1$	$D_{\text{Max}} = 0.0818$	$p < 0.05$
	Distances euclidiennes		Distances routières	
K-S milieu rural vs milieu urbain	$D_{\text{Max}} = 0.0737$	$p < 0.01$	$D_{\text{Max}} = 0.0807$	$p < 0.025$

3.2. Analyse du rapport de distances

Sauf cas fortuits, où les deux distances sont équivalentes, la distance le long de la route est toujours supérieure à la distance euclidienne (tableau 1). Le rapport moyen entre la première et la seconde (D_r/D_e) est de 1.302 en milieu urbain, et atteint 1.352 en milieu rural (tableau 3). On ne peut cependant rejeter l'hypothèse d'égalité des moyennes, mais bien la similitude des distributions (tableau 3). On constate, d'autre part, que les valeurs maximales du rapport sont obtenues sur de courtes distances, surtout en milieu urbain (figure 3).

Tableau 3. Paramètres statistiques du rapport D_r/D_e

	Milieu rural	Milieu urbain
Effectif	167	770
Minimum	1.000	1.000
Maximum	3.500	5.41
Moyenne	1.352	1.302
Écart-type	0.457	0.382
Kurtosis	8.1740	51.7945
Normalité (K-S)	Dif _{Max} = 0.2364 $p < 0.01$	Dif _{Max} = 0.2220 $p < 0.01$
Comparaisons		
Moyennes	Student $t = 1.4827$	$p = 0.138$
Distributions	K-S D _{Max} = 0.0830	$p < 0.05$

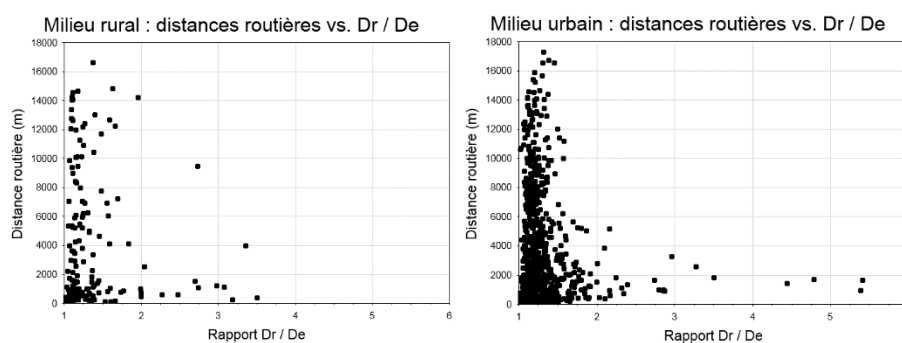


Figure 3. Relation entre le rapport D_r/D_e et les distances routières

Le réseau de voirie est sensiblement moins dense et surtout plus hétérogène en milieu rural, ce qui peut laisser penser à une distribution spatiale non uniforme du rapport D_r/D_e dans ce milieu. L'hypothèse paraît légitime lorsque l'on constate l'excès d'aplatissement négatif (kurtosis) de la distribution rurale du rapport vis-à-vis de la distribution urbaine (tableau 3).

3.3. Autocorrélation du rapport de distances en milieu rural

Une autocorrélation spatiale significative du rapport D_r/D_e permettrait de vérifier l'hypothèse mentionnée dans la section précédente, et de conclure du même coup à l'excès de simplification provoqué par l'usage de la seule distance euclidienne dans les applications menées dans le milieu rural. Nous avons dès lors calculé les indices d'autocorrélation spatiale globale et locale (Anselin, 2004). Une première matrice de voisinage est créée sur base de la distance moyenne attendue entre les destinations deux à deux [$1/(2\sqrt{\text{densité}}) = 3\,352$ m]. Cette métrique laisse plusieurs points isolés, de sorte qu'une seconde matrice basée sur les 6 premiers voisins de chaque point de destination a été aussi constituée pour comparaison. D'autre part, une tessellation de Voronoï (ou Thiessen) autour des points de destination a été utilisée pour calculer les indices selon un voisinage topologique entre polygones. La signification des indices et leur stabilité ont été évaluées sur base de 9 999 permutations et les autocorrélations locales ont été calculées selon les méthodes de Moran (LISA) et de Getis et Ord (G^*). Dans le cas des polygones de Voronoï, ce dernier indice est calculé sur base de la distance séparant les centres des polygones contigus.

