OpenEdition
Journals

Cybergeo: European Journal of Geography

A comparison of two methods for classifying trajectories: a case study on neighborhood poverty at the intra-metropolitan level in Montreal

Comparaison de deux méthodes d'identification des trajectoires de changement : le cas de la distribution spatiale de la pauvreté à Montréal

PHILIPPE APPARICIO, MYLÈNE RIVA ET ANNE-MARIE SÉGUIN
https://doi.org/10.4000/cybergeo.27035

👍 🍪

...ve examined changes in the distribution of poverty in the ... empirical work assessing neighborhood change between ... to make a methodological contribution to the study of ...ring two classification methods, one classical (k-means ... Latent Class Growth Modelling; LCGM) to identify groups ...nilar trajectories of poverty in the Montreal metropolitan ... poverty are measured over a twenty-year period, using ...rformance of the LCGM vs. the k-means clustering was ...inomial logistic regressions examining how different ...ociated with the trajectories of poverty. Results showed ...fied similar groups of census tracts characterized by ... poverty levels throughout the period, with LGCM only ...s clustering.

Durant les dernières années, plusieurs auteurs ont examiné les changements dans la répartition de la pauvreté au sein des villes nord-américaines. La plupart de ces travaux empiriques ont porté sur deux points dans le temps. Cet article vise à apporter une contribution méthodologique pour l'étude du changement dans les quartiers, en comparant deux méthodes de classification : l'une bien connue (le K-means) et l'autre plus récente (*Latent Class Growth Modelling*; LCGM) pour identifier les groupes de secteurs de recensement ayant suivi des trajectoires de pauvreté similaires dans la région métropolitaine de Montréal (Canada). Les trajectoires sont identifiées sur une période de vingt ans en utilisant cinq années de recensement (1986, 1991, 1996, 2001, 2006). La performance relative du LCGM versus le k-means a été évaluée à l'aide d'une série de régressions logistiques multinomiales examinant comment les différentes variables socio-économiques ont été associées avec les trajectoires de pauvreté. Les deux méthodes permettent d'identifier à la fois des trajectoires de pauvreté stables, ascendantes et descendantes. Toutefois, le LCGM est très légèrement plus performant que le k-means.

Entrées d'index

Texte intégral

Introduction

1   In the past decades, rapid urban expansion in conjunction with a growing income gap (Heisz *et al.*, 2006) and structural forces such as economic restructuring, demographic shifts, as well as changes in government policies have impacted on the socioeconomic divisions in North American cities (Kitchen *et al.*, 2009; Walks, 2001), as well in European cities (Van Kempen *et al.*, 2009). Hence the spatial dimension of urban and neighbourhood change in cities of developed countries is receiving increasing attention (Kitchen *et al.*, 2009). To date, most of the empirical work has examined neighbourhood change between two time points (see for example the recent works of Mikelbank. 2006: Reibel *et al.*, 2011; Vicino, 2008). Except the recent work

leveland-Akron metropolitan area, and the one of
real, few studies have identified trajectories of change
lowing the direction and magnitude of trends to vary
points, and how this change plays out at the intra-
es of gentrification (Berry, 1986; Bunting *et al.*, 1988;
impoverishment of inner suburbs documented in
2006; Jargowsky, 2003; Lee *et al.*, 2007; McConville
) are indicative of important transformations in the
important to grasp in their dimensions.
et al., 2012), we use a new clustering technique for
growth model (LCGM) – to identify trajectories of
ntreal over five consecutive census years (1986 to
tinomial logistic regression analysis for modelling the
M.
patial dataset of that previous study, our objective in
While the classification methods such as k-means
luster analysis are largely used in quantitative

geography, the LCGM approach is still little-known by many geographers. Consequently, the purpose of this study is to compare LCGM versus the k-means clustering to assess which method performs better in identifying trajectories of neighbourhood change. The potential benefits of the LCGM in quantitative research in urban geography are discussed as well. Indeed, this approach for classifying spatial longitudinal data can be applied to analyze the neighbourhood changes such as poverty, immigration, ethnic or social diversity, unemployment, etc.

4   For illustrative purposes, we analyze the geography of low-income population as measured at five time points over a 20-year period in Montreal, i.e. the trajectories of relative poverty at the neighbourhood level. In practical terms, the paper aims to identify groups of neighbourhoods characterized by a similar pattern of change in their poverty levels. The present study therefore attempts primarily to make a *methodological* contribution to the study of neighbourhood change.

Background

Studying neighbourhood poverty change

5   Several studies concerned with changes in high-poverty neighbourhood, at least in the USA, have focussed on the growth of these neighbourhoods in suburban areas, and especially in inner-ring suburbs. Some studies aim to identify the evolution in the distribution of poverty zones in metropolitan areas, notably by opposing the central city to inner-ring suburbs (Cooke *et al.*, 2006; Lee, 2011; Lee *et al.*, 2007). Other research, such as the study by McConville and Ong (2003), is interested by the trajectories of poor neighbourhood, i.e. whether they stayed poor, worsen or improved over time, in relation to change in other neighbourhood conditions, e.g. ethnic, immigration, education, employment and family profiles.

6   In Canada, few studies have examined transformations of poverty neighbourhood. In 2000, Ley and Smith (2000) noted the changing nature of some census tracts in Toronto, Montreal and Vancouver over a twenty year period, observing that deprived neighbourhoods in 1971 were not necessarily deprived by 1991 and inversely, none-deprived neighbourhood in 1971 could have become deprived by 1991. Their

several cumulative indicators associated with
esholds) measured first in 1971 and then in 1991.

McLeod (2006) showed that both the proportion of
nd their spatial distribution across the different
aried between 1981 and 2001. For example, they
981, low-income neighbourhoods in Montreal and
cated in inner city neighbourhoods by 2001 and more
Although this study by Heisz and McLeod identified
itan area, the authors did not identify neighbourhood

ms (2009) set in Saskatoon, a Canadian metropolitan
d neighbourhood change between 1991 and 2001,
991-1996 and 1996-2001. Their analyses considered
neighbourhoods at the beginning of the period (1991),
or high socioeconomic status (SES) neighbourhoods.
racterised as following three possible trajectories of
2001, i.e. decline, improvement or stability. Kitchen
eresting as it considered socioeconomic conditions at

the beginning of the study period and their change over a 10 year period. Yet ten years of observation might be too short to identify important changes in urban processes such as filtering down (out-migration of households to newer and more elaborate dwellings and in-migration of households of lesser wealth and lower social status), gentrification, and the suburbanization of poverty.

9   Our study proposes to analyse trajectories of neighbourhood poverty on a longer 20-year period. As Kitchen and Williams (2009), we consider poverty levels at the beginning of the period, i.e. 1986, but also neighbourhood poverty levels in 1991, 1996, 2001 and 2006. Thus, over time, a neighbourhood could be characterised by an ascending trajectory, then a descending trajectory, followed by a stable level of poverty, to a declining slope at the end of the study period. At each time point poverty is measured as a continuous variable. This will allow us to identify trajectories with more precision and to identify an optimal number of trajectories. Statistically, this will be achieved by applying and comparing two clustering techniques to group neighbourhoods having followed similar trajectories of change in poverty levels. Hence the objective of this paper is essentially a methodological one, i.e. to identify a clustering method most suited to measure neighbourhood change. We will see later that these approaches allow maximising variation between trajectories and minimising variation within trajectories. This step is crucial if the aim is to identify the determinants of neighbourhood change and to measure their relative importance. This paper builds on previous work by defining trajectories of poverty using five time points, allowing for the magnitude and direction change in poverty to vary at each time point.

10   This precision, however, comes with methodological challenges, including the construction of a longitudinal database at the intra-metropolitan scale (i.e. at the census tract level) with comparable and harmonised socioeconomic data and geographical boundaries across several census years. Another methodological challenge is to find the most accurate approach to group neighbourhoods characterized by a similar evolution of their poor population over time, with each group (i.e. trajectory) being most different amongst themselves. In the next section, we discuss two possible techniques.

