# Hybridising geographically weighted regression and multilevel models: a new approach to capture contextual effects in geographical analyses

Thierry Feuillet, Etienne Edouard Cossart, Hélène Charreire, Arnaud Banos, Hugo Pilkington, Virginie Chasles, Serge Hercberg, Mathilde Touvier, Jean-Michel Oppert

# Hybridising geographically weighted regression and multilevel models: a new approach to capture contextual effects in geographical analyses

T. Feuillet[1,2], E. Cossart[3], H. Charreire[2,4], A. Banos[5], H. Pilkington[6] H., V. Chasles[3], S. Hercberg[2], M. Touvier[2], J.M. Oppert[2,6]

[1] University of Caen, CNRS – UMR 6266 IDEES, Caen, France
[2] Université Sorbonne Paris Nord and Université Paris Cité, INSERM, INRAE, CNAM, Center of Research in Epidemiology and StatisticS (CRESS), Nutritional Epidemiology Research Team (EREN), 93017, Bobigny, France
[3] University Jean Moulin Lyon 3, CNRS – UMR 5600 Environnement Ville Société, Lyon, France
[4] MoISA, Univ Montpellier, CIRAD, CIHEAM-IAMM, INRAE, Institut Agro, IRD, Montpellier, France
[5] CNRS – UMR 6266 IDEES, Le Havre, France
[6] Department of Nutrition, Human Nutrition Research Center Ile-de-France (CRNH IdF), Pitié-Salpêtrière Hospital (AP-HP), Sorbonne University, Paris, France

**Abstract**

Multilevel models are one of the main statistical methods used in modelling contextual effects in social sciences. A common limitation of these methods is the use pre-set boundaries – usually administrative units – to define contexts, when these boundaries do not always match up with the “true” causally relevant contexts that may affect the outcomes of interest. In this study applied to the obesity geography in the Paris area (France), we propose a new spatially explicit two-step procedure to tackle this methodological issue. The first step consists in estimating a geographically weighted regression (GWR) model, then using it to reveal and delineate relevant non-stationarity-based data-driven spatial contexts, and finally including them as a random effect into a random slope multilevel model. In applying this hybrid methodology for modelling body mass index (BMI) within a sample of 9,089 French adults, we demonstrate that it outperforms administrative-based multilevel models in terms of decreasing Akaike information criteria (AIC), and is better at accounting for contextual effects through intraclass correlation coefficient (ICC) and increasing slope variance. We suggest that this procedure might be generalized to quantitative geographical analyses involving contextual effects.

## Introduction

Geospatial data analysts often face two major challenges when conducting quantitative geographical analyses. These challenges relate to the two main geographical effects that characterize all geographical data: the spatial effect (mainly controlled by distance) and the contextual (or 'platial' when contexts are geographical) effect (Arcaya et al., 2012; Wolf et al., 2021). From a theory-driven perspective, the spatial effect relates to spatial dependence, according to which the strength of spatial interaction between two things depends on their mutual distance (Tobler, 1970). The contextual or platial effect, on the other hand, is related to spatial heterogeneity – proposed by Goodchild (2004) to be the second law of Geography – since this effect is directly due to the fact that spatial contexts differ from one another, and therefore that belonging to a given spatial context implies specific consequences on objects and processes operating within it. From a statistical viewpoint, spatial dependence is usually linked to spatial autocorrelation, while spatial heterogeneity leads to spatial nonstationarity, i.e.~~,~~ to means, variances and covariances that vary from a place to another.

The demarcation line between spatial effects and contextual effects is not always a clear one, but a fundamental distinction between the two effects can be drawn depending on how we consider geographical space: either as continuous (thus revealing distance effects), or as discrete (thus capturing contextual effects). This dichotomy echoes the distinction between space and place discussed by human geographers for many years (Kearns & Joseph, 1993; Roche, 2016, Wolf et al., 2021). The place perspective uses location to form groups, while the spatial perspective looks at proximity between observations. As raised by Arcaya et al. (2012), the respective merits of these two approaches are rarely compared for one given dataset.

This dichotomy is also to be found in statistical techniques that are commonly used to model geographical data. Two major families of statistical modelling frameworks are used by most studies dealing with geographical effects. The first encompasses, within a place-based perspective, all the mixed models that include spatial context considered as a random effect (the so-called hierarchical or multilevel models, Goldstein, 2011). The second comprises, within a space-based perspective, spatially explicit models including spatially varying coefficient models, i.e., those considering geographical coordinates for each observation to estimate relationships, such as geographically weighted regressions (GWR, Brunsdon et al., 1996) upon which we will focus here.

From a theoretical viewpoint, each of these families of models has strengths and limitations. Multilevel models, popular in the education, social and health sciences since the 1990s

(Chaix et al., 2005; Duncan et al., 1998), provide through their hierarchical structure a sound inferential framework for modelling contextual effects while simultaneously controlling for individual effects. Moreover, random slope models allow the capture of heterogeneity in direction or intensity of relationships between places (Jones et al., 1992). Thus, these multilevel models are able to avoid both ecological and atomistic fallacies, as outlined by Owen et al. (2016). A major pitfall of these models is that the boundaries of spatial contexts must be established before the estimation of the model, and these boundaries determine in turn potential contextual effects. In the social sciences, administrative boundaries are commonly used to define spatial contexts and thus to group observations. Contextual effects can thus be over- or underestimated if administrative boundaries do not match with the actual and relevant geographical contexts related to the outcome under study (Merlo et al., 2005). According to Owen et al. (2016), this problem refers to spatial design. Many scholars, particularly in the field of neighbourhood effect studies, have outlined this spatial mismatch between administrative neighbourhood boundaries used by default as spatial contexts because of ease of access, and the true causally relevant context (Diez-Roux & Mair, 2010; Petrović et al., 2020). According to the modifiable areal unit problem (MAUP), such a spatial mismatch may have considerable consequences on subsequent statistical associations, including in multivariate models (Fotheringham and Wong, 1991), leading to seek out the “optimal” zoning prior to statistical modelling (Openshaw, 1978).

On the other hand, the increasingly popular GWR constitutes a powerful solution to explore spatial nonstationarity without any preset geographical contexts. Based on local regressions, GWR provides maps of regression coefficients that can be further used to identify spatial contexts with specific and sometimes opposing trends in relationships (Fotheringham et al., 2002). However, this method does not allow us to account for the aggregation of observations within places through variance component decomposition, like multilevel models do. Therefore, it ignores the effect of “belonging together” among those observations that fall within geographic boundaries (Figure 1).

To overcome the limitations of these methods, while retaining their strengths, this study investigates the implementation of a methodological workflow combining the two approaches into a two-step procedure. The key idea is to (i) delineate GWR-based spatial contexts, (ii) incorporate them into a multilevel modelling framework as a random effect, and (iii) compare the resulting model, in terms of statistical quality and fit, with the one based on administrative boundaries. These GWR-based spatial contexts are hypothesized to capture “true” causal spatial contexts (i.e., those that actually do affect human behaviour but that remain unknown) better than administrative boundaries, within a data-driven perspective

(Fotheringham & Sachdeva, 2022).