L'autocorrélation globale du rapport D_r/D_e calculée selon l'indice de Moran n'est pas significative, quelle que soit la matrice de voisinage utilisée. Par contre, toutes les mesures locales identifient entre 7% et 30% des points de destination présentant des valeurs significatives à très significatives d'indice local (tableau 4). Ces points se répartissent en groupes stables, d'autocorrélation positive ou négative, comme l'illustre à titre d'exemple la cartographie de l'indice LISA par polygone de Voronoï (figure 4).

Tableau 4. Autocorrélation spatiale globale et locale de D_r/D_e en milieu rural

	Distance moyenne attendue			6 Voisins			Voronoi		
	#	min	max	#	min	max	#	min	max
<i>I</i> de Moran	-0.0134			0.0122			0.0051		
LISA	50	-0.8629	0.8690	12	-1.4128	1.6939	16	-2.0234	1.650
G^*	49	0.0049	0.0079	15	0.0049	0.0084	16	0.0063	0.0127

: nombre d'entités significativement auto-corrélées à leur voisinage

Dans la figure 4, les autocorrélations significatives sont classées en 4 groupes, selon les quadrants du diagramme de Moran, en qualifiant d'abord la valeur standardisée de l'indice du polygone courant, puis celle de son voisinage. Cette même figure fait apparaître quelques éléments géographiques qui méritent d'être soulignés, tels que la proximité immédiate des groupes significatifs avec la zone urbaine (coin nord-ouest, voisin de l'agglomération de Namur, à la périurbanisation rapide en raison de son statut assez récent de capitale administrative régionale) et avec les frontières du Grand-Duché de Luxembourg et la France (partie sud et sud-est, constituant pour divers motifs socio-économiques une zone dense de migrants alternants vis-à-vis de l'agglomération de Luxembourg). On pourrait conclure à la nécessité de sur-segmenter l'espace rural mais, d'une part, les critères de sous-régionalisation n'agissent pas de façon ubiquiste (pas d'autocorrélation locale significative au sud de Liège ou le long de la frontière allemande, par exemple) et, d'autre part, on multiplierait ad libitum les modèles locaux reposant sur un nombre toujours plus réduit de données. On note aussi la concentration des sites de crimes et l'alignement des petits polygones associés le long des axes autoroutiers, en particulier l'autoroute Namur-Luxembourg. Ce dernier point suggérerait plutôt une comparaison entre la distance euclidienne et la distance-temps, plutôt que la distance le long de la voirie, dans les parties de la zone rurale traversées par d'importants axes routiers.

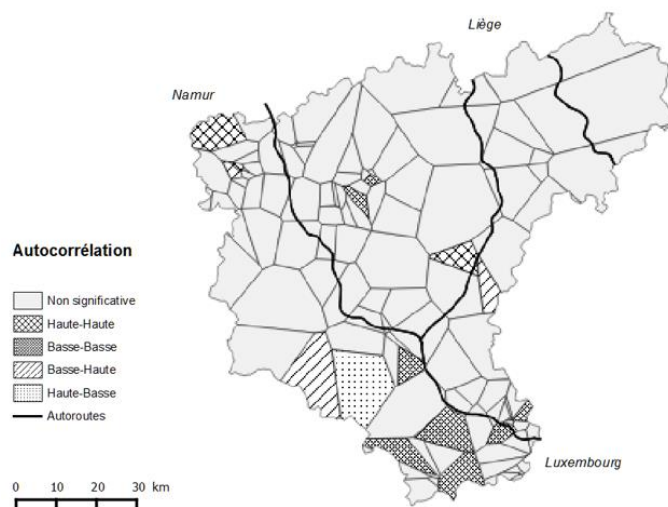


Figure 4. Clusters formés par les valeurs significatives de LISA en milieu rural
légende dans l'ordre : polygone- voisinage

L'autocorrélation locale d'une dizaine de pour-cent des secteurs confirme l'hétérogénéité du rapport D_r/D_e en milieu rural. Si l'on ajoute à cela les distances plus longues enregistrées dans ce milieu, on pressent une plus grande difficulté à

utiliser le modèle de décroissance avec la distance en milieu rural, telle que notée par ailleurs dans la littérature (Marshall *et al.*, 2005).