Identifying trajectories of neighbourhood change at the intra-metropolitan level: the possibilities
...s clustering and latent class

👍 🍪

Ce site utilise des cookies et
vous donne le contrôle sur
ceux que vous souhaitez
activer

✓ Tout accepter

✕ Tout refuser

Personnaliser

Politique de confidentialité

...ainly been tackled from time-series and econometric
...economic cycles and changes in labour markets at
...so possible to envision trajectories of change in terms
...owed a similar pattern of change along a variable (or
... To date unsupervised clustering techniques such as
...CA) and k-means clustering, or extensions of k-means
...al., 1998), partitioning around the medoid (Kaufman
...a description and a comparison of several extensions
...instreamed in urban geography and social sciences,
...ster spatial 'objects' with similar characteristics at one
...Mikelbank, 2004; Vicino et al., 2011) and less often at
..., 2007, 2011; Vicino, 2008).

12   However, it is possible to apply these 'classical' clustering techniques to a longitudinal dataset to identify neighbourhood trajectories. For example, Mikelbank (2011) applied an HCA to group census tracts along several demographic, housing and socioeconomic variables extracted from four censuses (1970, 1980, 1990, 2000). These variables were standardized (z-scores) at each time point, and were then appended in one table, on which an HCA was computed. This allowed identifying five types of neighbourhoods over the period: struggling, struggling African American, stability, new starts, and suburbia. Thus, a census tract could belong to the same neighborhood type during the four census years, or it could change type one or more times. Finally, Mikelbank (2011) built several transition tables to identify stability or change in neighbourhood trajectories across two years or all the time periods.

13   Whereas it is possible to apply classical clustering techniques to longitudinal spatial datasets, new semi-parametric analytical procedures have recently been developed to groups 'objects' having followed similar trends over time, namely Latent Class Growth Models (LCGM) (Nagin, 2005). To date, LCGM has mainly been applied in psychology (Nagin, 2005) and epidemiology, for example, to group individuals with similar trajectories of change in health-related behaviours (Barnett *et al.*, 2008; Brookmeyer *et al.*, 2009). To our knowledge, it has been applied to spatial data to examine trajectories of change within a metropolitan area or a country in only three studies (Apparicio *et al.*, 2012; Pearson *et al.*, 2013; Riva *et al.*, 2012).

14   K-means and LCGM are statistical clustering techniques that can be applied to classify objects (i.e. census tracts) into k number of clusters (i.e. trajectories) with similar change in a variable (i.e. poverty) over time. K-means is an exploratory statistical clustering technique that uses an allocation/re-allocation algorithm to optimally reassign census tracts to the nearest cluster centroid (Everitt *et al.*, 2001). The goal is to maximize between clusters variations and to minimize within cluster variations, and thus to group into k types of local areas having followed similar trajectories of change in poverty over the study period. Some statistical software, e.g. SAS, optimizes the choice of initial cluster centres; thus, the random selection of cluster centres, potentially leading to different solution when the model is re-run, is no longer being a 'problem'. Compared to HCA algorithms, the number of k must be chosen a priori in k-means clustering and there are several methods to identify the optimal cluster solution (Milligan *et al.*, 1985) including, amongst others: the Pseudo-F statistic (Calińskia *et al.*, 2007); the Cubic Clustering Criterion (Sarle, 1983); or a more discursive method (Tibshirani *et al.*, 2001) consisting of plotting the average r each cluster solution and visually identifying the there is a natural levelling off in the distribution (the

👍 🍪

proach to classification (Andruff *et al.*, 2009; Collins 09). Although each census tract will follow a unique levels, the heterogeneity in the distribution of census set of polynomial functions, each corresponding to a Andruff *et al.*, 2009; Nagin, 2005). Because the ge can freely vary between trajectories, a set of model lopes, is estimated for each trajectory (Andruff *et al.*, rajectory, the slope and intercept are treated as fixed n LCGM, the optimal number of classes is informed by whereby the modelling starts with a one-class model, ded to evaluate improvement in model fit. The model a is identified by interpreting and comparing several nodel with the lowest Bayesian Information Criterion ies of group membership (a maximum-probability

assignment rule is used to assign each individual to the trajectory to which it holds the highest posterior membership probability) (Andruff *et al.*, 2009). LCGM is now relatively easy to apply in software such as SAS (ProcTRAJ) (Jones *et al.*, 2007), Mplus (Muthén & Muthén), and LatentGOLD (Statistical Innovations). The main differences between k-means and LCGM are summarised in Table 1 (adapted from LatentGold website: http://www.statisticalinnovations.com/articles/kmeans2a.htm).

16   The underlying principles of k-means and LCGM are thus different: one is a descriptive/ exploratory technique whereas the other adopts a semi-parametric approach to classification. It is also worth noting that K-Means requires continuous or dichotomous variables while LCGM can be applied to any type of distribution (continuous, ordinal, nominal, count and binomial). In this study, the variable used for the classification –i.e. the location quotient– is continuous. Yet no studies have compared how the methods fare in generating groups of spatial units having followed similar trajectories of change.

Table 1. Differences between k-means clustering and latent class growth modelling

	K-means	LCGM
Approaches to classification	Use of allocation/re-allocation algorithm (and ad-hoc distance measure) to optimally reassign objects to nearest cluster centre.	Probability-based method of classification also producing information on misclassification of object into clusters.
Identification of number of optimal clusters	Lowest average distance to cluster centre; cubic clustering criterion; pseudo-F statistic.	Various modelled-based diagnostics such as the BIC statistic and the posterior probabilities of group membership.
Types of variables and standardization	Interval scale and dichotomous variables for which Euclidean distance measures can be calculated. Standardization of variables is recommended.	Continuous, categorical (nominal or ordinal), counts variables or any combination of these. Standardization of variables is not necessary. Covariates can be used to maximise the classification

Objectives of the study

👆 🍪

Ce site utilise des cookies et vous donne le contrôle sur ceux que vous souhaitez activer

✓ Tout accepter

✗ Tout refuser

Personnaliser

Politique de confidentialité

e trajectories of neighbourhood poverty change, the
and compare two clustering techniques to 20 years of
identify groups of neighbourhoods having followed
etween 1986 and 2006. We apply both k-means and
method performs better in identifying trajectories of
ost accurate classification represents a crucial step
models of socioeconomic changes operating in

trated that the clustering accuracy of the k-means is
lly when computed on large data sets (see for example
ults of HCA vary according to the distance metric
clidean distance, Mahalanobis distance, etc.) and the
e, median, centroid linkages, Ward's method, etc.). To
GM models to several variants of the HCA, we opted

Material and methods

Study area and data

19   This study is set in Canada, in the Montreal Census Metropolitan Area (CMA) comprising a population of about 3.64 million inhabitants spread over a territory of 4,259 km² in 2006 (Statistics Canada, 2007). Intra-metropolitan areas are defined using census tract boundaries. Administrative and census geographical boundaries in the Montreal CMA have changed considerably between 1986 and 2006: the number of census tracts increased from 698 to 825 over this period. Harmonisation of the geographic boundaries of census tracts was therefore necessary. Starting with the 1986 census tracts geographic boundaries (the earliest time point), this was achieved by aggregating contiguous census tracts in order to obtain comparable boundaries across all census years. A total of 611 census tracts was obtained.

