

Discontinuités spatiales, santé et mobilités

Analyses et typologies des Google POI et des Tweets pour caractériser les structures spatiales et les dynamiques d'attractivités de Bangkok (Thaïlande)

Cebeillac A.¹, Daudé E.¹², Vaguet A¹

¹ Université de Rouen, UMR IDEES, Rouen, France

² CNRS UMR IDEES

DENFREE
AEDESS



Contexte épidémique de la Dengue

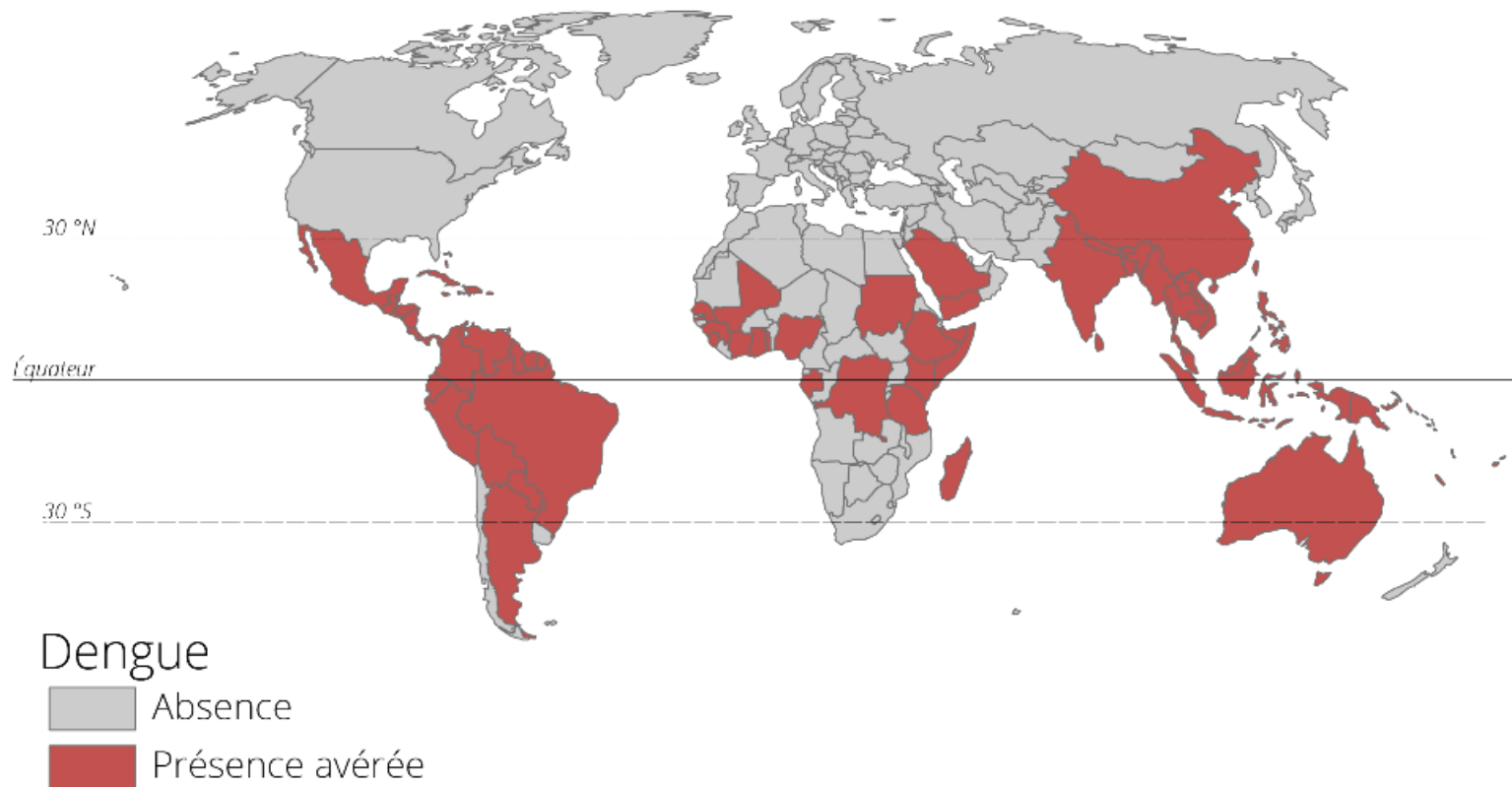


Figure 1 : La dengue dans le monde (OMS 2016)