4. Ajustement d'un modèle de profilage géographique

4.1. *Modèle de décroissance avec la distance*

L'application du modèle de décroissance avec la distance consiste à ajuster, sur l'ensemble des cas résolus du type de crimes considéré, une fonction de décroissance de la fréquence avec la distance séparant le site de crime du point d'ancrage de l'agresseur. Plusieurs types de fonctions sont utilisés dans la littérature, les plus fréquents étant les fonctions linéaires et les exponentielles négatives. La fonction choisie est alors utilisée pour construire une « carte de vraisemblance ou de probabilité »² centrée sur chaque site de crime attribué à la série. Les vraisemblances ainsi construites sont additionnées, et la zone la plus probable est censée contenir le point d'ancrage du criminel cherché.

Si les distances euclidiennes sont utilisées, les zones de vraisemblances correspondent à des cercles concentriques, centrés sur chaque site de crime. Par contre, si les distances sont mesurées le long de la route, les vraisemblances peuvent être propagées, depuis chaque site de crime, le long des voiries. La procédure est réalisée en mode maillé, après rasterisation du réseau.

On notera encore que le modèle n'a de prétention prédictive que dans des dispositions géométriques particulières. En première analyse, il faut accepter l'hypothèse que le point d'ancrage attendu soit situé entre les sites de crimes, soit de manière pragmatique, au sein du polygone convexe construit sur ces derniers. En dehors de cette situation, d'autres modèles prédictifs peuvent éventuellement être utilisés.

4.2. *Ajustement des fonctions de décroissance*

Les histogrammes de la figure 2 sont le support de l'ajustement de quatre fonctions exponentielles négatives. Les effectifs sont calculés par tranche de 250 m, à partir de 50 m, et ils sont limités aux cas présentant une distance euclidienne maximale de 13300 m. Ils sont soumis à une transformation logarithmique pour linéariser le modèle de régression (tableau 5).

Pour permettre une comparaison avec la solution effectivement mise en œuvre dans une investigation criminelle classique, l'ajustement des effectifs à la distance euclidienne a également été calculé pour l'ensemble des crimes enregistrés dans la

² Les termes de « probabilité » ou de « vraisemblance » communément utilisés dans ce contexte n'ont pas la signification statistique habituelle et sont employés par facilité. Les valeurs absolues et les unités n'ont pas d'utilité et seules les valeurs relatives et les extrema locaux de la distribution spatiale des valeurs sont analysés.

base de données, quels que soient les milieux des origines et des destinations. Cet ensemble de données reprend donc nos deux échantillons, urbain et rural, mais aussi les faits d'origine urbaine et de destination rurale et inversement. Les paramètres de cet ajustement global apparaissent à la dernière ligne du tableau 5.

Tableau 5. Ajustements des fonctions de décroissance de la distance de la forme $\ln(\text{effectif}) = b_0 + b_1 * \text{distance}$

	b_1	b_0	R^2 ajusté
Milieu rural – Distances routières	-0.000108	1.654481	0.3579
Milieu rural – Distances euclidiennes	-0.000125	1.735968	0.3691
Milieu urbain – Distances routières	-0.000209	3.626893	0.8471
Milieu urbain – Distances euclidiennes	-0.000240	3.740030	0.7648
Ajustement global – distances euclidiennes	-0.000146	3.673688	0.6899