Figure 1. The study area.
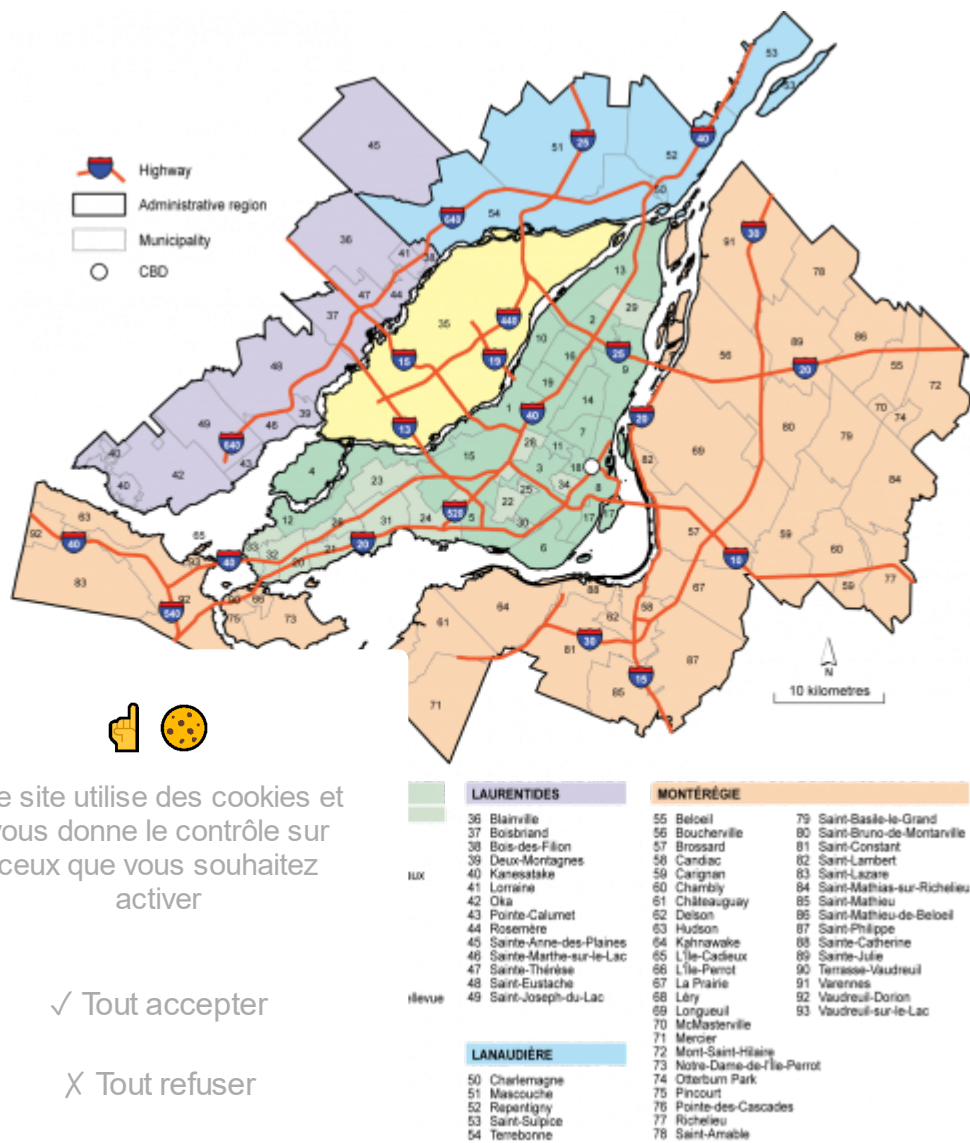
...d every five years between 1986 and 2006 using data ...ng the 'low-income cut-offs' variable calculated by ...corresponds to the income level at which a household

spend 20% or more of its income (before tax) on the basics (i.e. food, shelter and clothing) than does the average household of similar size (Statistics Canada, 2011).

21   This measure is the only one in the Canadian census that allows identifying low-income persons or families at a small geographical scale, e.g. census tracts (Apparicio *et al.*, 2007; Séguin *et al.*, 2012). Because comparing 'crude' poverty levels rates between census tracts and across time might be influenced by the changing economy (i.e. periods of recession or economic prosperity), poverty was modelled as a 'location quotient' so that, at each time point, the poverty rate of each census tract was divided by the rate observed for CMA as a whole; we are thus using a measure of relative poverty concentration. The proportion of low-income population in the CMA at each time point is showed in Table 2. The location quotient gives an overview of how, at any time, local poverty levels compare to the CMA average. This measure of concentration is largely used in urban and regional studies (Mikelbank, 2006; Shearmur *et al.*, 2008; Shearmur *et al.*, 2009; Vicino *et al.*, 2011; Walks *et al.*, 2008), and is computed as follow:

$$LQ_i = (x_i / t_i)/(X / T)$$

22   Where:

x_i = low income population in private households the census tract i;
t_i = total persons in private households in the census tract i;
X = low income population in private households in the CMA;
T = total persons in private households in the CMA.

23   A location quotient greater than 1 indicates a concentration of poverty (i.e. a percentage of low-income population higher than that of the CMA) whereas a value below 1 indicates an underrepresentation of poverty (i.e. a percentage of low-income population lower than that of the CMA).

Table 2. Description of indicators of poverty for the Montreal CMA between 1986 and 2006a

Census year	1986	1991	1996	2001	2006
Total population	2,826,270	3,019,350	3,125,545	3,208,860	3,363,975
Low-income population	609,175	666,680	863,745	723,670	728,220
	?1.55	22.08	27.64	22.55	21.65
	1.32	11.69	11.22	7.52	7.01
	5.92	15.73	17.55	18.23	18.24
	5.31	27.34	29.55	31.19	31.99
	.26	10.26	11.09	11.97	12.70
	.27	2.73	4.21	3.46	4.77
	9.76	34.96	31.51	25.87	21.61
	0.74	13.42	26.05	26.27	25.14
	5.54	53.65	52.19	50.45	47.51

CMA boundaries of 1986.

es: Population 15 years and over with less than grade 13
; For 2001 census: Population 20 years and over with less

than grade 13 without secondary school certificate; For 2006 census, population 15 years and over without diploma.

[c] Population 15 years and over, except for the 2001 census where this information is available for population aged 20 years and over.

LCGM and k-means clustering methods to identify trajectories of relative poverty

24   Generation of trajectories of neighbourhood poverty was first conducted using LCGM, as this technique provides diagnostic statistics regarding the optimal cluster solution. We set as an initial criterion that each cluster/trajectory needed a minimum of 5% of the 611 census tracts, so a minimum of 30 census tracts per trajectory. This was set to ensure a minimum of observations per trajectory in later validation stage (e.g. and in line with minimum requirement of observation for regression analysis). As we had no a priori for the optimal number of classes, LCGM was conducted for 5 to 20 classes; a minimum of 5 classes was set in order to have a minimum of differentiation between groups of census tracts. The optimal cluster solution is identified by: 1) a minimum of 5% of census tracts per trajectories; and 2) the lowest BIC value. Analyses were conducted in LatentGOLD software (Statistical Innovations).

25   K-means clustering was conducted in SAS 9.2 (SAS Institute Inc), specifying again 5 to 20 clusters. The value of the average distance to cluster centroid for each cluster solution was plotted to identify a 'natural break' in the distribution, indicating the optimal cluster solution. In the end, the choice of the optimal number of cluster solutions was informed by the LCGM solution providing the best fit to the data.

Statistical analyses

26   To assess the relative performance of the LCGM vs. the k-means clustering in identifying trajectories of relative poverty, two sets of analyses were conducted. First, in a multinomial logistic regression (MLR), the variables used for the classification (i.e. the location quotients from 1986 to 2006) were modelled as predictors of the trajectories (trajectories are modelled as a categorical dependent variable). This approach is some form of discriminant analysis, used to test the performance of ...n (see for example Magidson *et al.*, 2002) or different ...The idea here is to use the R-Square and model fit ... inform which of the k-means and LCGM cluster ... variation in poverty concentration.