Transmission de la dengue

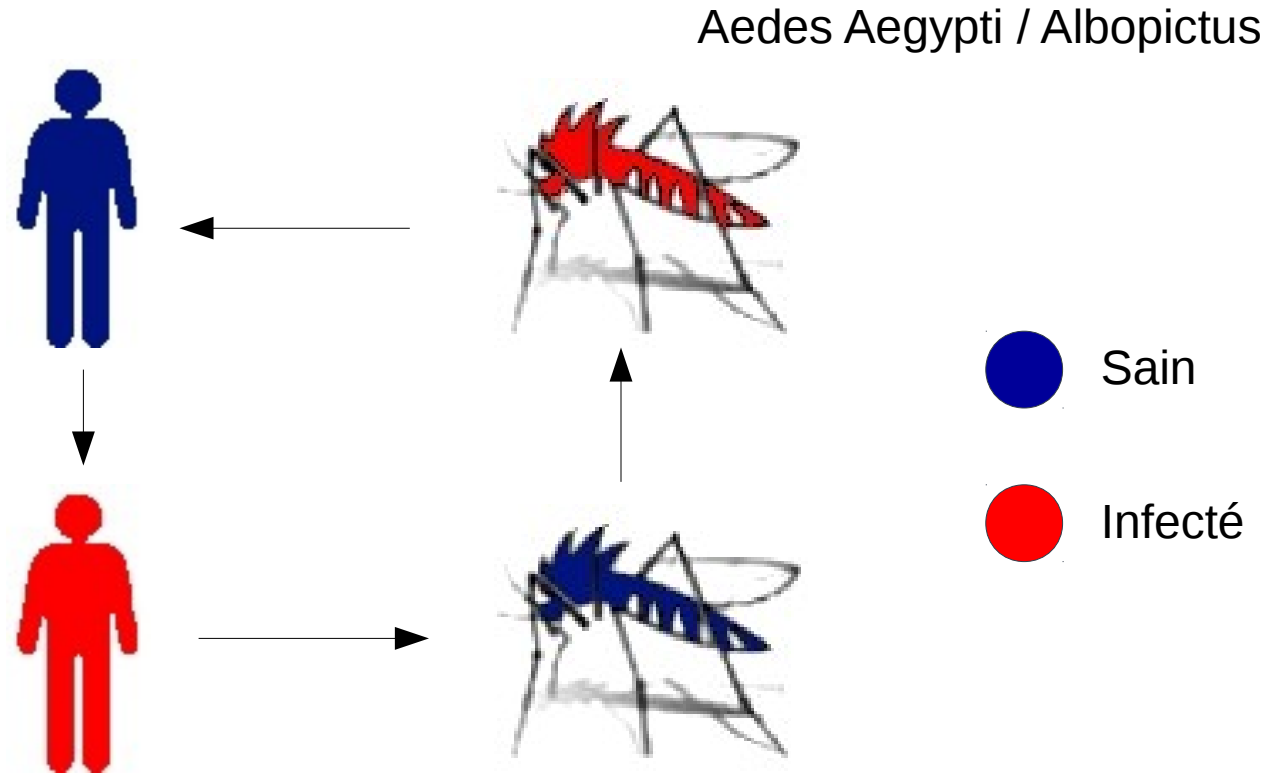


Figure 2 : Transmission très schématisée de la dengue

Cas de Bangkok

Bangkok 9,2M d'habitants en 2015 // Superficie ~ 1500 km²

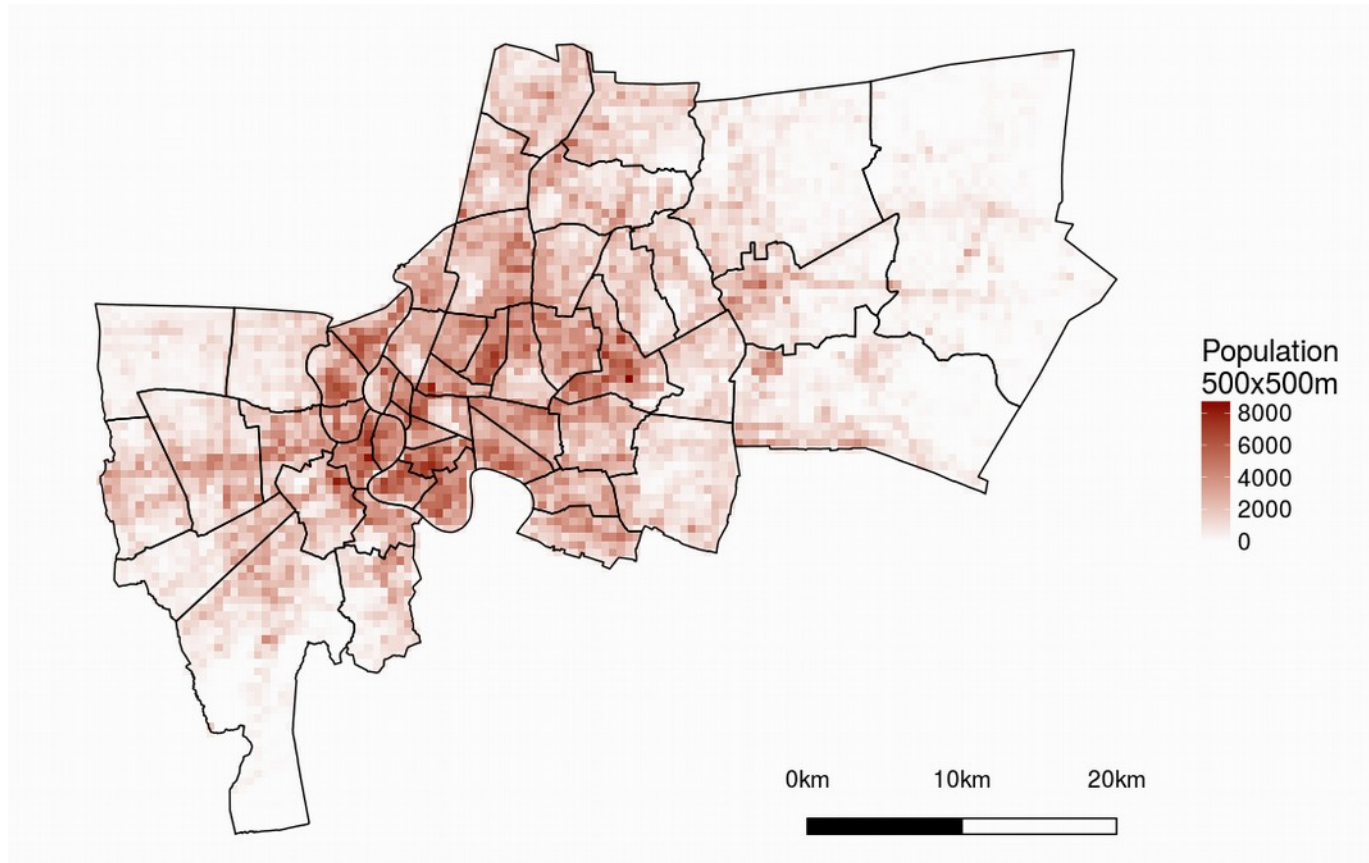