Some studies have previously explored varying degrees of hybridization between spatial and multilevel models, but not in the to the extent we describe here. For instance, Chen & Truong (2012) studied socio-ecological determinants of obesity in Taiwan by first estimating a multilevel model to adjust mean odds ratios (ORs) by administrative units, then using these ORs as a dependent variable in a GWR. Other authors used spatial models as a first step to better define residual spatial dependence structures in multilevel models (Dong et al., 2016; Janko et al., 2019; Park & Kim, 2014) or to account for spatial dependence in adding spatial autoregressive terms (Dong & Harris, 2015). Recently, Hu et al. (2022) proposed a combination of multilevel and GWR models (abbreviated as HLM-GWR) for modelling spatial data with identical geographical coordinates. In their study on real estate prices in China, a multilevel model was used to consider observations within the same locations as grouped, while simultaneously capturing spatially varying relationships of the group-level variables through GWR.

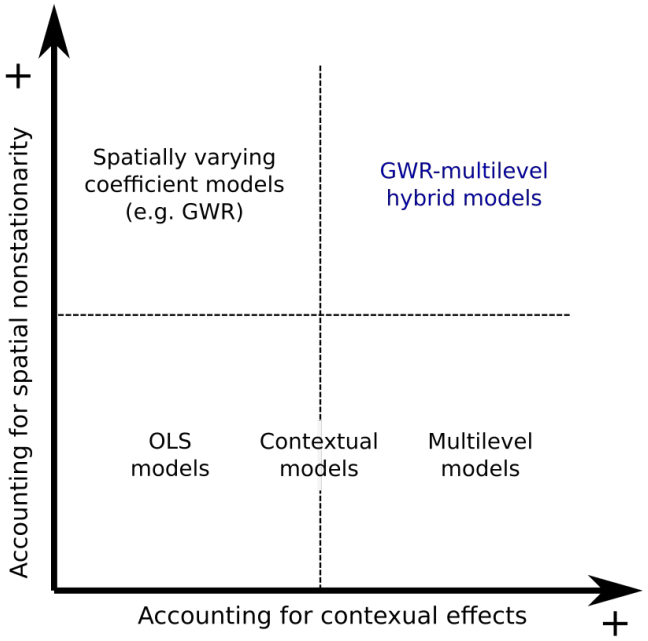


Figure 1. Ability of common statistical models used in geographical analysis to consider spatial non-stationarity and contextual effects. The GWR-multilevel hybrid procedure proposed in this study can simultaneously handle spatial nonstationarity and contextual effects. Note that “contextual models” refers to ordinary least square (OLS) models including contextual aggregated variables as predictors.

Closer to our study, Arcaya et al. (2012) incorporated spatially explicit information into their

multilevel models to define spatial contexts in a study on life expectancy in the US. Rather than considering counties (level-1) within states (level-2), they replaced the administrative level of US states with spatial patches solely based on proximity around counties. However, they still based their neighborhood structure scheme (level-1) on administrative units (counties).

The main originality of our article resides in the willingness to overcome pre-set geographical boundaries to construct spatial contexts. These contexts are built in a two-stage data-driven perspective, (i) First, letting GWR exhibit spatial non-stationarity, using it to draw contexts independently of administrative boundaries and (ii) including the resulting outputs in a multilevel model as a grouping variable.

This methodological procedure was applied to the geographical variation in a well-established indicator of obesity using data from participants of the French web-based Nutrinet-Santé cohort dataset residing in the Paris region. In this cohort, each participant was located at the residential address. Obesity is recognized as an important public health issue leading to increased morbidity and mortality (WHO Europe Obesity Report, 2022) and the study of obesity distribution fits well with a contextual analysis including both space and place effects. Indeed, a large body of research has highlighted that the prevalence of obesity reveals spatial patterns that differ according to scale (Swinburn et al., 1999), and that both direct (e.g. individual obesity-related behaviors such as diet and physical activity) and indirect (e.g. built and social environment) obesity drivers are spatially non-stationary (Chen & Truong, 2012; Feuillet et al., 2015, 2020; Oshan et al., 2020).

In the following sections, we first briefly describe the overall methodological workflow, before describing each step in more detail, as well as the data we use, and finally reporting and discussing the main results and possible extensions of the proposed methodology.

**Overall methodological workflow**

The methodological workflow is based on three main steps (Figure 2): (i) First a GWR model is estimated; (ii) A spatially constrained multivariate clustering is then applied to the GWR coefficient map in order to delineate data-driven spatial contexts; Finally (iii) a multilevel model including these contexts as a random effect is estimated, and then compared to a counterpart model incorporating administrative units as a grouping variable.

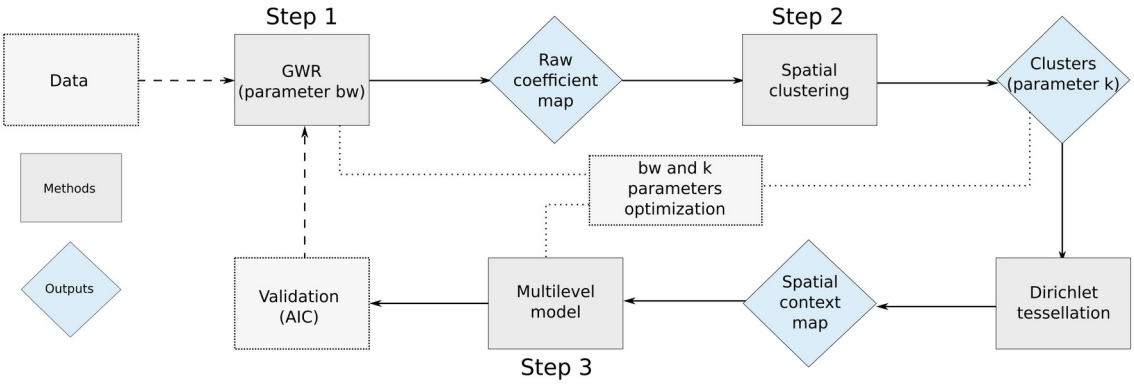


Figure 2. Overall methodological workflow followed in this study: bw is the bandwidth used in the GWR weighting scheme, k is the number of clusters, and AIC is the Akaike information criteria. The three steps marked in the figure are those that are specifically explained in the method section below.

**Dataset and variables**

In this study we use data from the ongoing French Nutrinet-Santé web-based cohort (Hercberg et al., 2010), restricting to participants residing in the Paris region. Launched in 2009, this cohort aims to provide information about relationships between nutrition and health among more than 100,000 participants aged 18 years or older, who completed through a secured website a set of questionnaires assessing their socioeconomic and health-related characteristics. Residential addresses were obtained from all participants, geocoded to the parcel or street levels. We focused on the area covering Paris inner city and its three surrounding “départements” called the “Petite Couronne” (Figure 3), 6.9 million inhabitants (2022 French census), i.e., a population density of ~8000 hab./km² (> 20,000 hab./km² in Paris inner city).

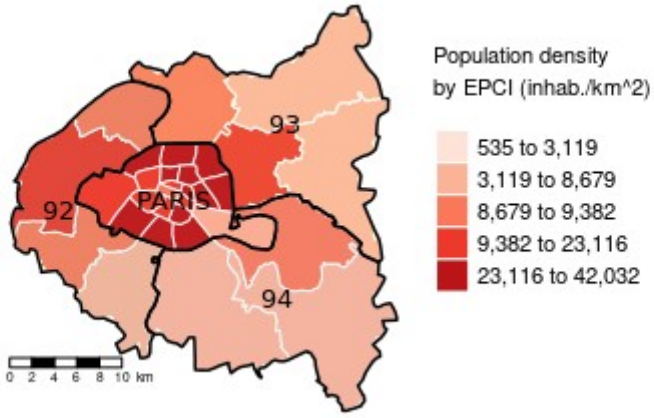


Figure 3. Location map of the study area. Black lines delineate “départements” (92 : Haut-de-Seine; 93 : Seine-Saint-Denis; 94 : Val-de-Marne) while white liness define contours of intercommunality administrative structure called *Etablissement public de coopération intercommunale* (EPCI), used as reference units in the subsequent multilevel models.