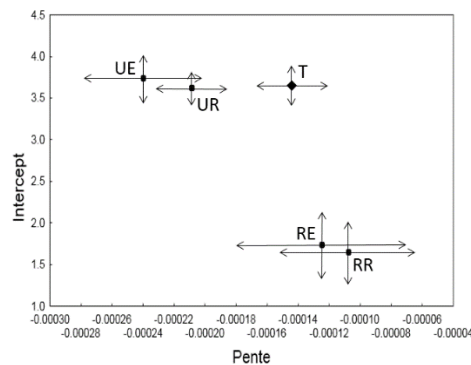


Figure 5. Intervalles de confiance $\pm 95\%$ des coefficients de régression
 UE : Urbain Euclidien – UR : Urbain Routier
 RE : Rural Euclidien – RR : Rural Routier
 T : ajustement global euclidien

Tous les coefficients sont significatifs au seuil $\alpha = 5\%$. Les ajustements en milieu urbain, d'une part, et en milieu rural, d'autre part, sont parfaitement différenciés, entre eux et vis-à-vis de l'ajustement global. De façon attendue, les pentes s'avèrent plus fortes dans le cas des distances euclidiennes, quel que soit le milieu. Cependant, si les ajustements sont pratiquement distincts en milieu urbain selon les deux types de distances, ils ne sont pas significativement différents en milieu rural (figure 5), où la qualité des ajustements est moindre. Ces comparaisons reflètent des considérations similaires déjà émises lors de l'examen des distributions (tableau 2). Mais on pourrait aussi penser que la portée maximale de 13 300 m utilisée dans cette

analyse – jugée plausible pour le milieu urbain et généralisée au milieu rural pour permettre les comparaisons - est trop courte pour le milieu rural. En milieu rural, en effet, les distances moyennes parcourues sont plus longues qu'en milieu urbain et le rapport D_r/D_e est plus élevé. Afin de tester cette éventualité, des ajustements ont été testés en boucle, par pas de 500 m pour garantir des effectifs suffisants, sur des portées croissantes jusqu'à la distance euclidienne maximale rencontrée dans la zone rurale considérée. La portée différenciant au mieux les pentes des ajustements avec les distances routières et les distances euclidiennes d'une part, et maximisant les R^2 ajustés d'autre part, est de l'ordre de 25 000 m (tableau 6), mais les coefficients ne sont toujours pas significativement différents entre les deux types de distances au seuil α de 5 %. La qualité des ajustements décroît très rapidement au-delà de cette portée, les observations devenant rarissimes.

Tableau 6. Ajustements des fonctions de décroissance de la distance en milieu rural maximisant le rapport des pentes b_1 et l'ajustement R^2 (portée 25 000m)

	b_1	b_0	R^2 ajusté
Milieu rural – Distances routières	-0.000066	1.718584	0.4112
Milieu rural – Distances euclidiennes	-0.000087	2.001422	0.4883

5. Discussion sur les effets de l'anisotropie du milieu

5.1. Milieux urbain et rural

On doit d'abord constater que la discrimination entre milieux rural et urbain s'est avérée confirmée à toutes les étapes de l'analyse. Quel que soit le type de distance utilisé, il est souhaitable d'exploiter un ajustement propre au milieu au sein duquel se perpétuent les faits criminels, plutôt que de se baser sur un modèle ubiquiste. Le milieu urbain fournit des ajustements très satisfaisants sur base de données étudiées. Or, englobant des quartiers générateurs ou attracteurs de criminalité, moins sensibles aux effets de la distance parcourue par l'agresseur, le milieu urbain est traité avec circonspection dans les analyses de profilage géographique. Cette étude tend à démontrer que, pour le type de faits criminels considéré et dans le respect des hypothèses d'application, les modèles de profilage géographique peuvent s'exprimer pleinement en milieu urbain.