Figure 3 : Répartition de la population à Bangkok - Misslin & Daudé, 2016

Dengue à Bangkok

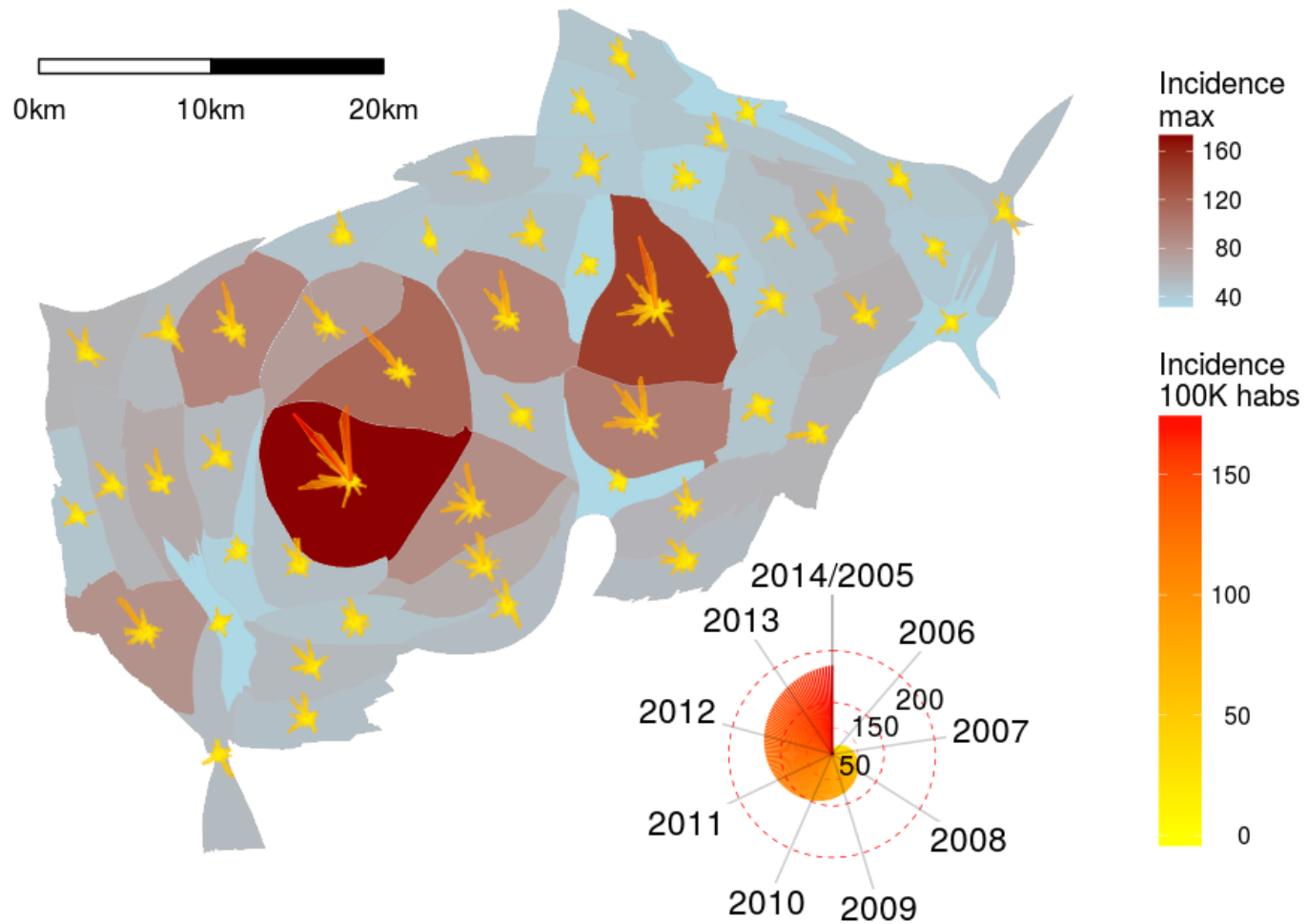
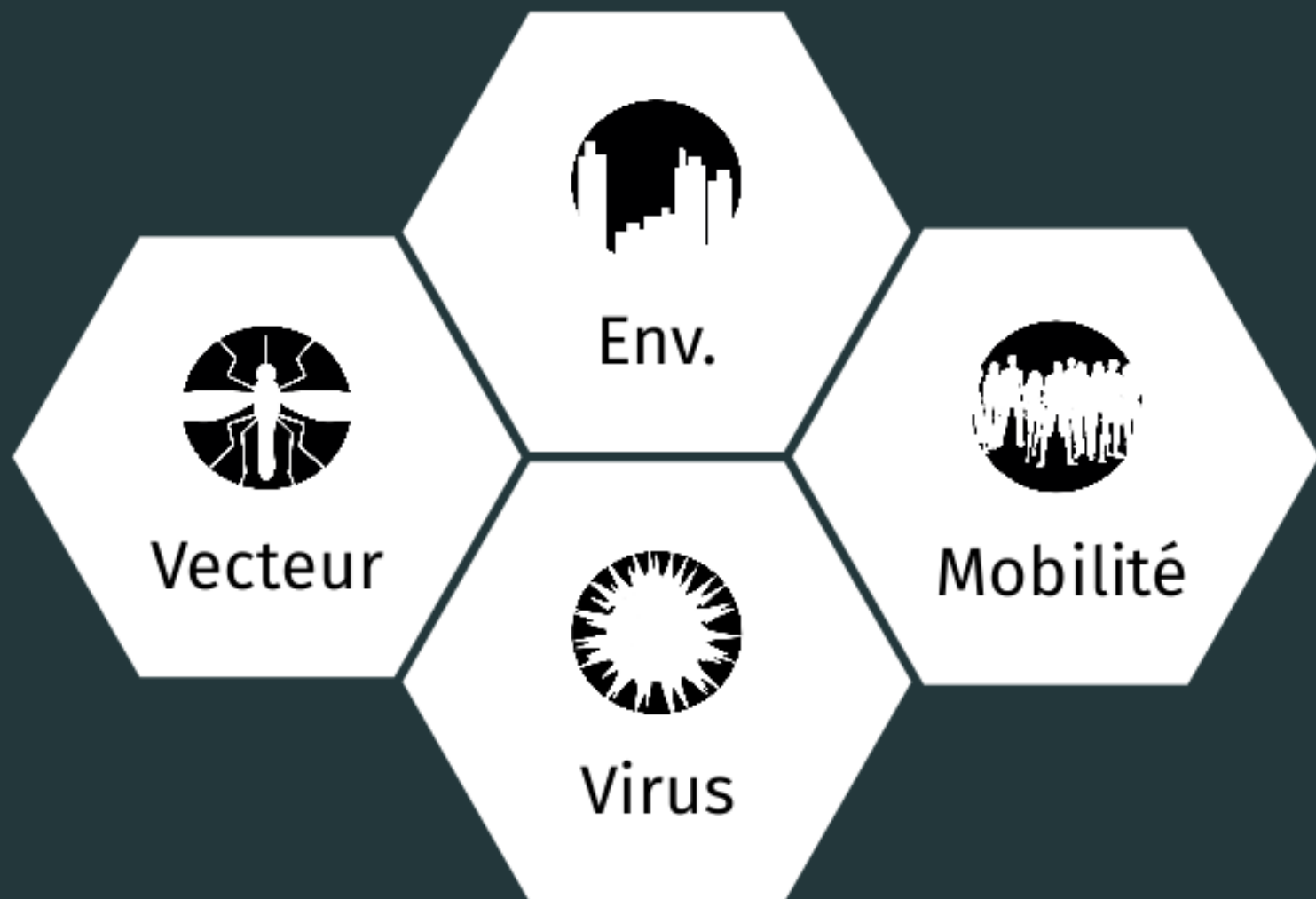
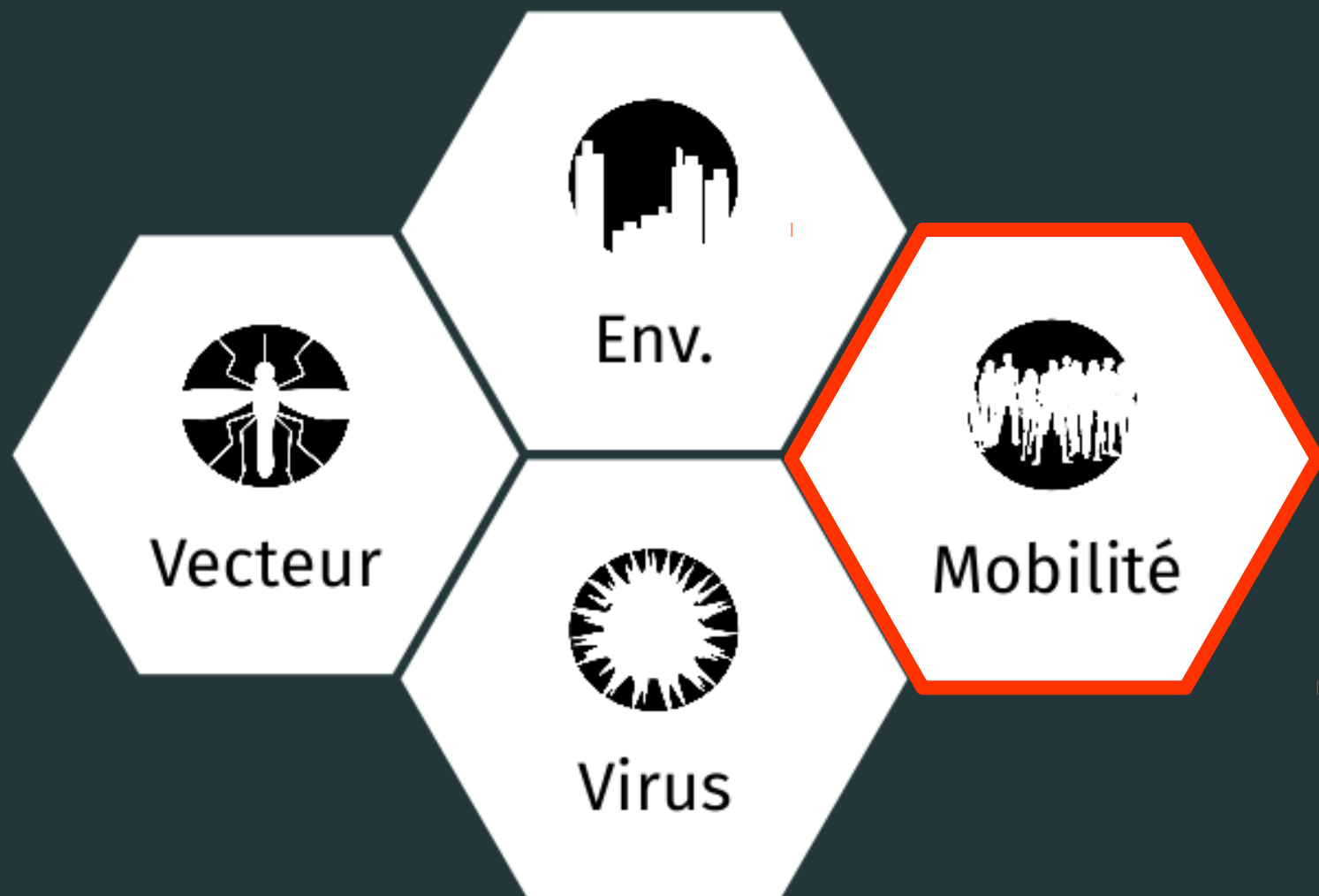