Data for 9,086 participants were available for analysis in the study area after the removal of missing residential addresses. The outcome variable is the body mass index (BMI), calculated as weight [kilograms]/height² [meters]. BMI is the indicator most commonly used in epidemiological and surveillance settings to assess body weight and excess adiposity defining obesity (WHO Europe Obesity Report, 2022). Although BMI is simply a function of weight and height it is strongly linked to ill health. Given that this article focuses on a methodological dimension, a minimal set of explanatory variables have been included, in order to emphasize computation and complexity (we will address the question of less parsimonious models in the discussion). Therefore, in the first step of the GWR, only two individual variables were considered: sex (categorical) and age (continuous and scaled beforehand) of participants. While socioeconomic profiles are known to be strongly associated with obesity, we did not consider them since they are typically spatially patterned. The linear functional form between age and BMI was visually checked using a spline-based generalized additive model.

**Step 1 - Geographically weighted regression**

We hypothesize that some predictors of a dependent variable have greater impact in some places than in others. GWR addresses this underlying assumption by fitting local regression models to each individual location (Brunsdon et al., 1996; Fotheringham et al., 2002). Each local model uses an inverse distance weighting scheme such as bisquare or Gaussian

kernel functions, assuming that closer observations from the calibration point are more influential than those further away. A basic GWR model extends traditional regression as follows:

$$y_i = \beta_{0(u_i,v_i)} + \sum_{m}^{k=1} \beta_k x_{ik_{(u_i,v_i)}} + \epsilon_{(u_i,v_i)} \quad \text{(Equation 1)}$$

Where $y_i$ is the dependent variable at location *i*, $(u_i, v_i)$ denotes the geographical coordinates of the *i*th location (i.e. individual), $x_{ik}$ is the *k*th independent variable at location *i*, $\beta_{0(u_i,v_i)}$ is the intercept at location *i*, $\beta_{k(u_i,v_i)}$ is the local regression coefficient for the *k*th independent variable at location *i*, and $\epsilon_{(u_i,v_i)}$ is a normally distributed error term at location *i*. The estimation of the local regression coefficients is given by the following equation (in a matrix form):

$$\beta_{(u_i,v_i)} = \left[\mathbf{X}^\top \mathbf{W}_i \mathbf{X}\right]^{-1} \mathbf{X}^\top \mathbf{W}_i \mathbf{y} \quad \text{(Equation 2)}$$

Where $\mathbf{X}$ is the sampling matrix of independent variables and $\mathbf{W}_i$ is the diagonal spatial weight matrix defining the weights given to each neighbour of an observation during the local calibration:

$$\mathrm{W}_i = \begin{bmatrix} w_{i,1} & a_{1,2} & \cdots & 0 \\ 0 & w_{i,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & w_{i,n} \end{bmatrix} \quad \text{(Equation 3)}$$

Where *n* is the number of neighbours at the *i*th location. Whether an observation is defined as a neighbour is given by two possible criteria: either a distance threshold, or a given number of *k* nearest neighbours (knn). The knn approach is advised when observations are irregular over space, to avoid too few observations in some local kernels (Feuillet et al., 2015). Subsequently, this method was used in this study. Thus, different kernel functions can be used to define neighbouring spatial weight matrices. We tested the five following kernel functions (Gollini et al., 2015):

$$\text{Gaussian:} \quad w_{ij} = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{knn}\right)^2\right)$$

$$\text{Exponential:} \quad w_{ij} = \exp\left(-\frac{|d_{ij}|}{knn}\right)$$

$$\text{Box-car:} \quad w_{ij} = \begin{cases} 1 & \text{if } |d_{ij}| < knn, \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Bi-square:} \quad w_{ij} = \begin{cases} \left(1 - (d_{ij}/knn)^2\right)^2 & \text{if } |d_{ij}| < knn, \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Tri-cube:} \quad w_{ij} = \begin{cases} \left(1 - (d_{ij}/knn)^3\right)^3 & \text{if } |d_{ij}| < knn, \\ 0 & \text{otherwise} \end{cases}$$

The bi-square function was found to be the best one to minimize the GWR AIC and was therefore kept in further analyses (note that the methods and results of the GWR-related parameter optimization – knn, but also the number of clusters – will be addressed in a dedicated final method section, because these are related to the overall workflow). Finally, each set of parameters was tested for spatial non-stationnarity using the $F_3$ test described in Leung et al. (2000) and implemented in the GWmodel R package (Gollini et al., 2015; Lu et al., 2014). The test statistics reflect the sample variance of the estimated values of $\beta_{ik}$ and can be approximated by a F-distribution. H0 is that all $\beta_{ik}$ are equal.

**Step 2 - Spatially constrained multivariate clustering of GWR estimates**

The next step consists in computing a spatially constrained multivariate clustering based on the vector of local GWR estimates. In addition to the coefficients associated to the two predictors (sex and age), local intercepts were also considered in the spatial clustering. These intercepts reflect the average BMI when sex and age are fixed and thus give the spatial distribution of the adjusted BMI. This age- and sex-adjusted BMI is itself potentially affected by contextual effects and therefore captures many unobserved spatially structured predictors, such as socioeconomic deprivation and the density of the built environment.

Several methods exist to build spatially constrained multivariate clustering. For example, in their study of housing market regionalization based on GWR coefficients, Helbich et al. (2013) used the SKATER (Spatial Klustering Analysis by Tree Edge Removal) algorithm, developed by Assunção et al. (2006). However, this method is resource-intensive and would make the whole parameter optimization burdensome. To overcome this issue, we used a more efficient method proposed by Chavent et al. (2018). This method consists in building a

hierarchical ascendant classification based on two matrices. The first is a dissimilarity matrix based on the vector of the three scaled GWR coefficients while the second is a neighbour matrix (contiguity) representing the spatial constraint. Once the spatial clustering was computed, the last step was to move from the punctual classified observations to polygons. For this purpose, we polygonised points through a Dirichlet tessellation, also known as Voronoi diagram. This regionalization procedure subdivides space in n cells (called Voronoi polygons), n being the number of points (i.e. home addresses), so that every location in a given cell is closer to its generating point than to any other. We then attributed cluster membership to each Voronoi polygon, and finally combined polygons according to clusters to form spatial contexts.

