



HAL
open science

Modélisation fine des dynamiques épidémiques - le projet ICI

Maxime Colomb, Carl Graham, Julien Perret, Denis Talay, Josselin Garnier

► **To cite this version:**

Maxime Colomb, Carl Graham, Julien Perret, Denis Talay, Josselin Garnier. Modélisation fine des dynamiques épidémiques - le projet ICI. 32e journée de la recherche de l'UGE-IGN-ENSG, Institut national de l'information géographique et forestière, Mar 2023, Champs sur Marne, France. hal-04654401

HAL Id: hal-04654401

<https://hal.science/hal-04654401v1>

Submitted on 27 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Projet ICI — Présentation Journées de la recherche

Maxime Colomb ^{1,2,3} Carl Graham ^{1,2} Josselin Garnier ¹ Julien Perret ^{3,4} Denis Talay ^{1, 2}

30 Mars 2023

¹ Centre de Mathématiques Appliquées de l'Ecole polytechnique (CMAP - École Polytechnique)

²Analyse d'interactions stochastiques intelligentes et coopératives (ASCI-INRIA)

³Laboratoire des Sciences et Technologies de l'Information Géographique (LASTIG-IGN)

⁴Laboratoire de démographie et d'histoire sociale (LDH-EHESS)

Introduction - Objectifs

ICI est un simulateur de propagation d'épidémies reposant sur une **modélisation très détaillée** d'un espace urbain et de sa population.

Modèles SIR/SEIR incapables de simuler l'effet de **politiques sanitaires**.

Conception d'un simulateur à l'échelle individuelle permettant de :

- Représenter des **chaînes de contaminations** réalistes.
- Simuler l'impact de **politiques sanitaires** en implémentant des mesures sur un environnement urbain **très détaillé**.

Composé de plusieurs modules :

I) Le générateur de données

1. Un recensement d'**activités** exhaustif.
 - L'ouverture et la fermeture des activités selon leurs types.
 - Une densité d'individus dans chaque activité.
2. Une **population** dans des ménages structurés.
3. L'**occupation** des individus dans ces différentes activités.

Composé de plusieurs modules :

I) Le générateur de données

1. Un recensement d'**activités** exhaustif.
 - L'ouverture et la fermeture des activités selon leurs types.
 - Une densité d'individus dans chaque activité.
2. Une **population** dans des ménages structurés.
3. L'**occupation** des individus dans ces différentes activités.

II) Le simulateur de diffusion épidémique

peu abordé dans cette partie

Modélisation urbaine

Recensement, fusion et enrichissement des activités

Modélisation de l'intérieur des bâtiments

Le **bâtiment** est le centre de notre modélisation.

Trois types d'activités contenues :

Lieu de travail

Activité professionnelle, classe, étude sup, etc.

Points d'intérêt (POI)

Commerces, loisirs, démarches, etc.

Logement

Vide ou occupé. Héberge un ménage.

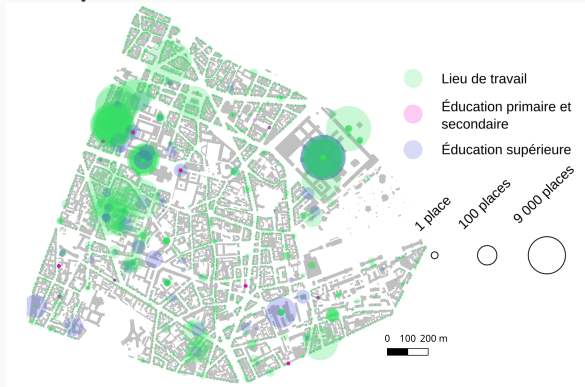
Lieux de travail

Recensement des activités économiques dans la base de données **SIRENE**.

Effectifs et nombre de classes des écoles du MEN.

Effectifs des établissements supérieurs du MESR.