Figure 4 : Incidence mensuelle des cas de dengue – 2005-2014 (source : BMA)

La dengue, un système complexe...



La dengue, un système complexe...



Quelles données pour la mobilité à Bangkok?

Pas d'EMD disponibles

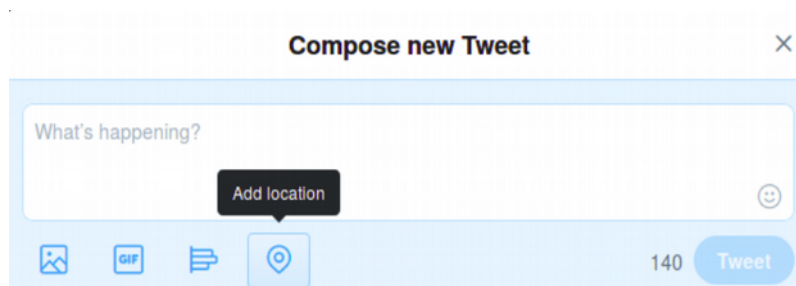
Enquêtes terrain délicates (~10M d'habitants)

Données téléphoniques : difficilement accessibles

Réseaux sociaux : Twitter

~ 310 Millions d'utilisateurs mensuels

3eme réseau social en Thaïlande (alexa), 4eme d'après google trend



Liste de lieux – Mall, ville, commerce etc..

Localisation précise – position GPS du téléphone

Collecte des données - Twitter

API Stream \Rightarrow au max 1 % des flux totaux

\Rightarrow Géolocalisation précise

Nom d'utilisateur ; Date & localisation du message

+ post par Instagram & foursquare

Entre Juin 2014 & décembre 2015 :

~ 307K utilisateurs pour ~30M de tweets

Supprimer des bots – critères de vitesses

Détection les domiciles

- Regroupement des Tweets (Dbscan)

- Dans des zones résidentielles (hors Mall, parcs, etc. via OSM)

- Un nombre et % de tweets entre 20h et 8h > seuil (hors lieux Instagram & Foursquare)

Bangkok ~ 25K personnes / 13M de tweets

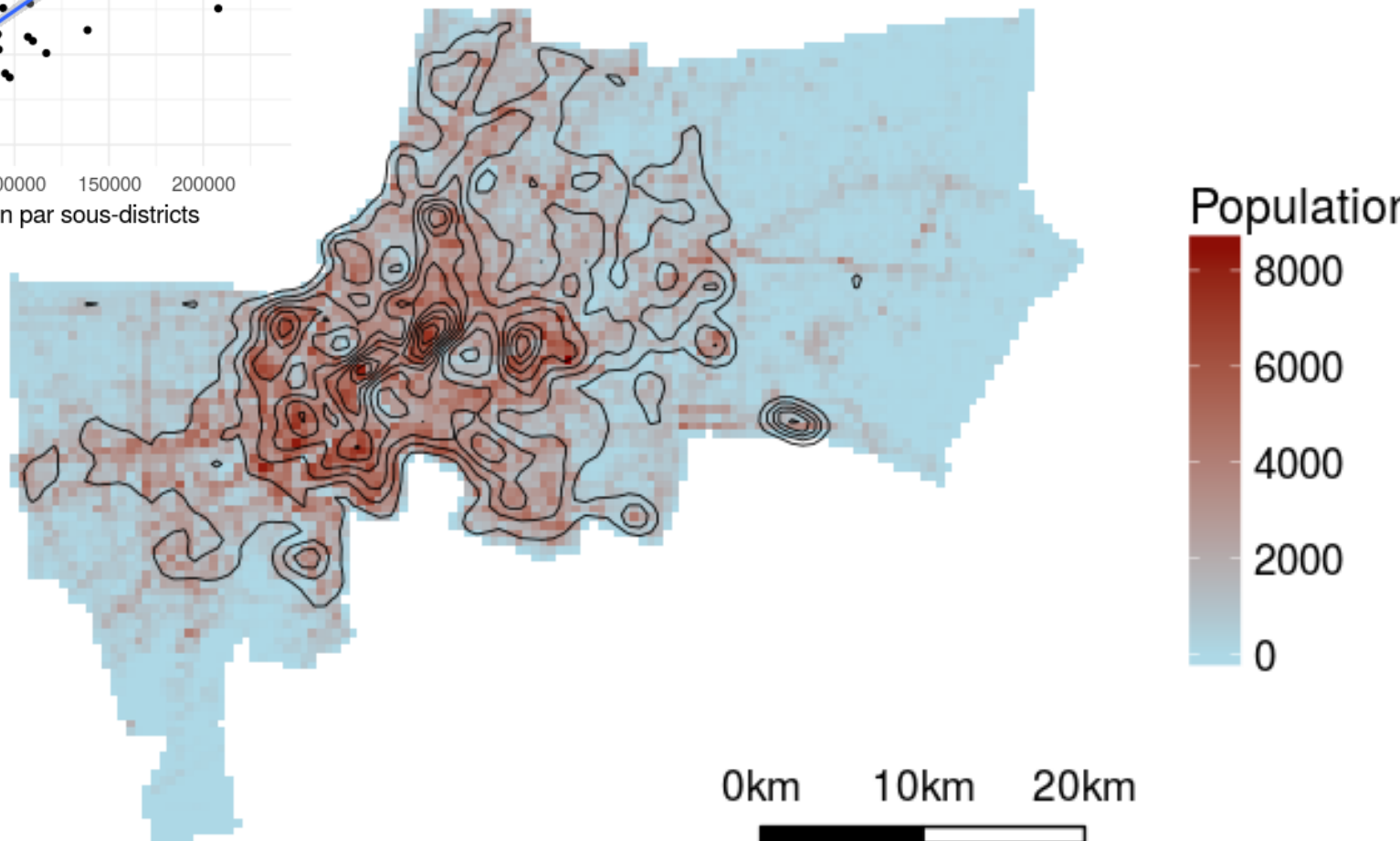
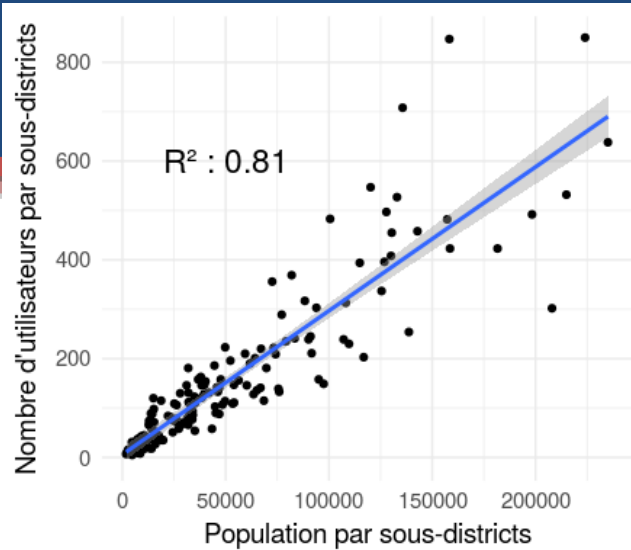
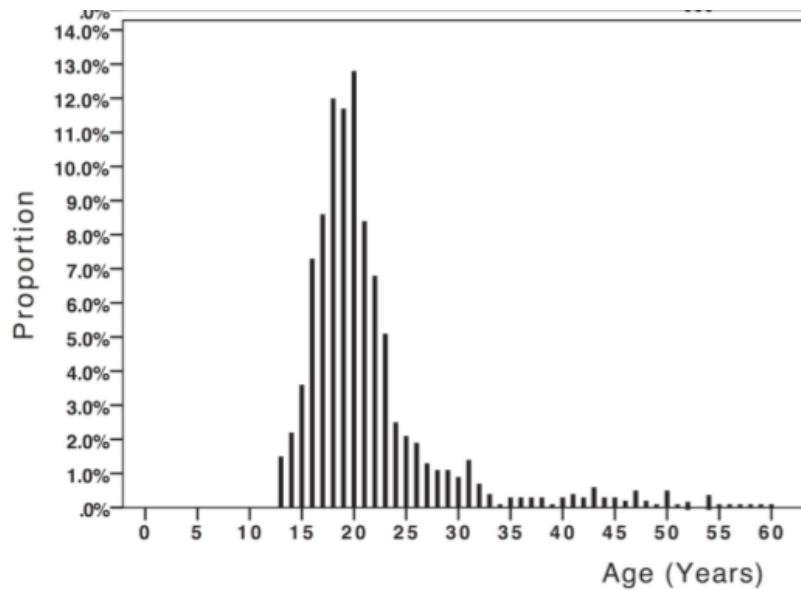