**Step 3 - Multilevel (or hierarchical) modelling**

The GWR-based regionalization resulting from the previous step was used to group observations in a subsequent multilevel model. Therefore we have observations *i = 1,...,n* clustered in groups *j = 1,...,J*, *j* representing each GWR-based spatial context. In multilevel models, we can allow groups to vary through the use of either intercepts (intercept-varying models), slopes (slope-varying models), or both (complete models). Here we estimated the following complete model, including two predictors at level-1 but no group-level predictors, expressed as follows:

$$\begin{aligned} y_{ij} &= \beta_{0j} + \beta_{1j}x_{ij1} + \beta_{2j}x_{ij2} + \epsilon_{ij} \\ \beta_{0j} &= \gamma_{00} + u_{0j} \\ \beta_{1j} &= \gamma_{10} + u_{1j} \\ \beta_{2j} &= \gamma_{20} + u_{2j} \end{aligned} \qquad \text{(Equation 4)}$$

Where $y_{ij}$ is the response variable for an individual i in a GWR-based spatial context *j*, $x_k$ are the predictors (sex and age), $\gamma_{00}$ is the mean intercept for all the contexts, $\gamma_{10}$ and $\gamma_{20}$ are the mean slopes, $\epsilon_{ij}$ is a normally distributed term, and $u_{kj}$ are the independent and normally distributed deviations between each context and the mean relationships.

**Parameter optimization and model validation**

Two essential parameters must be set during the model calibrations. The first is the GWR bandwidth (i.e., the local kernels based on knn) and the second is the number of clusters. GWR bandwidths are commonly selected using cross-validation based on prediction

accuracy. However, the purpose of this study is not to maximize the GWR prediction performance, but rather to maximize the final multilevel model derived from GWR estimate clustering. Consequently, we selected optimal parameters (bandwidth and number of clusters) in minimizing the Akaike information criteria (AIC) of the final multilevel model. Since the AIC function has potentially several local minima, it requires the implementation of a constrained optimization of a non-linear, multivariate and complex objective function. We used the generalized simulated annealing algorithm (Kirkpatrick et al., 1983), as being able to tackle these constraints and to converge more quickly than other similar methods (Xiang et al., 2017). Beyond this optimization, robustness analyses were performed on the number of clusters (from 20 to 60) and type of administrative units (EPCI and municipalities) and will be presented in the results section. Note that the administrative-based model was based on a French specific intercommunality administrative structure called *Etablissement public de coopération intercommunale* (EPCI, $k$ = 31). We also compared results using municipalities (known as communes) as a grouping variable ($j$ = 143).

Validation of the final GWR-based multilevel model was achieved by (i) comparing its AIC with the one of a usual administrative-based multilevel model and (ii) comparing the intraclass correlation coefficient (ICC, which is an indicator of the correlation between two observations taken randomly within a given spatial context or administrative unit) and slope variances (indicating the ability of the model to capture contextual effects) of the two multilevel models. The model exhibiting the highest values of ICC and slope variances was considered as the one including the most relevant spatial contexts among the contexts tested in this study.

All the analyses have been written in R language (R Core Team, 2023). The code core structure is shared on the following GitHub page: https://github.com/tfeuillet/gean2023, but note that given our analyses involved personal health data, the code cannot be be directly reproducible.

**Results**

The mean age of the participants was 42.4 years and 76.8% were women. This overrepresentation of women may reflect a volunteer bias, knowing that it is established that women are more likely to participate in research studies (Andreeva 2016, Galea, 2007) and are also more health- and nutrition-conscious than men (Barebring, 2020). Mean BMI was 23.4 kg/m2 (descriptive statistics are presented in Table 1).

Table 1. Characteristics of the sample (N = 9089)

| | N | Mean (SD) or % | Range |
|---|---|---|---|
| Age | 9089 | 42.4 (14.6) | [18-86] |
| Gender (female) | 9089 | 76.8 % | n/a |
| BMI [kg/m$^2$] | 9089 | 23.4 (4.5) | [12.2-76.2] |

**GWR model outputs**

GWR scaled coefficient surfaces are presented in Figure 4 and their descriptive statistics in Table 2. Coefficients are interpolated (inverse distance weighting method) for visualization purposes, and the spatial distribution of participants is mapped using kernel densities (Fig. 3A). The optimization algorithm yielded an optimal GWR bandwidth equaling 67 nearest neighbours of participants. The mean distance between each participant and its 67$^{th}$ nearest neighbour is about 1 km. While the OLS model reveals that BMI is on average positively and significantly correlated with age and with being a female, GWR shows significant spatial non-stationarity for the intercepts and age coefficients, with regular opposite signs according to location (Figure 4 B,C,D). There is a strong difference between the median GWR coefficient for age, and the OLS coefficient (0.98 vs. 0.07, respectively, see Table 2), indicating that many local coefficients are higher than the mean OLS coefficient. The intercept map also displays contrasted values throughout the study site.

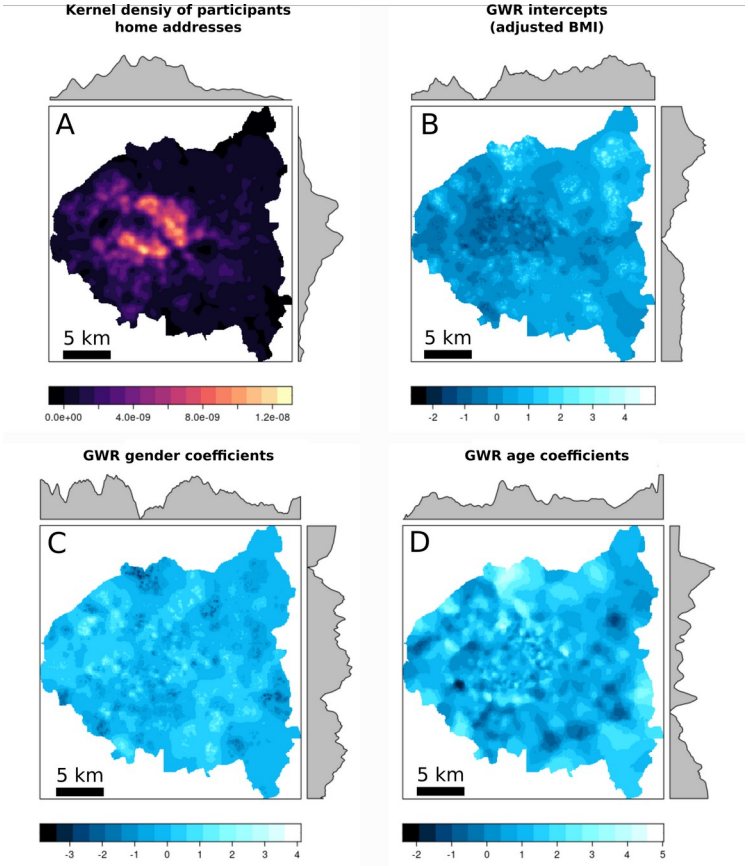


Figure 4. A: Spatial distribution of the participants (N = 9089) using kernel density. B, C and D: Maps of the scaled GWR coefficients after IDW interpolation (for intercepts, gender and age, respectively). Marginal plots (in grey) summarize the mean raster values (kernel density) using zonal statistics.

Table 2. GWR raw coefficient summary. OLS coefficients have been added for comparison purposes.

| | GWR model (non-scaled coefficients) | | | | | | OLS model | |
|---|---|---|---|---|---|---|---|---|
| | Min | 1st Qu | Median | 3rd Qu | Max | Leung's $F_3$-test for spatial non-stationnarity (p-value) | Beta | p-value |
| Intercept | 20.21 | 22.33 | 23.00 | 23.77 | 28.85 | <0.001 | 20.26 | <0.001 |
| Gender (ref = male) | -4.88 | 0.23 | 1.23 | 2.24 | 7.63 | >0.05 | 1.19 | <0.001 |
| Age | -2.09 | 0.55 | 0.98 | 1.44 | 4.77 | <0.001 | 0.07 | <0.001 |
| Number of observations | 9089 | | | | | | 9089 | |
| $R^2$ | 0.23 | | | | | | 0.07 | |

Regarding the GWR-based regionalization, the results provide 60 spatial contexts, mapped in Figure 5, alongside the 31 administrative units used in the comparison model.