Exemple sur le Vème arrondissement de Paris



POI : Fusion de données de références

Composition des POI

légende : *nombre d'entrées dans la nomenclature* *position géographique* *Indice de taille*

Commerces recensés par la
municipalité Parisienne



218 types d'établissement
Classe de surface de
l'établissement
Centre commercial
Entrée des établissement
contrôlée manuellement

Activités déclarées



368 types d'établissement
Tranche d'effectif



732 types d'établissement
Tranche d'effectif

Géocodage

Plateforme collaborative
de données géographiques



176 catégories (sur le V^{ème})
Capacité (rare) Horaire
Entrée des établissement
contrôlée manuellement

Base de données des
équipements



158 types d'établissement
Pas de mesure de taille
Géocodage par l'INSEE

Création d'une **nomenclature** spéciale pour le **projet** comportant
34 classes, adaptés aux décrets de fermeture des confinements.

Fusion automatisée et enrichissement de ces données.

Population Synthétique

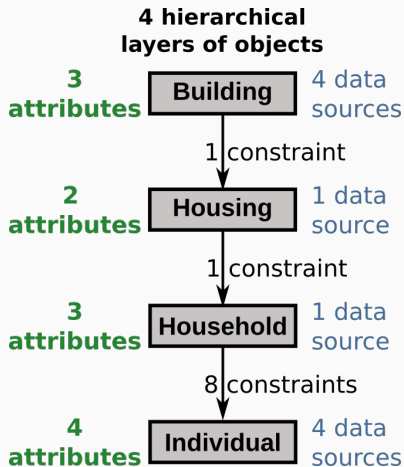
Génération de ménages structurés

Population synthétique quadri-couche

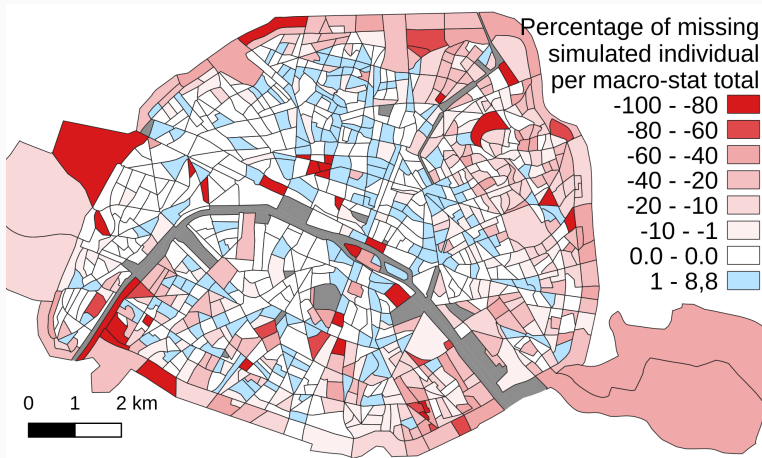
Un générateur original

- Fortement **géolocalisé**.
- Configurations **hiérarchisés**, plausibles et respectant des contraintes macro ou globales,
- **Flexible** aux données disponibles et la génération d'attributs.

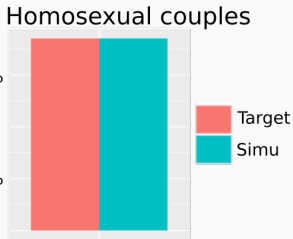
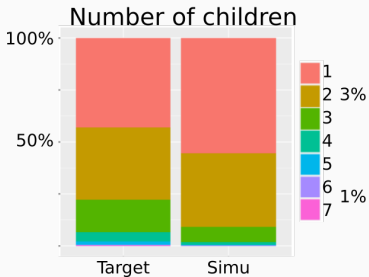
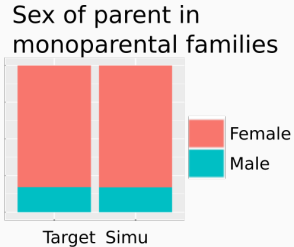
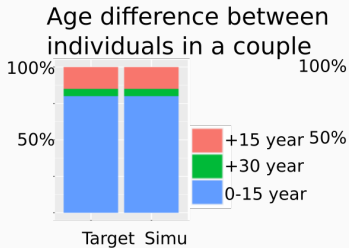
Méthode proposée de **rejet stochastique sous contrainte**.