Figure 5 : Répartition de la population à Bangkok (bleu ciel \Rightarrow rouge, Landsat 8 & Census, Misslin & Daudé 2016) & densité de domicile d'utilisateur de Twitter (isolignes)

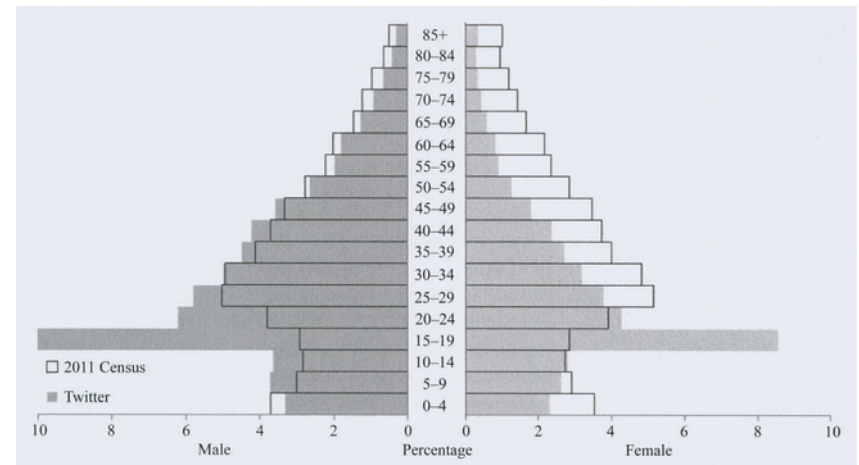
Age des utilisateurs

UK :



Sloan et al, 2015 - *PloS ONE*

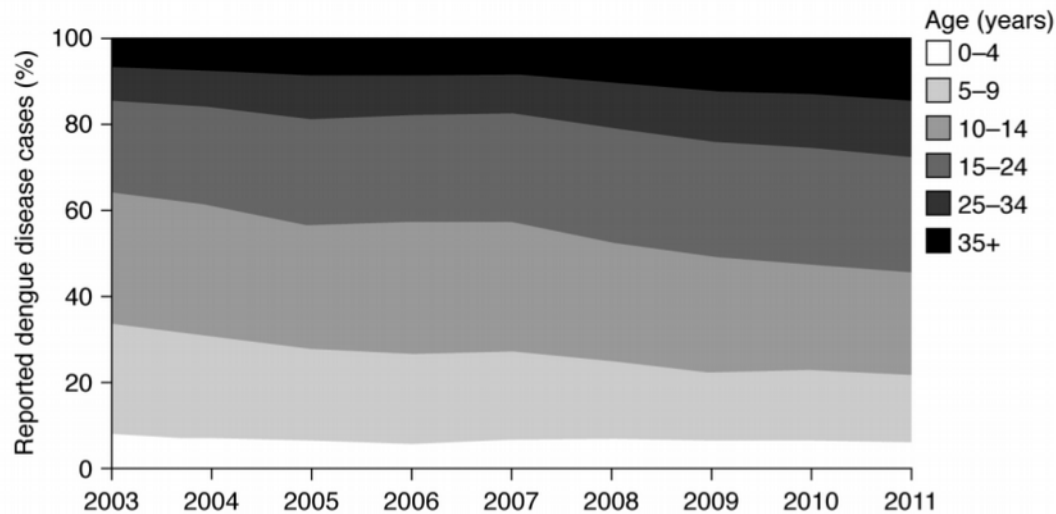
A Londres :