**Multilevel model estimation**

Results of the two multilevel models (administrative-based and GWR-based, respectively, the municipality level being discussed later) are provided in Table 3 and Figure 5. The first finding is that the AIC of the administrative-based model is higher than the GWR-based model (52326 vs. 52156), indicating a better quality of the hybrid model. The second finding is that the intraclass correlation coefficient (ICC) is more than twice as high in the GWR-based model as that for the other (5.9% vs. 2.8%), reflecting higher correlations among observations in the GWR spatial contexts than in the administrative units. Finally, the last interesting observation is the difference in slope variances (noted $\tau_{11}$) between the two models, equaling 0.032 in the first and 0.212 in the GWR-based model. This difference is illustrated in Figure 5 for the partial effect of age. This means that heterogeneity of relationships is stronger in GWR contexts than in administrative ones.

| | **Administrative-based multilevel model** | | | **GWR-based multilevel model** | | |
|---|---|---|---|---|---|---|
| *Predictors* | *Estimates* | *95% CI* | *p* | *Estimates* | *95% CI* | *p* |
| Intercept | 23.04 | 22.76 – 23.32 | **<0.001** | 23.48 | 23.21 – 23.75 | **<0.001** |
| Gender (ref = male) | 1.23 | 1.02 – 1.44 | **<0.001** | 1.25 | 1.04 – 1.46 | **<0.001** |
| Age | 0.99 | 0.88 – 1.11 | **<0.001** | 1.05 | 0.90 – 1.21 | **<0.001** |
| **Random effects** | | | | | | |
| $\sigma^2$ | 18.344 | | | 17.804 | | |
| $\tau_{00}$ | $0.504_{epci}$ | | | $0.912_{gwr}$ | | |
| $\tau_{11}$ | $0.032_{epci.age}$ | | | $0.212_{gwr.age}$ | | |
| $\rho_{01}$ | $0.387_{epci}$ | | | $0.395_{gwr}$ | | |
| ICC | 0.028 | | | 0.059 | | |
| N | $31_{epci}$ | | | $60_{gwr}$ | | |
| Observations | 9089 | | | 9089 | | |
| Deviance | 52303.8 | | | 52134.2 | | |
| AIC | 523226.4 | | | 52156.4 | | |
| log-Likelihood | -26156.2 | | | -26071.2 | | |

Table 3. Results of the multilevel models. Left: model using administrative units (N = 31) as a grouping variable (i.e. random effect). Right: model using GWR-based spatial contexts as a grouping variable (N = 60). $\tau_{00}$ refers to the intercept variance, and $\tau_{11}$ to the slope variance (for age).

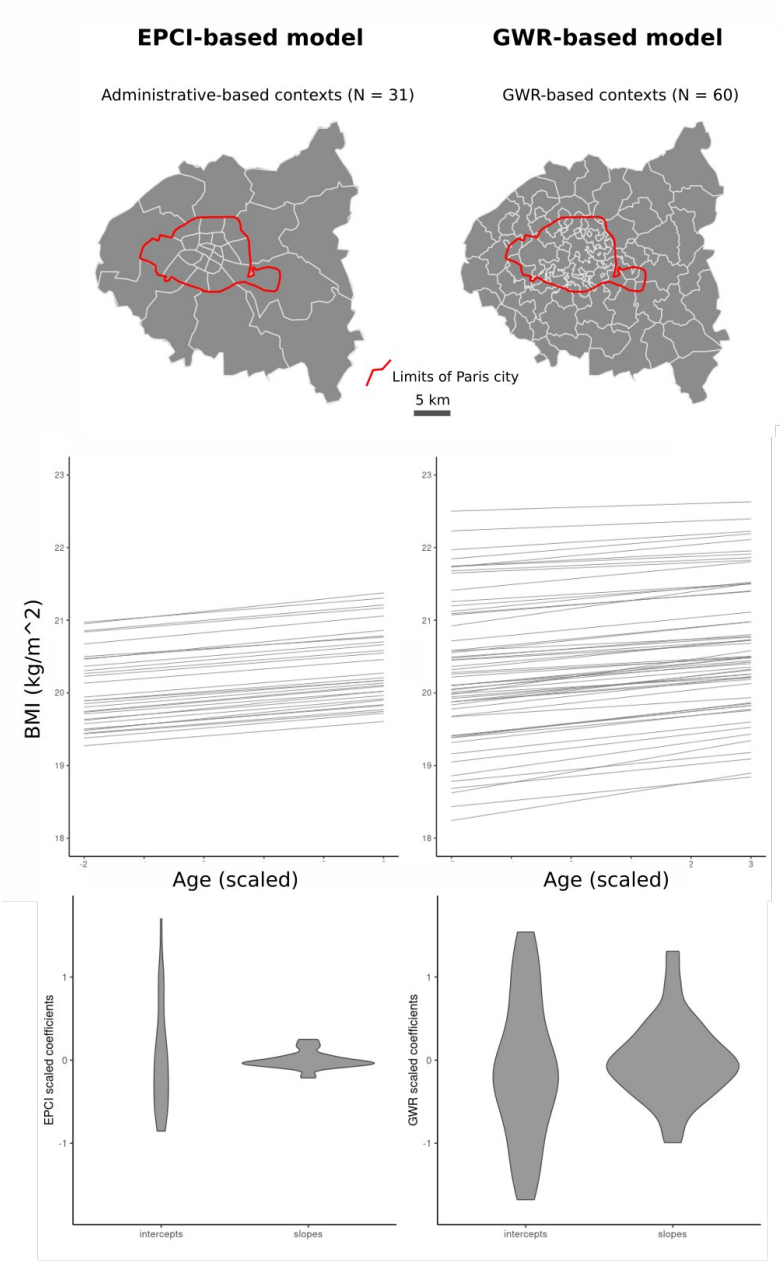


Figure 5. Maps of spatial contexts (EPCI and GWR-based) used in the multilevel models. Graphs display the relationships between age and BMI in the two models. The violin plots (bottom) show that both intercept and slope variances are higher in the GWR-based model.

**Model robustness under alternative parameters**

In multilevel modelling, random effect variances depend on the number of units in the grouping variable. To check for such a possible bias, we (i) conducted robustness analyses by varying the number of GWR clusters from 20 to 60 and (ii) estimated a multilevel model based on municipalities ($j$ = 143) instead of intercommunalities (EPCI) as used previously. Results showed that AIC of GWR-based models remained lower than administrative-based

models notwithstanding the number of clusters (Figure 6). Note that the AIC plateaued from 45 clusters. Likewise, the model based on municipalities exhibited a higher AIC than the one based on intercommunality units, as well as lower ICC and slope variance (data not shown).