Par opposition, le milieu rural montre une plus grande sensibilité aux méthodes de profilage géographique. Vis-à-vis du milieu urbain, les distances parcourues, tous types confondus, sont plus grandes et le rapport entre la distance routière et la distance euclidienne est plus élevé. L'ajustement des fréquences des faits criminels avec la distance est moins efficace qu'en milieu urbain, et les données utilisées ne peuvent différencier les ajustements avec la distance routière et la distance euclidienne. D'autre part, une faible variation des caractéristiques locales dans un secteur paraît modifier les caractéristiques de distances. La proximité d'une

agglomération, d'une frontière ou d'un axe routier important seraient au nombre de tels éléments perturbateurs, comme l'indique l'analyse de l'autocorrélation locale du rapport des distances, mais sans pouvoir généraliser ces effets. Plusieurs pistes peuvent être envisagées pour tenter de lever ces difficultés. D'une part, on peut penser à un découpage plus fin de l'espace rural, au risque de tendre à une analyse idiographique perdant toute espérance de généralisation. D'autre part, il est possible d'envisager un ou plusieurs modèles alternatifs à celui de la décroissance de la distance (Canter *et al.*, 2012 ; Trotta, 2012). Cette dernière piste paraît plus prometteuse, même si chaque modèle repose sur des hypothèses et présente des limites intrinsèques.

5.2. Distances euclidiennes et routières

Trois effets conjugués résultent de l'utilisation de la distance routière à la place de la distance euclidienne. Tout d'abord, la proximité des sites n'est plus dictée par leurs positions géographiques, mais par la connectivité du réseau. Ensuite, on constate une diminution de la pente de l'ajustement qui s'explique aisément par des distances routières systématiquement plus grandes que les distances euclidiennes. Enfin, l'espace candidat à la localisation de la zone d'ancrage de l'agresseur est réduit très sensiblement, en l'occurrence aux seuls tronçons de voirie participant aux connexions entre les sites de crimes.

En construisant la surface cumulée de vraisemblance au départ des sites de crimes avec la distance euclidienne, la pente plus forte de l'ajustement conserve des probabilités fortes à proximité des sites eux-mêmes, qui ne peuvent être dépassées qu'occasionnellement lors du cumul de vraisemblances associées à des sites proches. L'usage de la distance routière, en adoucissant la pente d'ajustement, étale les valeurs de probabilités importantes et permet plus facilement d'obtenir des valeurs maximales à l'endroit des vraisemblances cumulées, même entre sites criminels éloignés.

Afin d'illustrer ce principe, la figure 6 a été construite sur une série simulée utilisant les paramètres d'ajustement du tableau 6, mais sans utiliser le réseau de voirie. Toutes les valeurs sont propagées dans l'espace continu autour des sites de crimes pour permettre une comparaison de l'effet de la pente « toutes autres choses étant égales ».

Lorsque l'on introduit la géométrie du réseau pour reporter et limiter la propagation des vraisemblances construites sur la distance routière, les valeurs maximales apparaissent sur les tronçons participant à plusieurs chemins les plus courts entre sites de crimes deux à deux, et tout particulièrement aux carrefours de ces chemins. La réduction de la zone d'investigation qui s'ensuit est illustrée à la figure 7, utilisant la même série que dans l'illustration précédente. Les surfaces de vraisemblances sont cette fois cumulées au départ des 5 sites de crimes, de façon isotrope pour les distances euclidiennes, et le long du réseau pour les distances

routières. Ces surfaces sont seuillées, depuis leurs extrema locaux respectifs jusqu'à la valeur incluant le site d'ancrage, préalablement connu dans cette simulation.

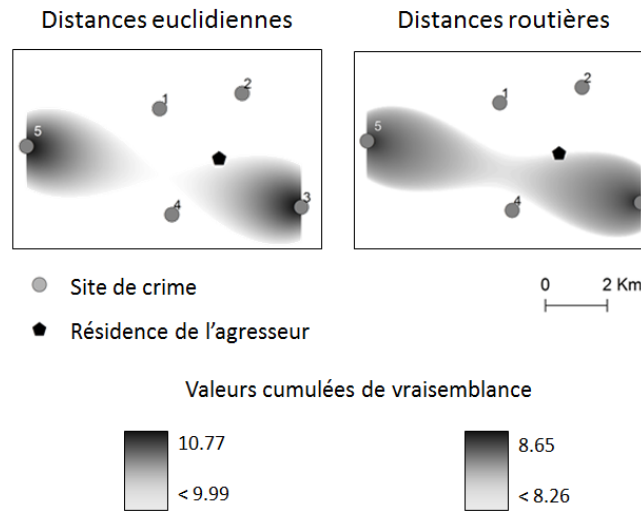


Figure 6. Cumul des vraisemblances depuis deux sites de crimes distants (3 & 5)