Résultats de population synthétique



Résultats de population synthétique



Présence de population dans les activités

Qui fait quoi, où et quand

Flux de population

Présence plutôt que déplacements.

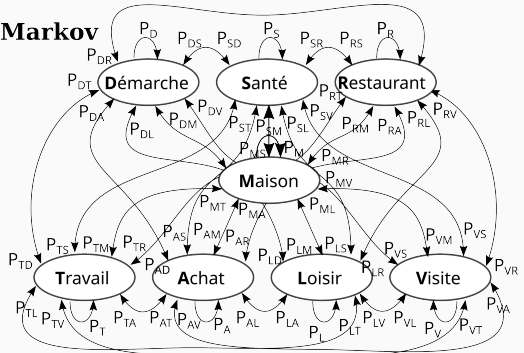
Fusion de données hétérogènes :

- Profils et raisons de présences des individus grâce à l'**EGT**.
- Localisation fine des présences grâce à des traces de **données mobiles**.
- Flux pendulaires et occasionnels grâce aux fréquentations de la RATP.

Attribution **stochastique** des activités individuelles.

Génération d'un agenda

Chaine de Markov

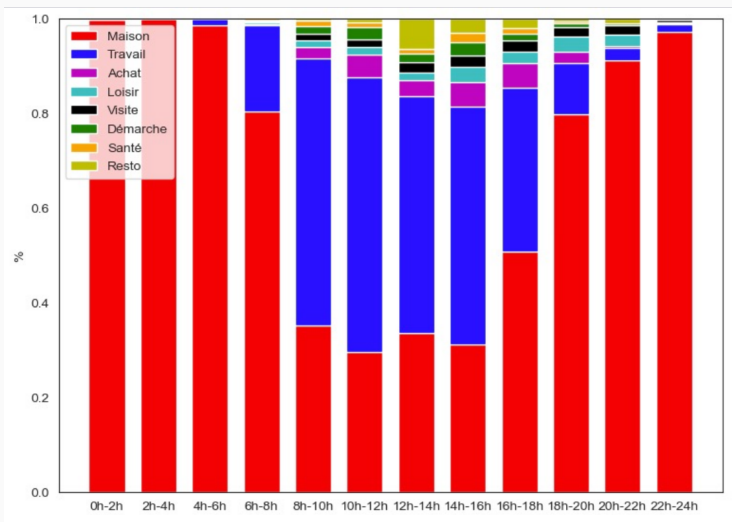


Déclinaisons du paramétrage

Tranche x Âge x Occupation x Social Status x Loc x Type					
horaire	0-2	Élève	Agriculteur	Paris	jour
custom	3-5				
3h	6-10	Étudiant	Libéral	1 ^{ère} couronne	Semaine
2h	11-17				
1h	18-24	Actif	Intellectuel	2 nd couronne	Samedi
30min	25-39				
	40-54	Mi-temps	Intermédiaire		Dimanche
	55-64				
	65-79	Chômeur	Ouvrier		
	80+				
		Retraité	Retraité		
			Autre		

*Paramétrage pour chaque catégorie de population
calculé à partir de l'enquête globale des transports (EGT)*

Génération d'un agenda



Exemple des proportions du temps passé dans chaque activité pour l'ensemble des catégories de population.

Conclusions et perspectives

Un modèle modulaire de génération de données

Un modèle de données générique et interopérable

Trois modules *relativement* indépendants

Flexibilité quant au besoin de données d'entrée.

Diffusion des données dans le web sémantique

Génération d'un graph suivant le formalisme RDF.