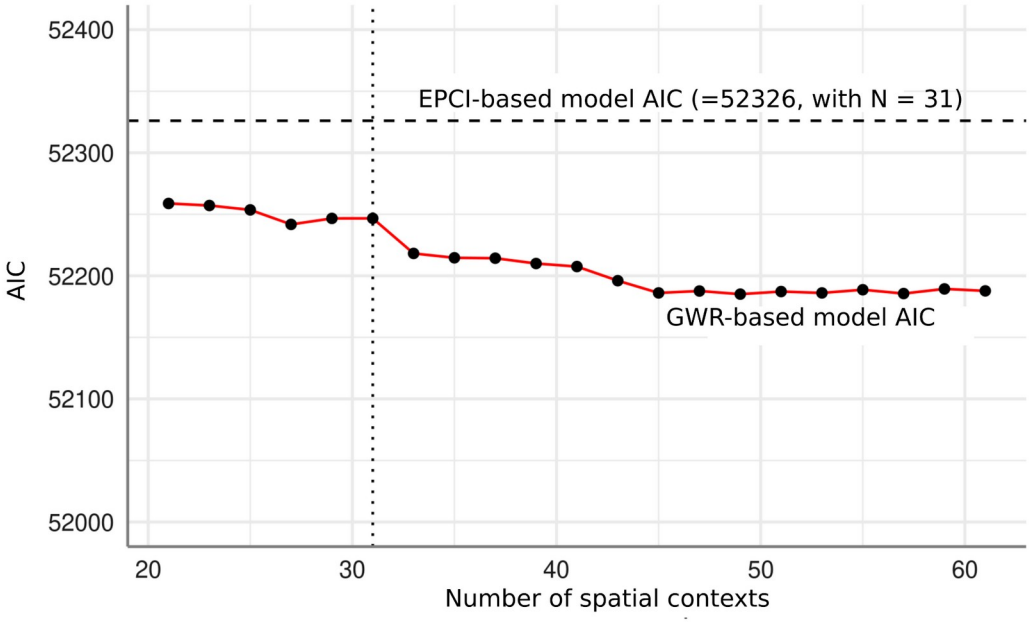


Figure 6. Robustness analyses on the number of GWR-based spatial contexts. The hybrid model outperformed the EPCI-based model whatever the number of contexts in terms of AIC minimization.

**Discussion**

In this study we focused on developing a novel modelling approach combining two major and complementary modelling frameworks commonly used in contextual studies (GWR and multilevel models). We applied this approach to the geography of human corpulence as assessed by the BMI. When integrating GWR-based spatial contexts as contextual settings rather than administrative units (commonly used in multilevel models), we overcame the well-known spatial design issue in multilevel modelling (Owen et al., 2016). We demonstrated in our sample that our hybrid approach outperforms administrative-based multilevel models as measured through decreasing AIC and account better for contextual effects through ICC and slope variance increases.

The spatial design issue in multilevel modelling refers to the mismatch between the geographical boundaries of contextual setting used in the model and the true causally relevant spatial contexts affecting the studied outcome, that remain, in fact, unknown (Diez-Roux & Mair, 2010; Owen et al., 2016). In many studies, particularly neighbourhood studies,

administrative boundaries are used by default and this choice affects subsequent inferential results about contextual effect occurrence and strength. Therefore, delineating relevant spatial contexts in multilevel models is a crucial challenge. Conceptual thinking about sociospatial context boundaries has led to interesting debates in human geography (Feuillet et al., 2016; Kwan, 2012, 2018; Petrović et al., 2020). Recommendations are often made to consider idiosyncratic contexts, since the way people respond to environmental influences (i.e., context effects) are highly personal (Kwan, 2018). Likewise, activity spaces are individual, and some contextual studies have used self-defined spatial contexts (Charreire et al., 2016; Perchoux et al., 2013). However, such an individual view of contextual effects is not without strong methodological issues, both in terms of data acquisition in large samples, and in terms of compatibility with the multilevel modelling framework.

Therefore, the solution we proposed to define spatial contexts, based on a spatial non-stationarity and a data-driven strategy, may be considered as a workable compromise between administrative units (partially disconnected from actual contextual effects), and idiosyncratic contexts (impossible to integrate into multilevel models as a random effect). GWR-based spatial contexts seem relevant in that they can capture unobserved spatially structured factors that affect the outcome under study (here, measures of obesity such as BMI). Such multidimensional and complex factors that shape spatial contexts are typically difficult to measure quantitatively at a supra-local scale (i.e., at a scale ensuring sufficient heterogeneity), as they relate to complex and spatially heterogeneous interactions between places and people (Feuillet et al., 2016; Fotheringham et al., 2021; Fotheringham & Sachdeva, 2022). According to Fotheringham & Sachdeva (2022), spatial context “is a shorthand term for the impact of the largely unmeasurable effects of location on one’s actions [...] and hence a multifaceted concept incorporating the influence of local media, family, friends, and local organizations as well as notions of traditions, persistent adverse or beneficial conditions, customs, lifestyles and psychological profiles common to an area that affect social norms, which in turn affect individual behavior” (Fotheringham & Sachdeva, 2022, p. 3). In ecological studies of obesity, the complexity of obesogenic contexts leading to a high BMI (so-called obesogenic environments) including personal and socioeconomic factors, and the built environment, into a comprehensive social ecological system has been emphasized (Swinburn et al., 1999) and can be transposed to other fields of human-related behaviour. We argue here that local modelling techniques such as GWR can be used to efficiently reveal spatial contexts otherwise unmeasurable, and consequently be included as a grouping variable in subsequent multilevel models to relevantly capture contextual effect in a sound inferential framework. That said, it is worth noting that spatial and contextual effects are sometimes embedded, and that “unmeasurable effects of location” can actually manifest

as both spatial heterogeneity and spatial dependence. This complexity leads to be cautious with the interpretations, in particular with salient effects that could require further qualitative work to reach a sound conclusion.

Caution is also crucial because of the other possible causes of GWR-derived spatial non-stationarity, leading to potential inferential bias. The first other cause raised by Fotheringham & Sachdeva (2022) is noise due to sampling variation during the subset local calibrations, inherent to kernel-based local techniques such as GWR. Such a sampling variation may result in the spatial variability of parameters, even if relationships are constant over space. The second cause is due to the assumption of an incorrect functional form in the relationships under study. Some studies have demonstrated that nonlinear relationships modelled as linear in spatially varying coefficient models may exhibit spatial non-stationarity (Sachdeva et al., 2022). Note that a nonlinear relationship can also reflect some spatial non-stationarity. This is typically the case when a predictor is strongly spatially patterned, e.g., in following a centre-periphery gradient, such as density or socioeconomic deprivation in European cities (Feuillet et al., 2021). In this study we checked the linear form of the relation between age and BMI through using generalized additive models, a highly recommended approach (Hastie, 2017). However, the functional form of a relationship can be linear on average, but non-linear in specific local kernels. Taken together, these points (noise and model misspecification) constitute a pitfall that modelers must be aware of.

Another point of discussion arises from the choice to limit the number of predictors in the GWR model, as we did in this study. We deliberately selected a limited number of explanatory variables which were only individual level variables We justify this choice by the fact that including contextual variables (e.g., characteristics of the built environment expected to influence obesity-related behaviours) in GWR would require defining individual exposure to such characteristics (typically buffers around home addresses, or administrative boundaries). This would bring us back to the initial problem raised in this study, namely, how to delimit the individual exposure to a spatial context. Thus, we recommend adding contextual variables at the step of the multilevel model, i.e., after having built the GWR spatial contexts. Moreover, this solution allows us to be parsimonious in GWR and to make the regionalization procedure easier, while also avoiding the redundancy of spatial information in GWR. Indeed, including spatially structured variables in GWR makes interpretation somewhat complicated since it layers over the spatially explicit information derived from local kernels.