... then conducted to empirically examine how a set of ...ciated with poverty explain each trajectory: ...rent families (%), one-person households (%), the ...nt immigrants (%), low education population (%), ...nters (%) (see Table 2 for a description of the values ...1986 and 2006 for the study region). These variables ...studies have demonstrated that they are strongly ...ribution of poverty across the Montreal CMA at the ...t al., 2007; Séguin *et al.*, 2012). In separate MLR ...odelled at the start of the period, i.e. 1986, at end, i.e. ...een 1986 and 2006 (e.g. unemployment rate 2006 - ...al model including baseline predictors and variation ...n. For each model, the focus is on the strength (R-

Ce site utilise des cookies et vous donne le contrôle sur ceux que vous souhaitez activer

✓ Tout accepter

✗ Tout refuser

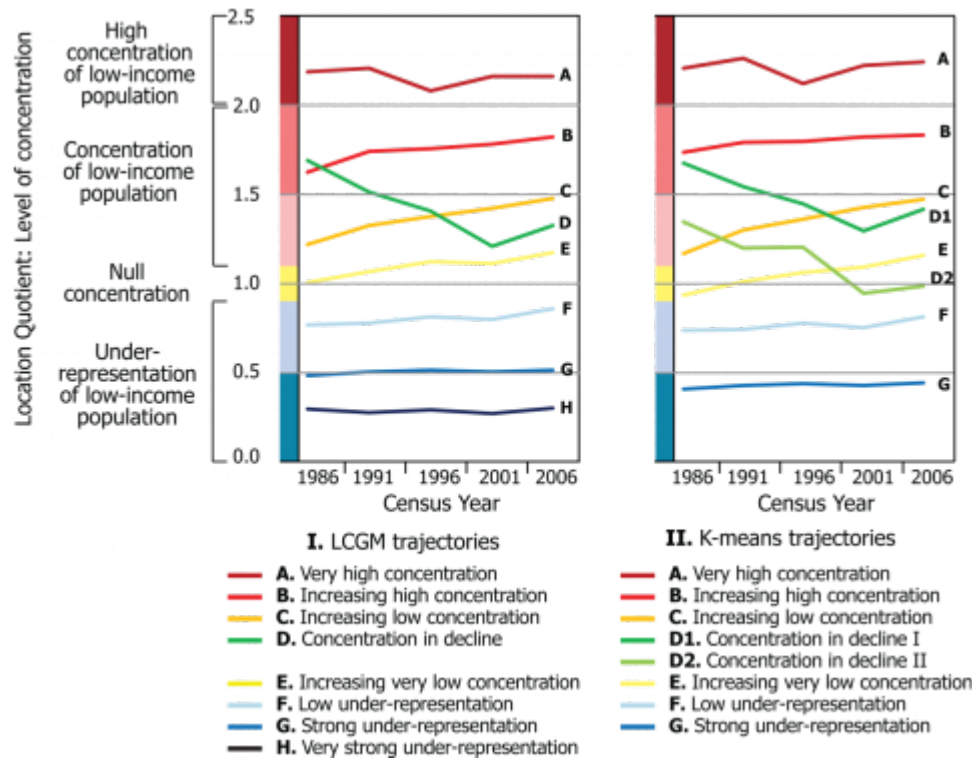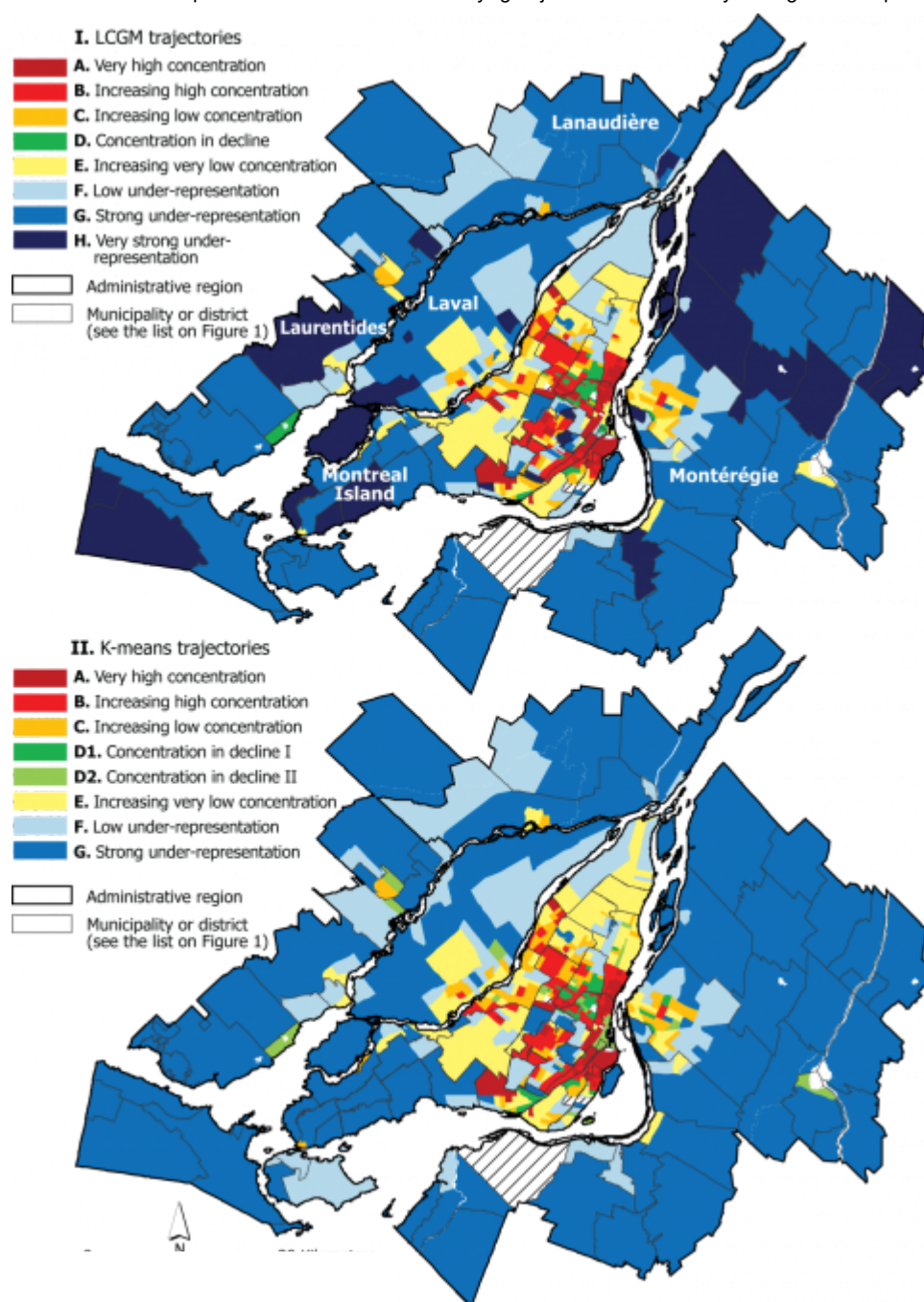Personnaliser

Politique de confidentialité

Square) and the fit (Akaike Information Criterion [AIC] and Bayesian Information Criterion [BIC]; lowest value of the AIC and BIC are indicative of better model fit) of the model.

Results

28   Location quotients of the low-income population for the five census years at the census tract level (CT) are mapped in Figure 2. As reported in previous studies (Apparicio *et al.*, 2007; Séguin *et al.*, 2012), CTs displaying a concentration of poverty are mainly located in the central part of the Island of Montreal, corresponding to inner-city neighbourhoods, whereas CTs characterized by an under-representation of poverty are observed in Laval and the North and South shores, corresponding to suburban areas. Over the 1986-2006 period, the presence of poverty in many central CTs (inner-city neighbourhoods) became less pronounced, while poverty gained grounds outside the central CTs on the island of Montreal in areas urbanized during the 1950's and 1960's. Some areas of concentrated poverty located in the northern periphery of the CMA disappeared over the period under study (these are old village centres in municipalities that have witnessed the arrival of new, wealthier populations).

