An agent-based model to investigate the effects of social segregation around the clock on social disparities in dietary behaviour

AUTHORS
Clémentine COTTINEAU, Julien PERRET, Romain REUILLON,
Sébastien REY-COYREHOURCQ, Julie VALLÉE

RÉSUMÉ
Les comportements de santé des individus dépendent étroitement de leur statut socio-économique et des interactions sociales qu'ils peuvent avoir avec leur entourage. Ces interactions résultent de la ségrégation à l'œuvre dans l'ensemble des quartiers qu'ils fréquentent au quotidien. Afin d'évaluer l'impact de la ségrégation sociale tout au long de la journée sur les comportements de santé, nous proposons un modèle multi-agents des dynamiques de comportements alimentaires en Ile-de-France basé sur des données empiriques. Le modèle comprend un mécanisme de mise en interaction, de diffusion d'opinion et de changement de comportement sous contrainte. À partir d'une population synthétique de 8 millions d'agents, nous combinons deux scénarios de localisation résidentielle des agents (aléatoire et observée) et deux scénarios de localisation des lieux quotidiens d'activité (aléatoire et observée) afin de comparer l'impact des représentations spatio-temporelles des territoires (et des interactions sociales associées) sur nos modèles de simulation de la dynamique des comportements alimentaires.

MOTS CLÉS
Santé, alimentation, dynamique, ségrégation socio-spatiale, modèle multi-agents, mobilité quotidienne

ABSTRACT
Evidence suggests that socioeconomic profile and social interactions with neighbours affect health behaviours. Social interactions result from the way individuals are spatially distributed in the city (i.e. the level of segregation) not only at their residences but also at their daily activities/locations. In order to evaluate the impact of social segregation "around the clock" on health behaviours, we present a data-driven agent-based model of dietary dynamics in the Paris Region. The model includes a mechanism which puts agents in interaction, a mechanism of opinion diffusion and a mechanism of behaviour change under constraint. From a synthetic representative population of 8 million agents, we combine two scenarios of agents' residential location (random location and census-based location) and two scenarios of agents' daily activities location (random location and survey-based location). With these four scenarios, we aim to compare the impact of the spatiotemporal representation of territories (and the related social interactions) on the outputs of a simulation model of dietary dynamics.

KEYWORDS
Health, Diet, Dynamics, Socio-spatial Segregation, ABM, Daily mobility

INTRODUCTION
Since the pioneering work of social epidemiologists and health geographers, it has been regularly shown and confirmed that health and dietary behaviours are unequally distributed across social groups. High income earners and highly educated individuals live longer, healthier, they

eat better and exercise more often than their less educated and less wealthy counterparts. Marmot (2015) even demonstrated that each increment in the social ladder corresponds to a gain in life expectancy. Economic inequality is thus deemed to be a driver of poorer health and health inequality (Wilkinson & Pickett, 2009), partly because of the spatial concentration of similar individuals resulting from the socio-spatial segregation of individuals (a direct consequence of inequality). Spatial concentration of people with similar resources, norms and behaviour gives way to neighbourhood effects which can be seen as a retroaction of the local society's features on individual behaviours. Local homogeneity of behaviour, in health as well as in other domains such as education or income, happens through various channels, among which social interactions (contagion, socialisation, competition), environmental processes (physical constrains, spatial mismatch), or institutions (stigmas, market actors) (Galster, 2010). In the context of dietary behaviour for example, several studies in the US and the UK have found evidence of unequal access to healthy foods (fresh fruits, vegetables, etc.) between racially and economically segregated neighbourhoods. Similarly, the sharing of social norms or the construction of social capital can affect the dietary habits of whole communities. This means that the way individuals are spatially distributed in the city (i.e. the level of segregation) can affect social gradients in exposure, opinion and finally behaviour. It is important, in that matter, to recall that neighbourhood effects on health behaviours are not only residential-based since individuals are not static and tied to their neighbourhood of residence (Cummins, 2007; Vallée & Chauvin, 2012). When exploring dynamics in health behaviours, it makes sense to consider the daily trajectories of people and the full spectrum of segregation experienced by people "around the clock" (Le Roux *et al.*, 2017). Besides, it also makes sense to take into account the daily trajectories of places and to overcome "jetlag" suffered by place attributes when they are labelled with frozen attributes over a 24-hour period (Vallée, 2017). Daily trajectories of both people and places are then relevant to model inequalities in health and dietary behaviours across social groups.