Longley et al, 2015 - *Environment and Planning*

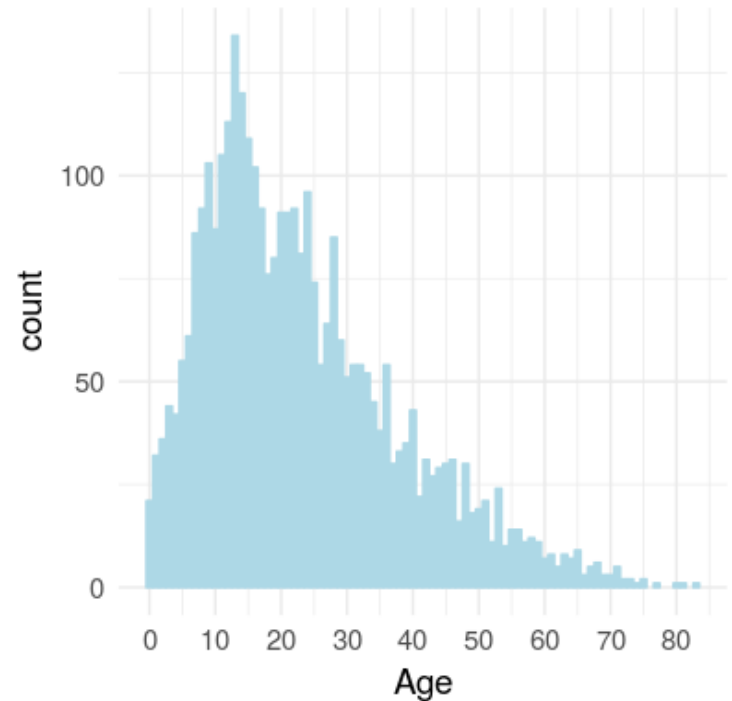
Age des malades de la dengue

Thaïlande



Limkittikul K., et al, 2014 - PLoS Neglected Tropical Diseases

Bangkok



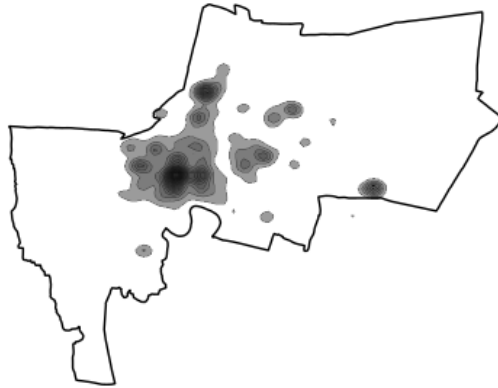
Nombre de cas présentant des symptômes dengueux investigué par la BMA en 2013