Figure 2. Location quotients of relative poverty in Montreal CMA at census tract level (1986-2006).

👆🍪

Ce site utilise des cookies et vous donne le contrôle sur ceux que vous souhaitez activer

✓ Tout accepter

✗ Tout refuser

Personnaliser

Politique de confidentialité

Source: Statistics Canada.
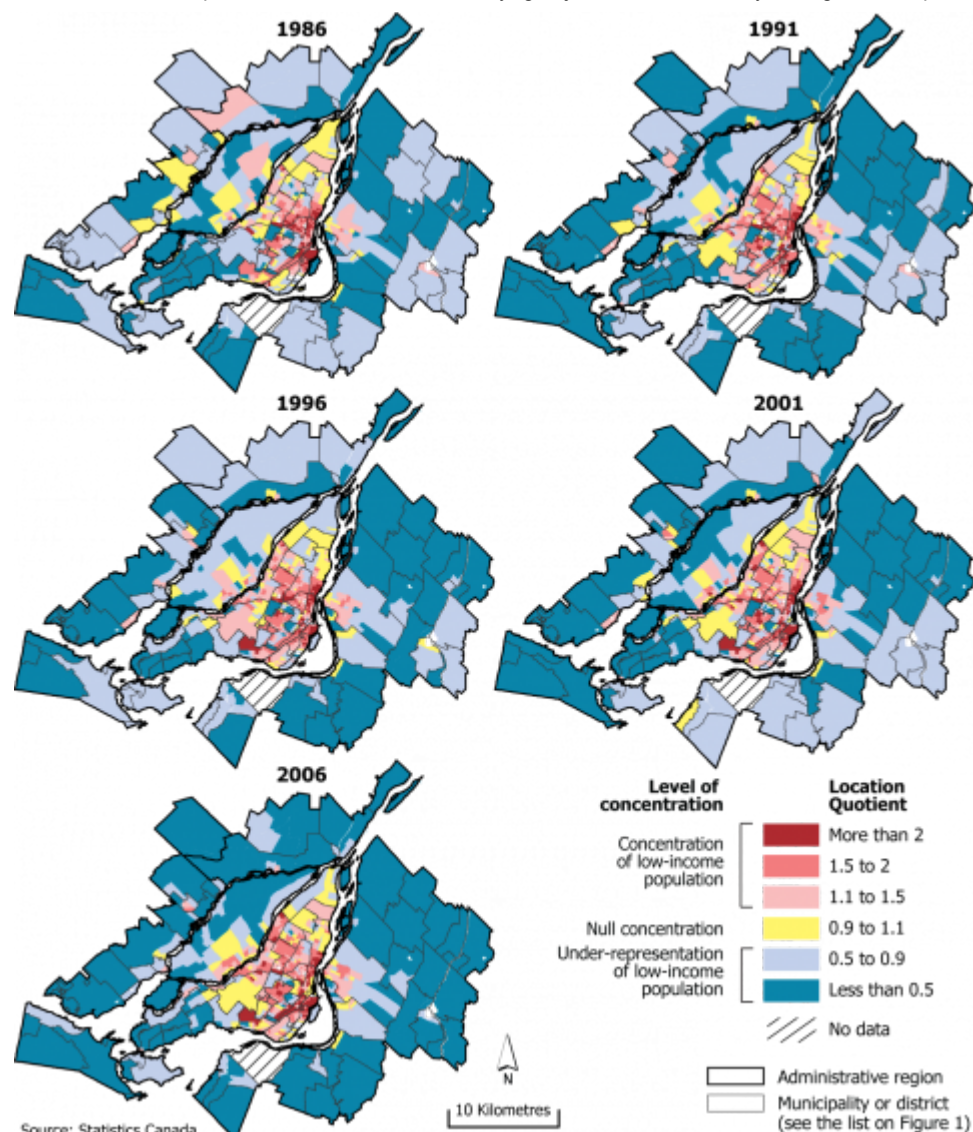
29   Optimal cluster solution for both the LCGM and k-means clustering are presented in Figure 3. According to the LCGM analysis, the 611 census tracts were optimally classified in eight trajectories, identified by the 'cluster solution' insuring a minimum of 5% census tracts per trajectory and with the lowest BIC values. The eight-cluster ... he data well in the k-means analysis. However, this ... up of areas comprising less than 5% of census tracts ... e elected to use the LCGM model with the best fit to ... al number of clusters, analyses are conducted on the ... e LCGM and k-means.

... nalysis, trajectories obtained from the k-means and ... ravity centres' of classes i.e. the mean values of the ... ent classes (Figure 3); the graph easily characterizes ... he cartography of census tracts according to their ... the Figure 4. As shown in Figure 4, both methods of ... identify groups of census tracts having followed ... creasing levels of poverty. This indicates changes in ... opulation in some neighbourhoods in the Montreal ... f census tracts, relative poverty levels did not change ... s indicated by groups of census tracts with relatively ... the period.

... by LCGM and k-means.
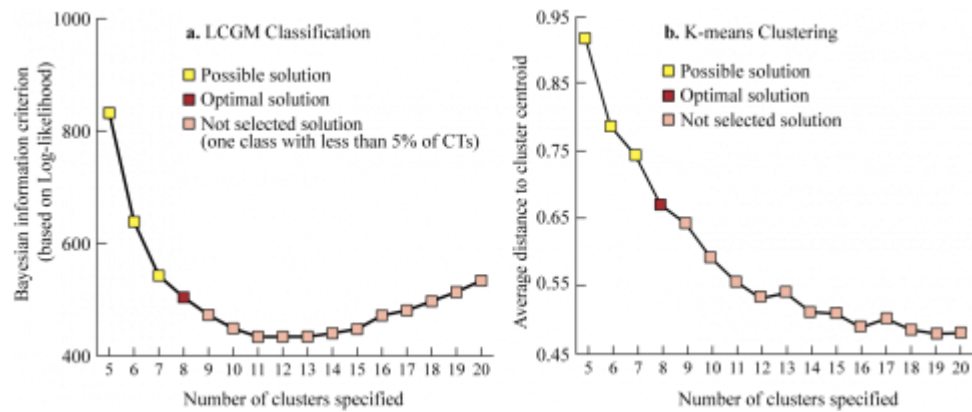
Figure 4. Trajectories of relative deprivation concentration obtained by k-means and LCGM methods.

s of relative deprivation concentration obtained by k-means

👆 🍪

Ce site utilise des cookies et
vous donne le contrôle sur
ceux que vous souhaitez
activer

✓ Tout accepter

✗ Tout refuser

Personnaliser

Politique de confidentialité

I. LCGM trajectories

- **A.** Very high concentration
- **B.** Increasing high concentration
- **C.** Increasing low concentration
- **D.** Concentration in decline
- **E.** Increasing very low concentration
- **F.** Low under-representation
- **G.** Strong under-representation
- **H.** Very strong under-representation
- Administrative region
- Municipality or district (see the list on Figure 1)

II. K-means trajectories

- **A.** Very high concentration
- **B.** Increasing high concentration
- **C.** Increasing low concentration
- **D1.** Concentration in decline I
- **D2.** Concentration in decline II
- **E.** Increasing very low concentration
- **F.** Low under-representation
- **G.** Strong under-representation
- Administrative region
- Municipality or district (see the list on Figure 1)

611 census tracts were cross-classified in different and LCGM cluster solutions; a rapid inspection of the shows that both methods produced relatively similar trend in poverty concentration in trajectory B is an in the k-means. Whereas there appears to be two erty concentration in k-means (D1 and D2), these one trajectory in LCGM (trajectory D). In addition, ween groups of low poverty concentration. Although olutions do not exactly replicate each other, there are results of the two approaches of classification. For LCGM trajectory D are distributed across k-means)2 (17.8%). Likewise, all observations in k-means ng trajectories G (70.4%) and H (29.6%) of LCGM.