Because of the spatio-temporal features of dietary dynamics, both the theoretical framework of time-geography and the methodological framework of spatial agent-based models (ABM) seem particularly suited. Auchincloss *et al.* (2011) provided a first example of ABM of income inequalities in diet in the context of residential segregation. They focused on the interaction between income differences and the differences in the supply of food through the uneven location of cheap and expensive shops and through extreme cases of segregation patterns (all agents of the left part of the artificial world being rich, the other half being poor, etc.).

In this contribution, we aim to take advantage of the computational possibilities of high performance computing (HPC) to model dynamics of social inequalities in diet in more details in the Paris Region, taking into account daily trajectories of both people and places. More precisely, we aim to explore the different opportunities for opinion diffusion for people depending whether they are located at their residence, place of work and leisure activities, and depending how social context of the places they are located in may change over a 24-hour period. We use nutrition survey data and census cross-category tables to generate a synthetic population of 8.16 million agents aged 16+ in the Paris Region. This population is then fed as input to a model of diet dynamics which involves spatial interactions, opinion changes and behaviour shifts. We finally assess the sensitivity of the model's results to spatial-temporal initial conditions by running four scenarios. The four scenarios result from crossing two options of agents' residential location (random location or census-based location to reproduce actual residential segregation patterns in the Paris region) and two options of agents' daily activities location

(random location or survey-based location to reproduce activity-based segregation patterns in the Paris region). By comparing these four scenarios, we want to question and to compare four spatio-temporal representations of the territory on which interactions between individuals occur.

1. GENERATING A SPATIALISED SYNTHETIC POPULATION

The synthetic population is calibrated on the population of the Paris Region (Île-de-France). Around 8.16 million synthetic individuals are artificially created. They are described by socio-demographic, daily mobility and dietary attributes.

1.1. Socio-demographic attributes

Socio-demographic attributes include sex (male/female), age groups (15-29; 30-59; 60+) and level of education (low; middle; high). The share of each category matches the census data and the distribution of joint categories (for example: a senior female of middle education level) matches joint distributions of the Paris Region.

1.2. Daily mobility attributes

Because opinion dynamics and dietary behaviour exhibit high stability over long period, the time we model is abstract. It represents the aggregation of dynamics over longer periods of time at similar hours during a summary "day". A "day" is thus composed of three time slices: night (00:00-08:00), day (08:00-16:00) and evening (16:00-24:00). We model the spatial distribution of agents at each time slice rather than the actual mobility between locations (which other models –such as MATSIM (Horni *et al.*, 2016)– do, being focused the modelling of individual mobilities and modal choices).

We first assign a fixed home location for each agent for the "night" slice (00:00-08:00). Based on the detailed 2012 census, direct sampling is used to compute the number of inhabitants in each of the socio-demographic categories in every census block ("IRIS") since the detailed survey is not available at this scale for confidential reasons. Agents are then distributed from their census block of residence into cells with 1 km side taking into account residential densities. From this process, the 8.16 million agents are located in 8,540 inhabited cells at "night" slice.

The allocation of an activity location for the "day" and "evening" time slices is computed by crossing the socio-demographic categories and the 2010 origin-destination survey for the Paris region, EGT (Enquête Globale Transport – DRIEA, STIF, OMNIL). Based on the 101,000 weekday trips between 127,000 locations performed by the 25,000 respondents of the survey aged 16+ (Le Roux et al, 2017), we select the list of locations of EGT respondents occurring within each of the two time-slices (day and evening). This partial list is then spatially interpolated (using the neighboring cells) in order to have a denser distribution of potential locations. We then compute the probability of each combination of socio-demographic attributes and residential location to be located in them. This probability was used to allocate a 'day-time' and 'evening-time' location to each synthetic agent.

The locations of the agents during the "night" and "day" time slices are defined at initialization and will be the same throughout the simulation. Among the 8.16 million agents, 67% are located in a "day" cell which differs from the "night" cell (5.52 million agents). On the contrary, the locations of the agents during the 'evening' time slice can change at every model iteration with a new selection of the location among the list of potential locations in the movement distribution.