We tested this hybrid approach by applying it to geographical variation in BMI, the most

common indicator of obesity, but it would be worth considering generalizing it to other outcomes and other fields of geographical analyses. Neighbourhood effects are well suited to such a method, since they embrace complex spatial contexts that are not easily measurable by common means. Besides multilevel models, GWR-based regionalization has already been used for housing market segmentation and further hedonic modelling (Helbich et al., 2013), and we may suppose that such an approach could also be useful for some other zoning-based analyses, such as spatial interaction models. Also, potential extensions or complements could be tested, for instance in using the multiscale GWR proposed by Fotheringham et al. (2017), which is able to capture the different spatial scales at which predictors operate. The way the regionalization is done is also a parameter that could be compared and improved. For example, the spatially clustered regression recently proposed by Sugasawa & Murakami (2021) could be an interesting way to delineate spatial contexts as well. These issues represent open avenues as areas for research, to improve and consolidate this new hybrid approach of contextual effect modelling.

## Conclusion

The issue of the delimitation of a pertinent spatial context is crucial in multilevel modelling, in order to appropriately capture contextual effects. In the social sciences, administrative boundaries are often used by default and because of a relative ease of access, even if they do not match with the actual geography of the spatial processes at work. In this study, focusing on geographical distribution of the BMI as an indicator of obesity, we suggest regionalizing GWR coefficients in order to delineate unmeasurable spatial contexts assumed to affect BMI and the relationships between BMI and individual variables (age and gender). Finally, we use these contexts as a random effect in a subsequent random slope multilevel model. We show that the GWR-based multilevel model outperformed its administrative-based counterpart, in terms of quality and contextual effect modelling. From a theoretical perspective, this hybrid procedure also provides a means to reconcile space and place, in accounting for both the spatial effect through GWR, and the place effect through multilevel modelling. When characterizing geospatial data, these two kinds of effects are rarely considered simultaneously in statistical models despite existing theoretical soundness. It can be hoped that the methodology presented here will pave the way to further extensions and improvements, in particular in parameter calibration and robustness to other data samples and outcomes, at various scales.

## Acknowledgments

The authors warmly thank Pierre Barbillon and the two anonymous reviewers for their suggestions on the first version of the manuscript.

## References

Arcaya, M., Brewster, M., Zigler, C. M., & Subramanian, S. V. (2012). Area variations in health: A spatial multilevel modeling approach. *Health & Place*, *18*(4), 824–831. https://doi.org/10.1016/j.healthplace.2012.03.010

Assunção, R. M., Neves, M. C., Câmara, G., & Freitas, C. D. C. (2006). Efficient regionalization techniques for socio-economic geographical units using minimum spanning trees. *International Journal of Geographical Information Science*, *20*(7), 797–811. https://doi.org/10.1080/13658810600665111

Bärebring, L., Palmqvist, M., Winkvist, A. et al. Gender differences in perceived food healthiness and food avoidance in a Swedish population-based survey: a cross sectional study. Nutr J 19, 140. https://doi.org/10.1186/s12937-020-00659-0

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis*, *28*(4), 281–298. https://doi.org/10.1111/j.1538-4632.1996.tb00936.x

Chaix, B., Merlo, J., Subramanian, S. V., Lynch, J., & Chauvin, P. (2005). Comparison of a spatial perspective with the multilevel analytical approach in neighborhood studies: The case of mental and behavioral disorders due to psychoactive substance use in Malmö, Sweden, 2001. *American Journal of Epidemiology*, *162*(2), 171–182. https://doi.org/10.1093/aje/kwi175

Charreire, H., Feuillet, T., Roda, C., Mackenbach, J. D., Compernolle, S., Glonti, K., Bárdos, H., Le Vaillant, M., Rutter, H., McKee, M., De Bourdeaudhuij, I., Brug, J., Lakerveld, J., & Oppert, J.-M. (2016). Self-defined residential neighbourhoods: Size variations and correlates across five European urban regions: Self-defined residential neighbourhoods. *Obesity Reviews*, *17*, 9–18. https://doi.org/10.1111/obr.12380

Chavent, M., Kuentz-Simonet, V., Labenne, A., & Saracco, J. (2018). ClustGeo: An R package for hierarchical clustering with spatial constraints. *Computational Statistics*, *33*(4), 1799–1822. https://doi.org/10.1007/s00180-018-0791-1

Chen, D.-R., & Truong, K. (2012). Using multilevel modeling and geographically weighted regression to identify spatial variations in the relationship between place-level disadvantages and obesity in Taiwan. *Applied Geography*, *32*(2), 737–745. https://doi.org/10.1016/j.apgeog.2011.07.018

Diez-Roux, A. V., & Mair, C. (2010). Neighborhoods and health. *Annals of the New York Academy of Sciences*, *1186*, 125–145. https://doi.org/10.1111/j.1749-6632.2009.05333.x

Dong, G., & Harris, R. (2015). Spatial Autoregressive Models for geographically hierarchical data structures. *Geographical Analysis*, *47*(2), 173–191. https://doi.org/10.1111/gean.12049

Dong, G., Ma, J., Harris, R., & Pryce, G. (2016). Spatial random slope multilevel modeling using multivariate conditional autoregressive models: A case study of subjective travel satisfaction in Beijing. *Annals of the American Association of Geographers*, *106*(1), 19–35. https://doi.org/10.1080/00045608.2015.1094388

Duncan, C., Jones, K., & Moon, G. (1998). Context, composition and heterogeneity: Using multilevel models in health research. *Social Science & Medicine*, *46*(1), 97–117. https://doi.org/10.1016/S0277-9536(97)00148-2

Feuillet, T., Bulteau, J., & Dantan, S. (2021). Modelling context-specific relationships