sus tracts in the k-means and LCGM cluster solutions

	% (row) % (col.)	LCGM trajectories								
		A	B	C	D	E	F	G	H	Total
K-means trajectories	A	**65**	0	0	0	0	0	0	0	65
		100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	10.64
		80.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
	B	**16**	**66**	0	0	0	0	0	0	82
		19.5	**80.5**	0.0	0.0	0.0	0.0	0.0	0.0	13.42
		19.8	**91.7**	0.0	0.0	0.0	0.0	0.0	0.0	
	C	0	**3**	**60**	0	**12**	0	0	0	75
		0.0	**4.0**	**80.0**	0.0	**16.0**	0.0	0.0	0.0	12.27
		0.0	**4.2**	**87.0**	0.0	**12.2**	0.0	0.0	0.0	
	D1	0	**3**	**7**	**37**	0	0	0	0	47
		0.0	**6.4**	**14.9**	**78.7**	0.0	0.0	0.0	0.0	7.7
		0.0	**4.2**	**10.1**	**82.2**	0.0	0.0	0.0	0.0	
	D2	0	0	**2**	**8**	**14**	**2**	0	0	26
		0.0	0.0	**7.7**	**30.8**	**53.9**	**7.7**	0.0	0.0	4.3
		0.0	0.0	**2.9**	**17.8**	**14.3**	**2.0**	0.0	0.0	
	E	0	0	0	0	**72**	**13**	0	0	85
		0.0	0.0	0.0	0.0	**84.7**	**15.3**	0.0	0.0	13.91
		0.0	0.0	0.0	0.0	**73.5**	**13.0**	0.0	0.0	
	F	0	0	0	0	0	**85**	**11**	0	96
		0.0	0.0	0.0	0.0	0.0	**88.5**	**11.5**	0.0	15.7
			0.0	0.0	0.0	0.0	**85.0**	**10.4**	0.0	
		0	0	0	0	0	**95**	**40**		135
			0.0	0.0	0.0	0.0	0.0	**70.4**	**29.6**	22.1
			0.0	0.0	0.0	0.0	0.0	**89.6**	**100.0**	
			72	69	45	98	100	106	40	611
		.3	11.8	11.3	7.4	16.0	16.4	17.4	6.6	100.0

…ed (Figure 4), discrepancies in classification of census …GM trajectories appear minimal. However, a clear …levels of concentration of poverty as illustrated by …f the suburbs located in the North and South shores of …Laurentides, Lanaudière and Montégérie). In both k-…A and B capture the traditional poverty of inner city …city of Montreal (districts of Ville-Marie, Mercier-…eray-Saint-Michel-Parc-Extension, Côte-des-Neiges-

Notre-Dame-de-Grâce, Le Sud-Ouest); these areas are characterized by trajectories of high concentration of low-income population. It is worth noting that trajectory B is increasing during the period, indicating that in this group of census tracts, concentration of poverty has increased between 1986 and 2006. Census tracts in trajectories C and E (k-means and LCGM) are characterized by increasing poverty concentration over the period; they are mainly located in inner suburbs (municipalities of Saint-Laurent, Saint-Léonard. Anjou, Montréal-Est, Montréal-Nord, LaSalle), characterized by null concentration of poverty in 1986, but increasing over the period. This supports other studies reporting a suburbanization of poverty in some North American metropolitan areas (Cooke *et al.*, 2006; Lee, 2011; Madden, 2003). Trajectories D1 and D2 produced by k-means and trajectory D produced by LCGM correspond to gentrifying census tracts located in inner city areas, mostly in the districts of Plateau-Mont-Royal and Rosemont-La Petite-Patrie (Bourne, 1993; Ley, 1988; Rose, 1984; Walks *et al.*, 2008). These groups of census tracts are characterized by decreasing concentration of poverty over the past 20 years. Finally, trajectory F and G in k-means and F, G and H in LCGM group census tracts characterized by lower poverty levels; they are essentially located in suburban areas of the CMA (Laval, North and South Shores and the West of the Island of Montreal) and in traditionally wealthier municipalities in the central part of the Island of Montreal (e.g. Westmont, Outremont and Town Mont-Royal).

33   Results of multinomial logistic regressions further inform on which of the k-means and LCGM performs better in identifying trajectories of neighbourhood change. To do so, we only present fit statistics from MLR models (R-Square, max-rescaled R-Square, AIC and BIC); results are reported in the Table 4. Model A examines how the location quotients measured at the five time points (and used for the classification) predict 'belonging' to a trajectory. Results show that LCGM very marginally outperformed k-means as shown by lowest AIC and BIC values. This indicates that LCGM might be more accurate in identifying groups of census tracts having followed similar trajectories of change over the period.

34   When other predictors of poverty were used to predict trajectories (models B to E), LCGM again seems to perform better than k-means. Yet the difference between models is again very modest. Indeed, as shown in Table 4, the R-Square of multinomial regression are very strong for both the k-means and LCGM models, given that the predictors of the models are known to be associated with poverty (excessive multicolinearity was not a problem in our models, as indicated by variance inflation ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ t reported here). R-Square values are essentially the ~~~~~~~~~~~~~~~~~~~~~~~~~~~~ e, R-Squares for model A are 0.915 for k-means and ~~~~~~~~~~~~~~~~~~~~~~~~~~~ ept for the model D, all values of AIC and BIC are lower

	K-means	LCGM
...ltinomial logistic regression models		
...to 2006)		
	0.980	0.982
...e)	0.998	0.999
	136.792	114.645
...on Criterion)	322.226	300.079
...sured in 1986		

R-Square	0.915	0.925
Max-rescaled R-Square	0.932	0.941
AIC	1057.340	1025.564
BIC	1335.491	1303.715

Model C: Indicators of poverty measured in 2006

R-Square	0.896	0.907
Max-rescaled R-Square	0.913	0.923
AIC	1180.815	1154.890
BIC	1458.966	1433.042

Model D: Variations in indicators of poverty between 1986 and 2006

R-Square	0.736	0.745
Max-rescaled R-Square	0.750	0.758
AIC	1751.644	1773.066
BIC	2029.795	2051.217

Model E: Indicators of poverty measured in 1986 +

variations in indicators of poverty between 1986 and 2006

R-Square	0.950	0.955
Max-rescaled R-Square	0.968	0.972
AIC	842.559	828.084
BIC	1367.955	1353.481

Discussion and conclusion

👆 🍪

Ce site utilise des cookies et vous donne le contrôle sur ceux que vous souhaitez activer

✓ Tout accepter

✗ Tout refuser

Personnaliser

Politique de confidentialité

ompare two clustering techniques, k-means clustering
ls, in order to better characterize trajectories of
period of twenty years. Thus, the main contribution of
ghbourhood change is methodological, contributing
clustering techniques to longitudinal data at the area
th k-means and LCGM techniques were successful in
erized by similar trajectories of change in poverty
As observed in other studies, different trajectories of
lontreal CMA were captured, including trajectories of
me population in inner city areas; suburban areas
centration of poverty over the twenty-year period.
tudies reporting a suburbanization of poverty in some
areas (Cooke *et al.*, 2006; Madden, 2003); the
reas (Heisz *et al.*, 2006; Ley, 1986, 1988; Walks *et al.*,
verty concentration in other parts of the central city
urne, 1993; Madden, 2003). However, most census
ewhat stable poverty levels over the period, either in

terms of stable but high levels of poverty concentration, or stable but low representation of poverty. Yet, most studies of neighbourhood trajectories are mainly concerned with change, i.e. either increase or decrease in poverty levels. Nevertheless, our study shows that some areas have been persistently experiencing high levels of poverty between 1986 and 2006, whereas in others, poverty levels are continuously lower compared to the CMA as a whole.