1.3. Dietary attributes

The value of dietary attributes is probabilistically determined by the proportion of each socio-demographic groups in the Health and Nutrition Barometers (Baromètre Santé Nutrition) for the years 1996, 2002 and 2008. These surveys look at the attitudes, opinions and behaviours towards food of about 7,000 French individuals aged 15+ (1,409 in 1996, 2,097 in 2002 and 3,305 in 2008). We get the information about having or having not eaten 5 fruits and vegetable during the past day, based on the description of menus and reconstruction by the survey analysts. This characterises the behaviour we want to model, and in particular the increase of the proportion of people reaching the 5-a-day target since the eponym campaign launched in 2007. We use fruits and vegetable consumption in 1996 to initialize agent behaviour according to their socioeconomic category. We will use fruits and vegetable consumption in 2002 and 2008 to respectively parameter and validate our model.

To represent the respondents' opinion towards a healthy diet, we used the question "In your opinion, how many fruits and vegetables should one eat to be healthy?", and kept the ratio between the number answered and the recommended number 5.

Finally, we computed probabilities of interacting during the different time-slices using the answers to questions like "who do you have breakfast/lunch/dinner with?" (social environment of meals) and measured the proportion of individuals under dietary constraints by adding together people who answered "agree" and "strongly agree" to questions like "would you say that time/money/family habits is a constraint in the composition of your menus?".

This synthetic population with spatial locations, initial dietary behaviour, opinions and constrains is the population of agents on which the model of dynamics in dietary behaviour is run.

2. A SPATIO-TEMPORAL MODEL OF DIETARY DYNAMICS

We model dietary habits as a dynamic process of opinion change, which is then expressed – under constraints– into behaviour. The model is made of the repetition of a sequence of three mechanisms (fig. 1).

– Firstly, agents scan their environment and observe the current split of behaviour. Some agents can also interact with another individual in their cell and record his/her dietary opinion and behaviour. This personal interaction depends on the social environment of meals taken from the survey.

– Secondly, agents update their opinion about a healthy diet, taking into account their previous level of opinion (a continuous variable), the dominant behaviour and their potential partner's opinion and behaviour.

– Finally, every agent considers the opportunity to modify their behaviour according to its correspondence with the agent's current opinion and his/her constraints. For example, time increases the probability of healthy agents to switch to an unhealthy (quicker) behaviour if their opinion is unfavourable to a healthy diet, but it will decrease the probability of unhealthy agents to switch to a healthy behaviour if their opinion is favourable to a healthy diet, because this diet is assumed to be more time-consuming. Budget constraints work like time constraints, because a healthy diet is assumed to be more expensive than an unhealthy one. Family habits constraints on the other hand give more inertia to each behaviour and decrease the probability of switching in each direction.

A step ends with everyone moving to the location (if the location of the next activity is different

from the previous one). There are three activities per day and each simulation runs for at least five "days". This minimum number does not represent a realistic time-frame for diet dynamics but allows for some dynamics to occur in the model during abstract "days".

Figure 1. Description of interaction, opinion and behaviour dynamics
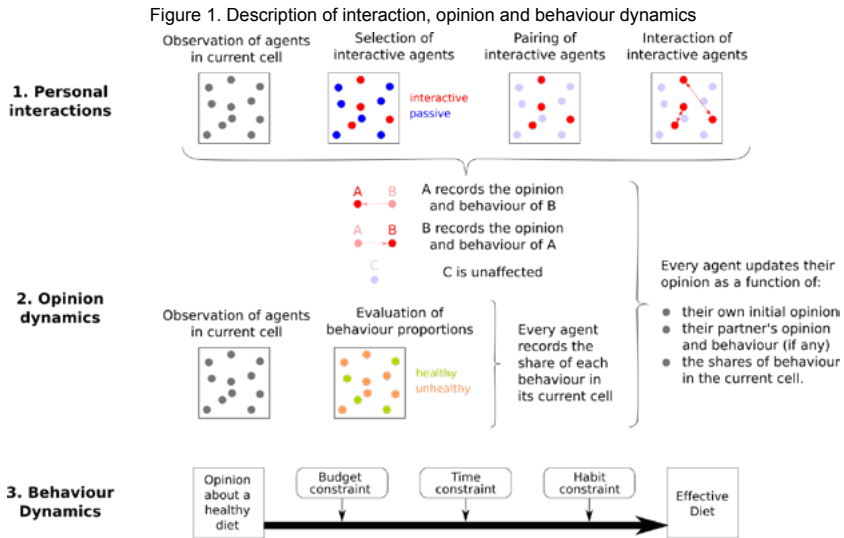
3. TESTING SENSITIVITY
TO SPATIO-TEMPORAL REPRESENTATION OF TERRITORIES

The model has three types of parameters: the ones which are initialised using the empirical values taken from Health and Nutrition Barometer, the ones which determine the spatial location of agents at different times of the day, regulating the spatiotemporal scenarios and some free parameters. The survey-calibrated parameters are not explored here, although they could be changed to model other national contexts. The free parameters are subject to a traditional sensitivity analysis, to see for example if the opinion inertia we add to the opinion updating mechanisms has an effect on the resulting segregation, because it controls the influence of others (which are encountered by co-presence) on one's opinion. Similarly, the parameter which controls the strength of constraints should affect social inequalities in diet, because different socio-demographic groups have different sets and levels of constraints.