Déplacements quotidiens

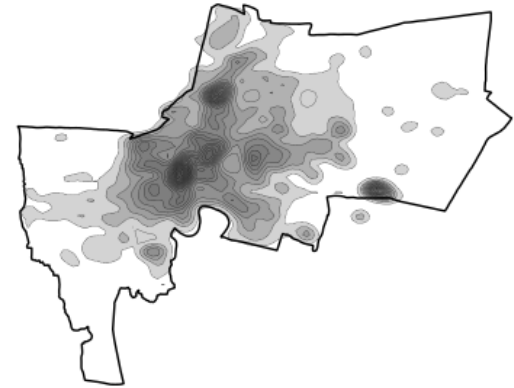
Tuesday 6-8



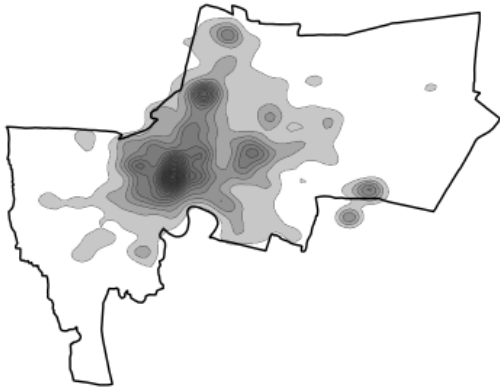
Tuesday 12-14



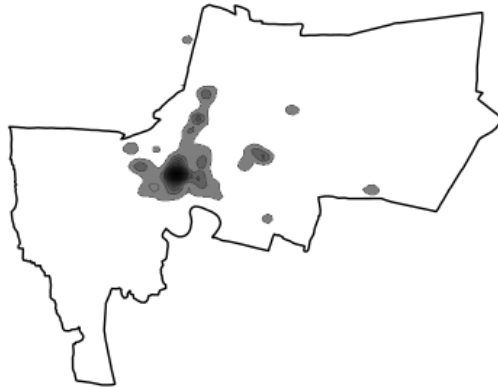
Tuesday 22-00



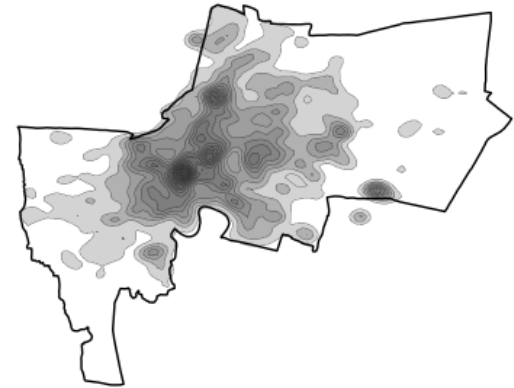
Sunday 6-8



Sunday 12-14



Sunday 22-00



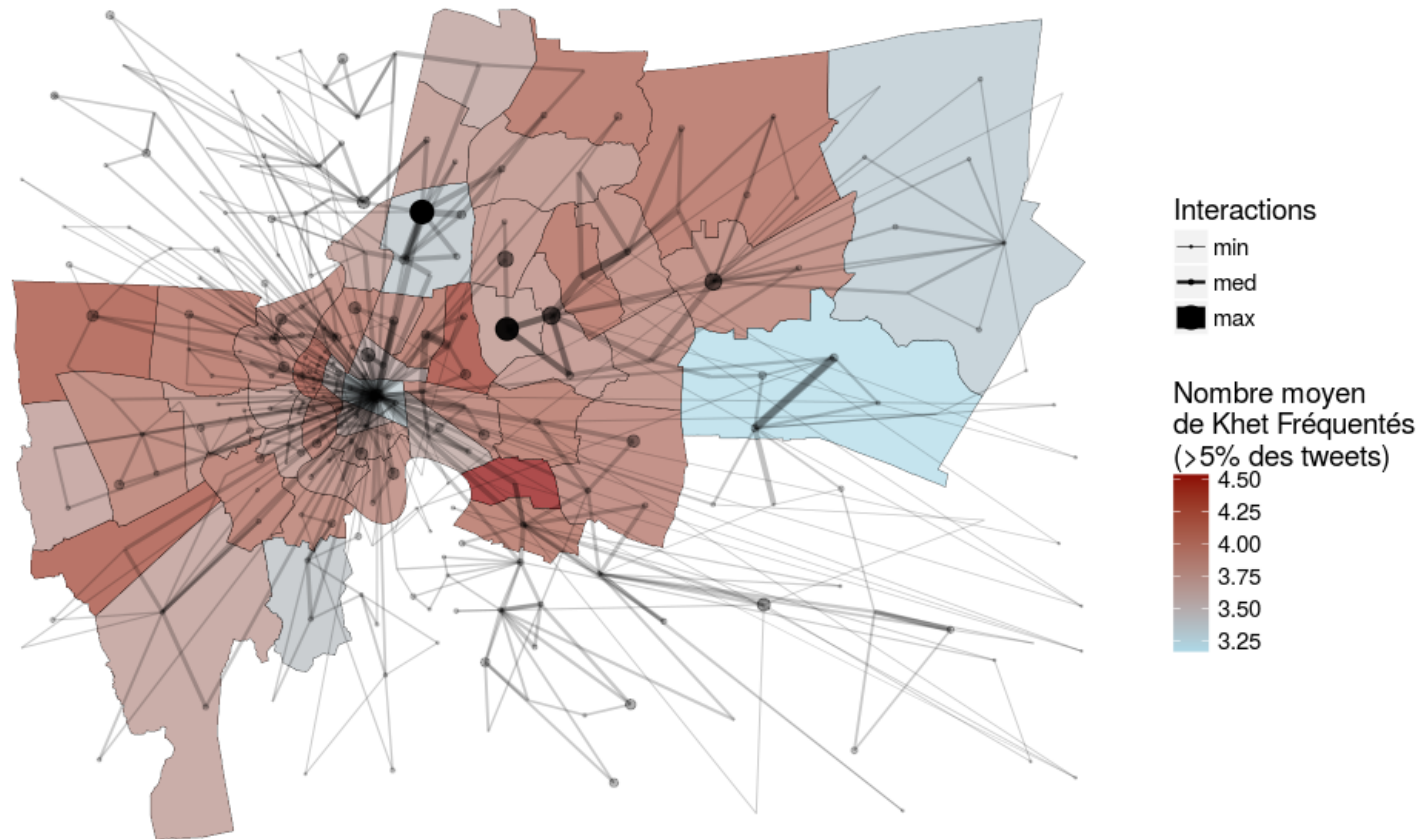
Density
max
min

Interactions spatiales

Par sous-district :

Pour chaque utilisateur \Rightarrow part de tweets dans les autres sous-districts

\Rightarrow Agrège



Activités à Bangkok

Fleury et al. – 2012 - *Définir les centralités commerciales au cœur d'une grande métropole : le cas de Paris intra-muros* – Cybergéo

À partir de la base de donnée des commerces à Paris (IRIS) + CAH

Aspect Morphologique :

- Densité, diversité, spécialisation et différentiel local

Aspect Fonctionnel :

- Typologie des activités commerciales

A Bangkok : pas de telle base de données (à notre connaissance)

Activité commerciales :

- ⇒ Donnée Google

Activité non commerciales (Ecole, Université, Hopitaux, etc.)

- ⇒ Google + Open Street Map

API Google Place :

Recherche de lieux dans un rayon donné

Système de priorité : « petit lieux » souvent moins visible

⇒ rayon de recherche très faible

⇒ Recherche fenêtre mobile (~200m)

⇒ Nom / localisation / catégories

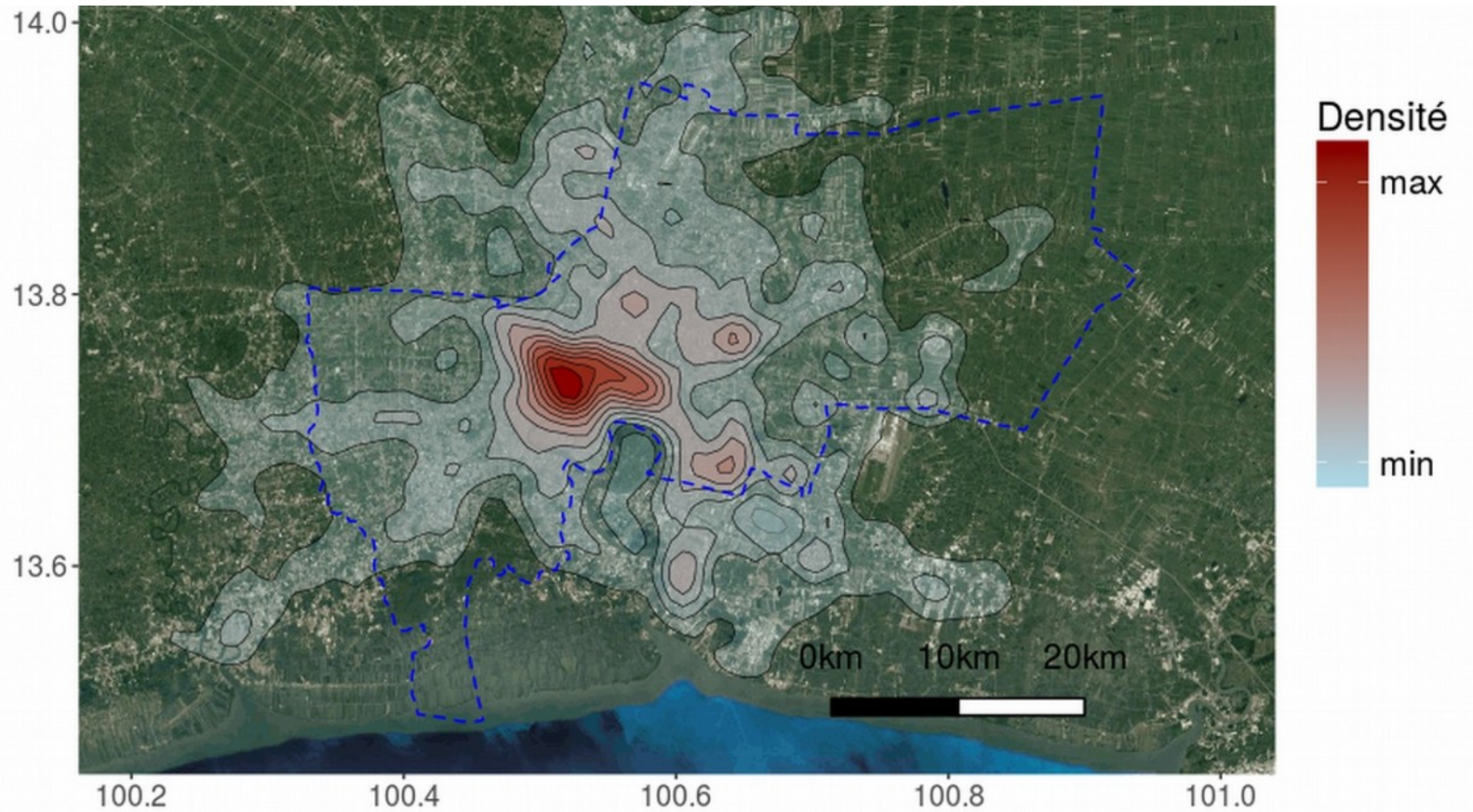
⇒ Suppression des doublons

⇒ Nombre de requêtes limitées
(2500 requêtes par jour & par clé)