37   Overall, the results of the two cluster solutions were relatively similar. However, we found some discrepancies between the maps of trajectories of relative deprivation obtained by k-means and LCGM methods. For the census tracts located in the suburbs, the LCGM solution identified three stable trajectories with low poverty levels against two for the k-means. In the central areas, the LCGM solution grouped into one trajectory 45 census tracts characterized by a decline of poverty concentration as a result of a gentrification process. Nonetheless, these 45 census tracts are divided into two trajectories of decreasing poverty concentration in k-means solution. In other words, the LCGM cluster solution provided more details in the suburbs while the k-means provided more details in central areas.

38   Finally, results of the multinomial logistic regressions showed that LCGM marginally outperformed k-means clustering. Since nowadays the LCGM method is more complex to implement than k-means clustering, the k-means still appears to be a robust and relevant approach to estimate trajectories of neighbourhood change where the clustering variable is measured continuously (e.g. location quotients). It should be recalled, however, that the k-means clustering could only be applied on continuous or binary variables whereas LCGM allows the identification of trajectories using variables with a range of different distributions (continuous, ordinal, nominal, count and binomial). Moreover, with the LCGM, it is also possible to maximize trajectories as a function of external variables (such as the predictors examined here for their association with the trajectories). Consequently, the LCGM approach is promising given its broader application. For instance, LCGM could be used to identify trajectories of neighbourhood change based on a count variable or a qualitative variable. Future studies are needed to test applications of LCGM to model neighbourhood change.

Bibliographie

👍 🍪

Ce site utilise des cookies et vous donne le contrôle sur ceux que vous souhaitez activer

✓ Tout accepter

✗ Tout refuser

Personnaliser

Politique de confidentialité

s between data clustering algorithms", *The International* ... *logy*, vol. 5, nᵒ 3, 320-325.

A., Gaudreau P., Louvet B., 2009, "Latent Class Growth ... *Quantitative Methods for Psychology*, vol. 5,11-24.

up X., 2007, "Modélisation spatiale de la pauvreté à ... e de la régression géographiquement pondérée", *The* ... *he canadien*, vol. 51, nᵒ 4, 412-427. 00189.x

., 2012, "Identifying, mapping and modelling trajectories ... od level: The case of Montréal, 1986-2006", *Applied*

C.-L., Katzmarzyk P. T., 2008, "Distinct trajectories of ... d predictors of trajectory class membership: a 22 year ... *l of Behavioral Nutrition and Physical Activity*, vol. 5:57,

utures: An American Dilemma Revisited", *Transactions of* ... vol. 5, nᵒ 1, 1-28.

Bourne L. S., 1993, "The myth and reality of gentrification: a commentary on emerging urban forms", *Urban Studies*, vol. 30, n° 1, 183-189.
DOI : 10.1080/00420989320080111

Brookmeyer K. A., Henrich C. C., 2009, "Disentangling adolescent pathways of sexual risk taking", *Journal of Primary Prevention*, vol. 30,677-696.
DOI : 10.1007/s10935-009-0196-6

Bunting T. E., Filion P., 1988, "Introduction: the movement towards the post-industrial society and the changing role of the inner city", in Bunting T. E., Filion P. (eds), *The changing Inner City*, Waterloo, University of Waterloo, 1-24.

Calińskia T., Harabasza J., 2007, "A dendrite method for cluster analysis", *Communications in Statistics*, vol. 3, n° 1, 1-27.

Clark W. A. V., 1987, "The Roepke Lecture in Economic Geography Urban Restructuring from a Demographic Perspective", *Economic Geography*, vol. 63, n° 2, 103-125.
DOI : 10.2307/144149

Collins L. N., Lanza S. T., 2009, *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*, Hoboken (New Jersey), John Wiley & Sons.
DOI : 10.1002/9780470567333

Cooke T., Marchant S., 2006, "The changing intrametropolitan location of high-poverty neighbourhoods in the US, 1990-2000", *Urban Studies*, vol. 43, n° 11, 1971-1989.
DOI : 10.1080/00420980600897818

Duncan T. E., Duncan S. E., Strycker L. A., 2009, *An Introduction to Latent Variable Growth Curve Modeling: Concepts, Issues, and Applications*, Mahwah (New Jersey), Lawrence Erlbaum Associates.
DOI : 10.4324/9780203879962

Everitt B., Landau S., Leese M., 2001, *Cluster Analysis*, London, Heineman Educational Books.
DOI : 10.1016/0191-8869(87)90161-9

Friedman M., Kandel A., 1998, *Introduction to pattern recognition : Statistical, structural, neural and fuzzy logic approaches*, Singapore, World Scientific Publishing Company.

Heisz A., McLeod L.,2006, *Low-income in Census Metropolitan Areas, 1980-2000.* Ottawa, Minister of Industry.

Jain A. K., 2007, "Data clustering: 50 years beyond K-means", *Pattern Recognition Letters*, vol. 31, n° 8, 651–666.
DOI : 10.1016/j.patrec.2009.09.011

Jargowsky P., 2003, *Stunning Progress, bidden problems: The dramatic decline of concentrated poverty in the 1990's,* Washington D.C., The Brookings Institution.

Jones B. L., Nagin D. S., 2007, "Advances in group-based trajectory modeling and an SAS procedure for estimating them", *Sociological Methods & Research*, vol. 35, n° 4, 542-571.
DOI : 10.1177/0049124106292364

005, *Finding groups in data: An introduction to cluster* s in Probability and Statistics.

"Measuring neighborhood social change in Saskatoon, *rban Geography*, vol. 30, n° 3, 261-288.

ra-metropolitan poverty differentiation: causes and to suburbs in the metropolitan Atlanta region", *Annals of* 57.
5

netropolitan spatial differentiation and decline of inner- ur US metropolitan areas", *Journal of Planning Education* 54.

planations for Inner-City Gentrification: A Canadian *ion of American Geographers*, vol. 76, n° 4, 521-535.
tb00134.x

six Canadian inner cities", *The Canadian Geographer / Le* 31-45.
tb00853.x

ng spatial concentration of income and poverty among reas", *Urban Studies*, vol. 40, n° 3, 481-503.

DOI : 10.1080/0042098032000053888

Magidson J., Vermunt J. K., 2002, "Latent class models for clustering: A comparison with K-means", *Canadian Journal of Marketing Research*, vol. 20, 37-44.

McConville S., Ong P., 2003, *The trajectory of poor neighborhoods in Southern California, 1970-2000,* Washington, DC, Center on Urban and Metropolitan Policy, The Brookings Institution.

Mikelbank B. A., 2004, "A typology of U.S. suburban places", *Housing Policy Debate*, vol. 15, 935-964.
DOI : 10.1080/10511482.2004.9521527

Mikelbank B. A., 2006, "Local Growth Suburbs: Investigating Change within the Metropolitan Context", *Opolis*, vol. 2, nº 1, 1-15.

Mikelbank B. A., 2011, "Neighborhood déjà Vu: Classification in metropolitan Cleveland, 1970-2000", *Urban Geography*, vol. 32, nº 3, 317-333.
DOI : 10.2747/0272-3638.32.3.317

Milligan G. W., Cooper M. C., 1985, "An examination of procedures for determining the number of clusters in a data set", *Psychometrika*, vol. 50, nº 2, 159-179.
DOI : 10.1007/BF02294245

Muthén & Muthén *Mplus*. Los Angeles, CA, USA.

Nagin D. S., 2005, *Group-based modelling of development*, Cambridge, MA, Harvard University Press.

Pearson A. L., Apparicio P., Riva M., 2013, "Cumulative disadvantage? Exploring relationships between neighbourhood deprivation trends (1991 to 2006) and mortality in New Zealand", *International Journal of Health Geographics*, vol. 12, nº 38, 1-8.