The parameters defining spatial location of agents at different times of the day are two-fold. First, we control the sensitivity of the model to residential segregation patterns (ie. at "night") by having a random residential location scenario (scenario 1) and a census-based residential location scenario (scenario 2). Second, we control the sensitivity of the model to daily mobility patterns by adding 'day' and 'evening' locations either randomly (scenario 3) or according socio-demographic attributes and residential location of agents (scenario 4). Crossing these two scenarios gives four non-equivalent levels of granularity to represent dietary dynamics in the Paris region. The sensitivity analysis aims to answer two sets of questions:
– Can we observe quicker changes in dietary opinions and behaviours (i) when "night" locations are spatially segregated rather than randomly located and (ii) when "day" and "evening"

locations are considered?

– Can we observe increasing social inequalities in dietary behaviours increase across the time (i) when "night" locations are spatially segregated rather than randomly located and (ii) when "day" and "evening" locations are considered?

In scenarios 1 and 2, the most common in the literature about neighborhood effects, agents have only one location which corresponds to their usual place of residence. By contrast, in scenarios 3 and 4, agents move around the city several times per day.

CONCLUSION

We expect the precision given to the spatiotemporal representation of territories to have a non-trivial effect in the simulation of dietary behavior and related social inequalities. In other words, we expect the results obtained with the four scenarios to differ, and we expect the results of scenario 4 to be the most accurate to simulate dynamics in social health inequalities in the Paris region. By including elements of socio-spatial segregation (because empirical studies point to neighbourhood effects) and daily mobility (because the spaces of interactions are not restricted to the residential location) in agent-based model, we aim to underline how dynamics in health behaviors would gain to be modeled from a precise representation of city rhythm.

REFERENCES

Auchincloss A. H., Riolo R. L., Brown D. G., Cook J., Roux A. V. D., 2011, "An agent-based model of income inequalities in diet in the context of residential segregation", *American journal of preventive medicine*, 40(3), p. 303-311.

Cummins S., 2007, "Commentary: investigating neighbourhood effects on health. Avoiding the 'local trap'", *International Journal of Epidemiology*, 36(2), p. 355-357.

Galster G., 2012, "The mechanism(s) of neighborhood effects: Theory, evidence, and policy implications", *in* M. Van Ham, D. Manley, N. Bailey, L. Simpson and D. Maclennan (eds.), *Neighbourhood Effects Research: New Perspectives*, Dordrecht, NL, Springer, p. 23-56.

Horni A., Nagel K., Axhausen K. W., 2016, *The multi-agent transport simulation MATSim*, London, Ubiquity Press.

Le Roux G., Vallée J., Commenges H., 2017, "Social segregation around the clock in the Paris region (France)", *Journal of Transport Geography*, n° 59, p. 134-145.

Marmot M., 2015, *The health gap: the challenge of an unequal world*, London, Bloomsbury Publishing.

Vallée J., 2017, "The daycourse of place", *Social Science & Medicine*, n° 194, p. 177-181.

Vallée J., Chauvin P., 2012, "Investigating the effects of medical density on health-seeking behaviours using a multiscale approach to residential and activity spaces. Results from a prospective cohort study in the Paris metropolitan area, France", *International Journal of Health Geographics*, vol. 11, art. n° 54.

Wilkinson R.G., Pickett K., 2009, *The spirit level: why equality is better for everyone*, London, Penguin Books.

THE AUTHORS

Clémentine Cottineau
University College London – CASA
c.cottineau@ucl.ac.uk

Julien Perret
Université Paris-Est
IGN ENSG – LaSTIG COGIT
julien.perret@ign.fr

Romain Reuillon
ISC PIF – Géographie-cités
romain.reuillon@iscpif.fr

Sébastien Rey-Coyrehourcq
IDEES
sebastien.rey-coyrehourcq@univ-rouen.fr

Julie Vallée
CNRS – Géographie-cités
julie.vallee@parisgeo.cnrs.fr