SPARQL endpoint is coming.

Calcul de la diffusion épidémique

Base pour le modèle de diffusion épidémique

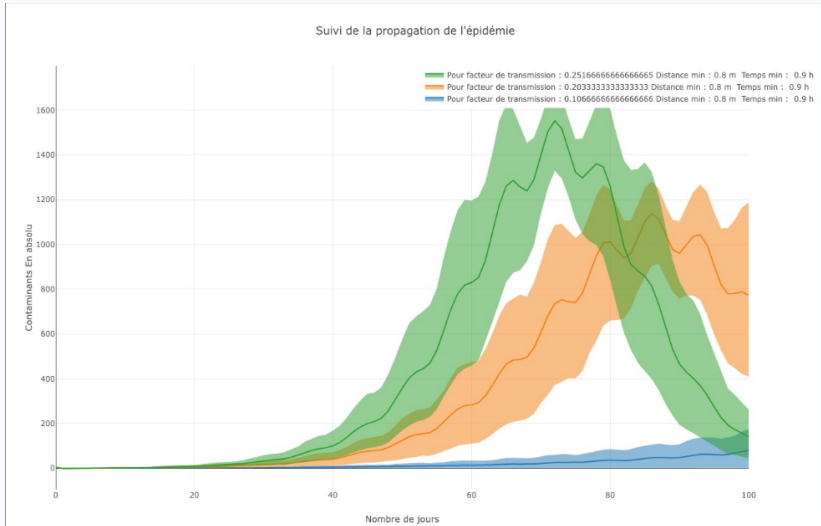
Simulation très variables et répliqués un grand nombre de fois.

Implémentation optimisée pour simuler rapidement de larges zones.

Possibilité d'implémentation de scénarios sanitaires

- Restrictions de déplacements (couvre-feu, confinement).
- Fermeture de magasins, d'écoles.
- Confinements de ménages infectés.
- Couvre-feu
- Campagne de vaccination

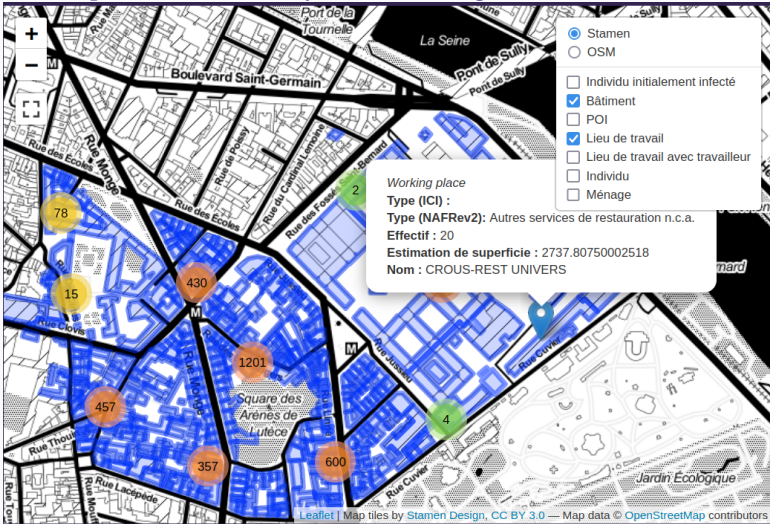
Résultats



Visualisation des données

Disponibles sur tout Paris ici :

https://ici.saclay.inria.fr/generic_arr.html/



- Intégration de **comportement différencié** selon les règles sanitaires (projet SafeCityMap (INSA Lyon)).
- **Visualisation** des résultats avec les indicateurs de LASTIG-Geoviz
- Collaboration avec l'INSERM sur les données et les applications épidémiologiques.
- Insertion de certains modules dans un catalogue de **jumeaux numériques** nationaux et souverains

Merci de votre attention

<https://ici.saclay.inria.fr>