Reibel M., Regelson M., 2007, "Quantifying neighborhood racial and ethnic transition clusters in multiethnic cities", *Urban Geography*, vol. 28, nº 4, 361-376.
DOI : 10.2747/0272-3638.28.4.361

Reibel M., Regelson M., 2011, "Neighborhood racial and ethnic change: the time dimension in segregation", *Urban Geography*, vol. 32, nº 3, 360-380.
DOI : 10.2747/0272-3638.32.3.360

Riva M., Curtis S. E., 2012, "Long-term local area employment rates as predictors of individual mortality and morbidity: A prospective study in England, spanning more than two decades", *Journal of Epidemiology and Community Health*,

Rose D., 1984, "Rethinking gentrification beyond the uneven development of marxist urban theory", *Environment and Planning D : Society and Space*, vol. 12, nº 1, 47-74.
DOI : 10.1068/d020047

Sarle W. S., 1983, *Cubic Clustering Criterion,* Cary, NC, SAS Institute Inc.

SAS Institute Inc *SAS version 9.2.* Cary, NC, USA.

M., 2012, "The Impact of geographical scale in identifying ed interventions to tackle poverty: the case of Montréal", vol. 5, nº 3, 231-251.
6

, "Urban hierarchy or local buzz? high-order producer isive business service location in Canada, 1991-2001", º 3, 333-355.

eak ties that bind: Do commutes bind Montreal's central *Affairs Review*, vol. 44, nº 4, 490-524.
92

., 2007, "The decline of inner suburbs: The new suburban *aphy Compass*, vol. 1/3, 641-656.
00020.x

D® 4.5. Belmont, MA, USA.

, *Quebec (Code462) (table). 2006 Community Profiles. 2006* Catalogue no. 92-591-XWE.

ncome Lines, 2009-2010, Ottawa, Minister of Industry

Tibshirani R., Walther G., Hastie T., 2001, "Estimation of number of clusters in a data set via the gap statistic", *Journal of the Royal Statistical Society: Series B*, vol. 63, n° 2, 411-423.

Van Kempen R., Murie A., 2009, "The new divided city: changing patterns in European cities", *Tijdschrift voor Economische en Sociale Geografie*, vol. 100, n° 4, 377-398. DOI : 10.1111/j.1467-9663.2009.00548.x

Vicino T., Hanlon B., Short J., 2011, "A typology of urban immigrant neighborhoods", *Urban Geography*, vol. 32, n° 3, 383-405. DOI : 10.2747/0272-3638.32.3.383

Vicino T. J., 2008, "The spatial transformation of first-tier suburbs, 1970 to 2000: The case of metropolitan Baltimore", *Housing Policy Debate*, vol. 19, n° 3, 479-518.
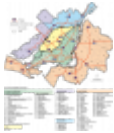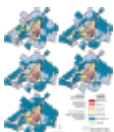
Walks R. A., 2001, "The social ecology of the post-fordist/global city? Economic restructuring and socio-spatial polarisation in the Toronto urban region", *Urban Studies*, vol. 38, n° 3, 407-447. DOI : 10.1080/00420980120027438

Walks R. A., Maaranen R., 2008, "Gentrification, social mix, and social polarization: Testing the linkages in large Canadian cities", *Urban Geography*, vol. 29, n° 4, 293-326. DOI : 10.2747/0272-3638.29.4.293

Table des illustrations

	Titre	Figure 1. The study area.
	URL	http://journals.openedition.org/cybergeo/docannexe/image/27035/img-1.png
	Fichier	image/png, 649k
	Titre	Figure 2. Location quotients of relative poverty in Montreal CMA at census tract level (1986-2006).
	URL	http://journals.openedition.org/cybergeo/docannexe/image/27035/img-2.png
	Fichier	image/png, 897k
	Titre	Figure 3. Cluster solutions identified by LCGM and k-means.
	URL	http://journals.openedition.org/cybergeo/docannexe/image/27035/img-3.png
	Fichier	image/png, 101k
	Titre	Figure 4. Trajectories of relative deprivation concentration obtained by k-means and LCGM methods.

...opnedition.org/cybergeo/docannexe/image/27035/img-

55k

...ography of trajectories of relative deprivation concentration ...means and LCGM methods.

...openedition.org/cybergeo/docannexe/image/27035/img-

12k

...ne-Marie Séguin, « A comparison of two methods for ...on neighborhood poverty at the intra-metropolitan level in ...*urnal of Geography* [En ligne], Espace, Société, Territoire, ...n 2015, consulté le 28 janvier 2022. URL : ...geo/27035 ; DOI : https://doi.org/10.4000/cybergeo.27035

Cet article est cité par

- Libório, Matheus Pereira. Martinuci, Oseias da Silva. Machado, Alexei Manso Correa. Lyrio, Renata de Mello. Bernardes, Patrícia. (2021) Time–Space Analysis of Multidimensional Phenomena: A Composite Indicator of Social Exclusion Through k-Means. *Social Indicators Research*. DOI: 10.1007/s11205-021-02763-y

- Ajayi, Olasupo O.. Bagula, Antoine B.. Maluleke, Hloniphani C.. (2020) Africa 3: A Continental Network Model to Enable the African Fourth Industrial Revolution. *IEEE Access*, 8. DOI: 10.1109/ACCESS.2020.3034144

- Gelb, Jérémy. Apparicio, Philippe. (2021) Apport de la classification floue c-means spatiale en géographie : essai de taxinomie socio-résidentielle et environnementale à Lyon. *Cybergeo*. DOI: 10.4000/cybergeo.36414

- Ling, Chenjun. Delmelle, Elizabeth C. (2016) Classifying multidimensional trajectories of neighbourhood change: a self-organizing map andk-means approach. *Annals of GIS*. DOI: 10.1080/19475683.2016.1191545

Auteurs

Philippe Apparicio
Professor
INRS Urbanisation Culture Société
385 rue Sherbrooke Est, Montréal (QC), H2X 1E3, Canada
Philippe_apparicio@ucs.inrs.ca

Articles du même auteur

Apport de la classification floue c-means spatiale en géographie : essai de taxinomie socio-résidentielle et environnementale à Lyon [Texte intégral]
Article 972
Paru dans *Cybergeo: European Journal of Geography*, Cartographie, Imagerie, SIG

Un atlas-web pour comparer l'exposition individuelle aux pollutions atmosphérique et sonore selon le mode de transport [Texte intégral]
Article 903
Paru dans *Cybergeo: European Journal of Geography*, GeOpenMod - Modèles et logiciels

👍 🍪

Ce site utilise des cookies et vous donne le contrôle sur ceux que vous souhaitez activer

✓ Tout accepter

✗ Tout refuser

Personnaliser

Politique de confidentialité

ée dans huit métropoles canadiennes de 1981 à 2006 : un gué [Texte intégral]

al of Geography, Espace, Société, Territoire

pplication for calculating residential segregation indices

al of Geography, Systèmes, Modélisation, Géostatistiques

ur Internet : nouvelles avenues avec le *Scalable Vector* intégral]

al of Geography, Dossiers

entielle : un outil intégré dans un système d'information

al of Geography, Espace, Société, Territoire

Medicine, Population Health and Optimal Practices in Health

Research Unit, CHU de Québec Research Centre, Université Laval, Canada
Mylene.Riva@crchudequebec.ulaval.ca

Anne-Marie Séguin
Professor
INRS Urbanisation Culture Société
385 rue Sherbrooke Est, Montréal (QC), H2X 1E3, Canada
anne-marie_seguin@ucs.inrs.ca

Articles du même auteur

La répartition de la population âgée dans huit métropoles canadiennes de 1981 à 2006 : un groupe de moins en moins ségrégué [Texte intégral]
Article 639
Paru dans *Cybergeo: European Journal of Geography*, Espace, Société, Territoire

---

Droits d'auteur

👆🍪

Ce site utilise des cookies et vous donne le contrôle sur ceux que vous souhaitez activer

✓ Tout accepter

✗ Tout refuser

Personnaliser

Politique de confidentialité