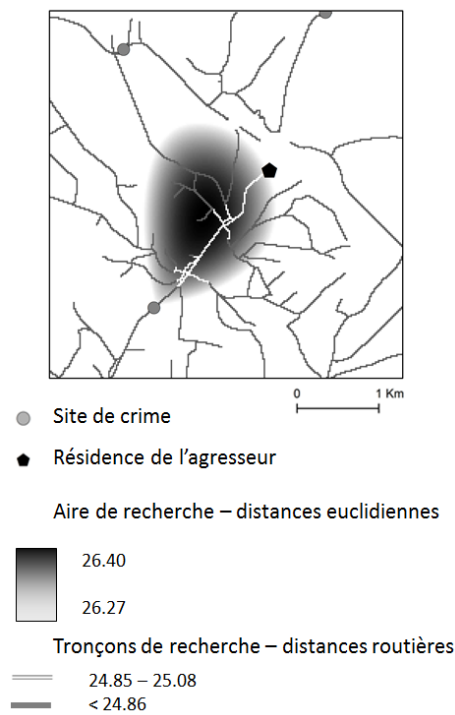


Figure 7. Comparaison des zones d'investigation limitées aux valeurs de vraisemblance permettant d'atteindre le site d'ancrage

6. Conclusions et perspectives

L'objectif de ce travail consiste à vérifier si l'usage d'un espace isotrope introduit un biais dans les méthodes de profilage géographique, et plus particulièrement dans le modèle de décroissance avec la distance largement utilisé en pratique dans la recherche de la zone d'ancrage des criminels en série. La recherche utilise une base de données officielle d'où ont été extraits 867 délits de type sexuel enregistrés en Belgique depuis 2004. Elle analyse l'influence du type de distance pour deux milieux distincts (urbain et rural) selon une division territoriale simple mais validée par la littérature géographique nationale. Au vu des résultats, l'usage de fonctions d'ajustement distinctes selon les milieux urbain ou rural paraît souhaitable. En outre, en milieu urbain, le recours aux distances routières, plutôt qu'euclidiennes, favorise une pente plus forte, diminuant les recouvrements possibles entre les surfaces de vraisemblance construites à partir de chacun des sites. Il reste que les conditions d'application du modèle de décroissance avec la distance doivent être rencontrées ce qui, en milieu urbain, est loin d'être garanti au départ d'une investigation criminelle. L'identification des comportements maraudeurs dans des environnements attractifs est en effet délicate. En milieu rural, les résultats obtenus sont très nuancés. Ils auraient plutôt tendance à remettre en cause la validité du modèle de décroissance avec la distance lui-même, plutôt que ses modalités d'application. La variabilité entre agresseurs n'est pas adaptée à l'usage d'une décroissance calibrée sur des faits agrégés. Ce constat va dans le sens d'études récentes proposant des modèles alternatifs (Canter *et al.*, 2012 ; O'Leary, 2011 ; Trotta, 2012).

Le modèle de décroissance avec la distance qui a été utilisé dans cette recherche n'est ni le seul modèle de profilage géographique, ni la seule technique à faire usage du concept de distance en cartographie criminelle. Si l'évolution des modèles de profilage est récente, en cartographie criminelle, par contre, plusieurs techniques classiques de mise en évidence des « points chauds » (*hotspots*) de criminalité utilisent également la distance, mesurée cette fois entre faits criminels d'un même type. Parmi ces techniques, l'estimation sur base de noyaux de densité (KDE pour *Kernel Density Estimator*) est largement répandue. Son application réclame la fixation de deux paramètres : d'une part, la résolution, dont la littérature semble minimiser le rôle (Chainey, 2013) et, d'autre part, la bande passante ou largeur de la fenêtre de convolution, directement liée au concept de distance, est qui jouerait un rôle fondamental. Une extension de la présente recherche consistera en une évaluation des différences causées par l'utilisation de la distance routière en lieu et place de la distance euclidienne dans la définition de la bande passante lors de l'application de cette technique.