between neighbourhood socioeconomic disadvantage and private car use. *Journal of Transport Geography*, *93*, 103060. https://doi.org/10.1016/j.jtrangeo.2021.103060
Feuillet, T., Charreire, H., Menai, M., Salze, P., Simon, C., Dugas, J., Hercberg, S., Andreeva, V. A., Enaux, C., Weber, C., & Oppert, J.-M. (2015). Spatial heterogeneity of the relationships between environmental characteristics and active commuting: Towards a locally varying social ecological model. *International Journal of Health Geographics*, *14*(1), Article 1. https://doi.org/10.1186/s12942-015-0002-z
Feuillet, T., Salze, P., Charreire, H., Menai, M., Enaux, C., Perchoux, C., Hess, F., Kesse-Guyot, E., Hercberg, S., Simon, C., Weber, C., & Oppert, J.-M. (2016). Built environment in local relation with walking: Why here and not there? *Journal of Transport & Health*, *3*(4), Article 4. https://doi.org/10.1016/j.jth.2015.12.004
Feuillet, T., Valette, J. F., Charreire, H., Kesse-Guyot, E., Julia, C., Vernez-Moudon, A., Hercberg, S., Touvier, M., & Oppert, J. M. (2020). Influence of the urban context on the relationship between neighbourhood deprivation and obesity. *Social Science & Medicine*, *265*, 113537. https://doi.org/10.1016/j.socscimed.2020.113537
Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically weighted regression: The analysis of spatially varying relationships*. Wiley–Blackwell.
Fotheringham, A. S., Li, Z., & Wolf, L. J. (2021). Scale, context, and heterogeneity: A spatial analytical perspective on the 2016 US presidential election. *Annals of the American Association of Geographers*, *111*(6), 1602–1621.
Fotheringham, A. S., & Sachdeva, M. (2022). On the importance of thinking locally for statistics and society. *Spatial Statistics*, *50*, 100601. https://doi.org/10.1016/j.spasta.2022.100601
Fotheringham, A. S., Yang, W., & Kang, W. (2017). Multiscale Geographically Weighted Regression (MGWR). *Annals of the American Association of Geographers*, *107*(6), 1247–1265. https://doi.org/10.1080/24694452.2017.1352480
Fotheringham, A. S., & Wong, D. W. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and planning A*, *23*(7), 1025-1044.
Galea S, Tracy M. Participation rates in epidemiologic studies. *Ann Epidemiol., 17*(9):643-53.
Goldstein, H. (2011). *Multilevel statistical models*. John Wiley & Sons.
Gollini, I., Lu, B., Charlton, M., Brunsdon, C., & Harris, P. (2015). GWmodel: An R package for exploring spatial heterogeneity using geographically weighted models. *Journal of Statistical Software*, *63*(1), 1–50. https://doi.org/10.18637/jss.v063.i17
Goodchild, M. F. (2004). GIScience, Geography, Form, and Process. *Annals of the Association of American Geographers*, *94*(4), 709–714.
Hastie, T. J. (2017). Generalized additive models. In *Statistical models in S* (pp. 249-307). Routledge.
Helbich, M., Brunauer, W., Hagenauer, J., & Leitner, M. (2013). Data-driven regionalization of housing markets. *Annals of the Association of American Geographers*, *103*(4), 871–889. https://doi.org/10.1080/00045608.2012.707587
Hercberg, S., Castetbon, K., Czernichow, S., Malon, A., Mejean, C., Kesse, E., Touvier, M., & Galan, P. (2010). The Nutrinet-Santé Study: A web-based prospective study on the relationship between nutrition and health and determinants of dietary patterns and nutritional status. *BMC Public Health*, *10*(1), 242. https://doi.org/10.1186/1471-2458-10-242
Hu, Y., Lu, B., Ge, Y., & Dong, G. (2022). Uncovering spatial heterogeneity in real estate prices via combined hierarchical linear model and geographically weighted regression. *Environment and Planning B: Urban Analytics and City Science*, *49*(6), 1715–1740. https://doi.org/10.1177/23998083211063885
Janko, M., Goel, V., & Emch, M. (2019). Extending multilevel spatial models to include spatially varying coefficients. *Health & Place*, *60*, 102235. https://doi.org/10.1016/j.healthplace.2019.102235
Jones, K., Johnston, R., & Pattie, C. J. (1992). People, places and regions: Exploring the use of multi-level modelling in the analysis of electoral data. *British Journal of*

*Political Science*, *22*(3), 343–380.
Kearns, R. A., & Joseph, A. E. (1993). Space in its place: Developing the link in medical geography. *Social Science & Medicine*, *37*(6), 711–717. https://doi.org/10.1016/0277-9536(93)90364-A
Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, *220*(4598), 671–680. https://doi.org/10.1126/science.220.4598.671
Kwan, M.-P. (2012). The uncertain geographic context problem. *Annals of the Association of American Geographers*, *102*(5), 958–968. https://doi.org/10.1080/00045608.2012.687349
Kwan, M.-P. (2018). The limits of the neighborhood effect: Contextual uncertainties in geographic, environmental health, and social science research. *Annals of the American Association of Geographers*, *108*(6), 1482–1490. https://doi.org/10.1080/24694452.2018.1453777
Leung, Y., Mei, C.-L., & Zhang, W.-X. (2000). Statistical tests for spatial nonstationarity based on the geographically weighted regression model. *Environment and Planning A: Economy and Space*, *32*(1), 9–32. https://doi.org/10.1068/a3162
Lu, B., Harris, P., Charlton, M., & Brunsdon, C. (2014). The GWmodel R package: Further topics for exploring spatial heterogeneity using geographically weighted models. *Geo-Spatial Information Science*, *17*(2), 85–101. https://doi.org/10.1080/10095020.2014.917453
Merlo, J., Chaix, B., Yang, M., Lynch, J., & Råstam, L. (2005). A brief conceptual tutorial of multilevel analysis in social epidemiology: Linking the statistical concept of clustering to the idea of contextual phenomenon. *Journal of Epidemiology & Community Health*, *59*(6), 443–449. https://doi.org/10.1136/jech.2004.023473
Openshaw, S. (1978). An empirical study of some zone-design criteria. *Environment and planning A*, *10*(7), 781-794.
Oshan, T. M., Smith, J. P., & Fotheringham, A. S. (2020). Targeting the spatial context of obesity determinants via multiscale geographically weighted regression. *International Journal of Health Geographics*, *19*. https://doi.org/10.1186/s12942-020-00204-6
Owen, G., Harris, R., & Jones, K. (2016). Under examination: Multilevel models, geography and health research. *Progress in Human Geography*, *40*(3), 394–412. https://doi.org/10.1177/0309132515580814
Park, Y. M., & Kim, Y. (2014). A spatially filtered multilevel model to account for spatial dependency: Application to self-rated health status in South Korea. *International Journal of Health Geographics*, *13*(1), 6. https://doi.org/10.1186/1476-072X-13-6
Perchoux, C., Chaix, B., Cummins, S., & Kestens, Y. (2013). Conceptualization and measurement of environmental exposure in epidemiology: Accounting for activity space related to daily mobility. *Health & Place*, *21*, 86–93. https://doi.org/10.1016/j.healthplace.2013.01.005
Petrović, A., Manley, D., & van Ham, M. (2020). Freedom from the tyranny of neighbourhood: Rethinking sociospatial context effects. *Progress in Human Geography*, *44*(6), 1103–1123. https://doi.org/10.1177/0309132519868767
R Core Team (2023). R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
Roche, S. (2016). Geographic information science II: Less space, more places in smart cities. *Progress in Human Geography*, *40*(4), 565–573. https://doi.org/10.1177/0309132515586296
Sachdeva, M., Fotheringham, A. S., Li, Z., & Yu, H. (2022). Are we modelling spatially varying processes or non-linear relationships? *Geographical Analysis*, *54*(4), 715–738. https://doi.org/10.1111/gean.12297
Sugasawa, S., & Murakami, D. (2021). Spatially clustered regression. *Spatial Statistics*, *44*, 100525. https://doi.org/10.1016/j.spasta.2021.100525
Swinburn, B., Egger, G., & Raza, F. (1999). Dissecting obesogenic environments: The development and application of a framework for identifying and prioritizing environmental interventions for obesity. *Preventive Medicine*, *29*(6), 563–570.

https://doi.org/10.1006/pmed.1999.0585
Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, *46*(sup1), 234–240. https://doi.org/10.2307/143141
Wolf, L. J., Anselin, L., Arribas-Bel, D., & Mobley, L. R. (2021). On spatial and platial dependence: Examining shrinkage in spatially dependent multilevel models. *Annals of the American Association of Geographers*, *111*(6), 1679-1691.
Xiang, Y., Gubian, S., Martin, F., Xiang, Y., Gubian, S., & Martin, F. (2017). Generalized simulated annealing. In *Computational Optimization in Engineering—Paradigms and Applications*. IntechOpen. https://doi.org/10.5772/66071