Remerciements

Les auteurs remercient la Police fédérale de Belgique pour avoir permis l'accès à la base de données ViCLAS. Les recherches de Marie Trotta sont réalisées avec l'aide financière du Fonds de la Recherche Scientifique-FNRS.

Bibliographie

- Anselin L. (2004). *GeoDa 0.9.5-i Release Notes*. Spatial Analysis Laboratory, University of Illinois, Urbana-Champaign.
- Brantingham P.J. & Brantingham P.L. (1990). *Environmental Criminology*. Waveland Press, Long Grove (Ill.).
- Canter D., Coffey T., Huntley M. & Missen C. (2000). Predicting Serial Killers' Home Base Using a Decision Support System. *Journal of Quantitative Criminology*, 16(4), p. 457-475.
- Canter D., Hammond L., Youngs D., & Juszcak P. (2012). The Efficacy of Ideographic Models for Geographical Offender Profiling. *Journal of Quantitative Criminology*, p. 1-24.
- Chainey S. (2013). Examining the influence of cell size and bandwidth size on kernel density estimation crime hotspot maps for predicting spatial patterns of crime. *Bulletin de la Société Géographique de Liège*, 60, p. 7-19.
- Dupuis B. (2005). Délits sexuels et violents. ViCLAS élargit l'horizon des enquêtes. *Inforevue. Magazine de la Police intégrée*, 2005(4), p. 22-24.
- ESRI (2012). *Algorithmes utilisés par Network Analyst.*, <http://help.arcgis.com/fr/arcgisdesktop/10.0/help/index.html#/004700000053000000>
- NAVTECH. (2000). NAVSTREETS™ Street Data. Arcview Format.
- Goodwill A.M., van der Kemp J.J. & Winter J.M. (2014). Applied Geographical Profiling. Bruinsma G. & Weisburd G. (eds.). *Encyclopedia of Criminology and Criminal Justice*. New York, Springer, p. 86-99.
- Kent J. (2003). *Using functional distance measures when calibrating journey-to-crime decay algorithms*. Master of Natural Sciences. Louisiana State University and Agricultural and Mechanical College.
- Kent J., Leitner M. & Curtis A. (2006). Evaluating the usefulness of functional distance measures when calibrating journey-to-crime distance decay functions. *Computers, Environment and Urban Systems*, 30(2), p. 181-200.
- Marshall, B., & Johnson, S. (2005). *Crime in rural areas: a review of the literature for the Rural Evidence Research Centre*. Londres, Jill Dando Institute of Crime Science, University College.
- O'Leary, M. (2011). Modeling criminal distance decay. *Cityscape: A journal of Policy Development and Research*, 13(3), p. 161-198.
- Rossmo K. (2000). *Geographic profiling*. CRC Press, Boca Raton.
- Trotta M. (2012). New hypotheses on serial offender's spatial behaviour. Pundt H. & Bernard L. (eds.), *Proceedings 1st AGILE PhD School*, Wernigerode, Shaker Verlag, p. 1-9.
- Trotta M. (2013). Selection of a spatial hypothesis in geographic profiling using graph theory. *Geographic Information Science at the Heart of Europe – AGILE 2013*, Leuven.
- Trotta M., Lemaître A., Donnay J.P. (2013). Operationality of geographic profiling through a hypothetico-deductive method. A review of constraints and factors. *Bulletin de la Société Géographique de Liège*, 60, p. 45-57.
- Van Hecke E., Halleux J.-M., Decroly J.-M. & Mérenne B. (2009). *Noyaux d'habitat et Régions urbaines dans une Belgique urbanisée*. Enquêtes socio-économiques 2001. Monographies. Bruxelles, Direction générale Statistique et Information économique.