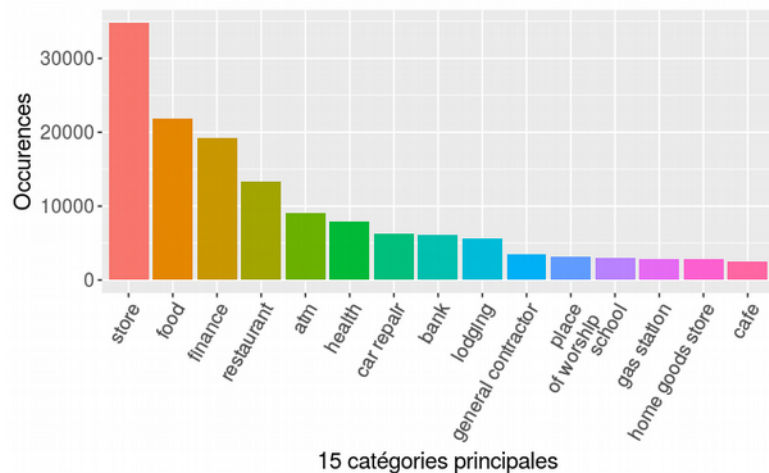
⇒ 143 430 points exploitables

```
1:
  geometry:
    location:
      lat: 13.7461272
      lng: 100.6032677
    viewport: Object
  icon: "https://maps.gstatic.com"
  id: "73b9bc395ce1cb5c13d1d0e"
  name: "Sophia Restaurant"
  opening_hours: Object
  photos: [1]
  place_id: "ChIJYYbHfDme4jAR7dyipv"
  rating: 4.3
  reference: "CmRRAAAAlpPuSPYrnGnFVbi"
  scope: "GOOGLE"
  types:
    0: "restaurant"
    1: "food"
    2: "point_of_interest"
    3: "establishment"
  vicinity: "1 ซอย รามคำแหง 5, Ramk
2:
```

Données Google

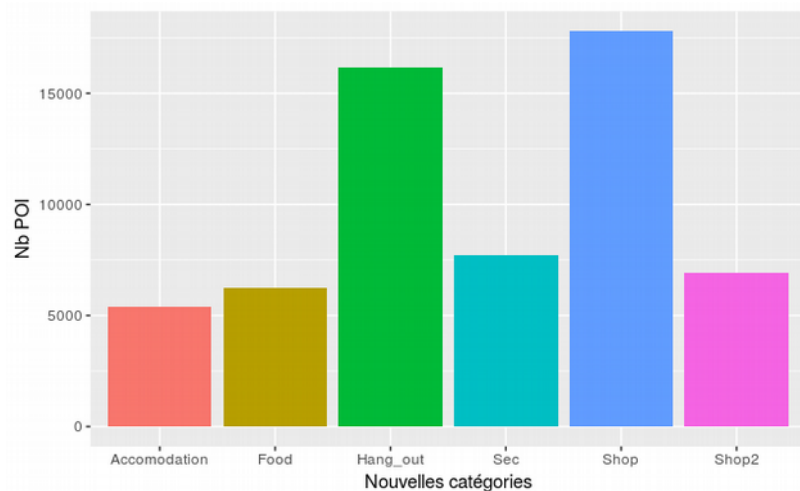


Données Google : activités commerciales



Regroupement en 6 grandes classes :

- **Lieux d'hébergements** : Hôtel, guest houses, condo etc.
- **Lieux de restauration non différenciés** (« Food »)
- **Lieux de sorties** : Restaurant, bars, boîtes de nuits, café
- **Entreprises** - « company », « co. Ltd » et liés aux voitures (car, car shop) etc.
- Commerces indifférenciés** : « Shop »
- Commerces différenciés** – furniture store, electronic store, clothing, etc.



Classification des activités commerciales

1) Regrouper les POI \Rightarrow Quelle unité spatiale ?

\Rightarrow Cellule de base de 30m^2

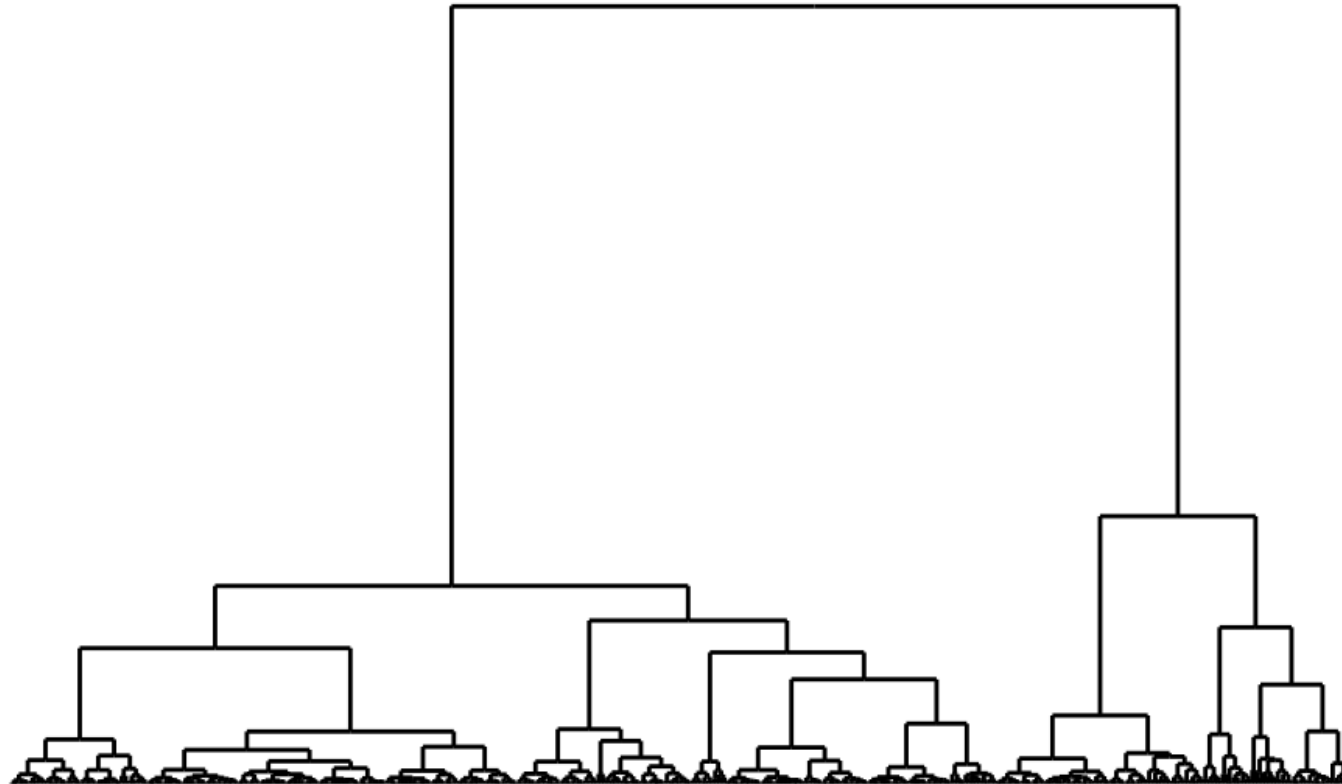
\Rightarrow Images Landsat8 // déplacement du moustique dans le modèle

180m : « bon » multiple de 30 \Rightarrow ~ un pâté de maison

30913 mailles avec des 66,8K POI

2) Regroupement de maille \Rightarrow Kmeans 500 classes

3) CAH





- ⇒ Discriminer au mieux les zones aux fortes densité de POI
- ⇒ Arriver à nuancer les zones avec peu de POI
- ⇒ Si on coupe trop bas ⇒ trop de classes

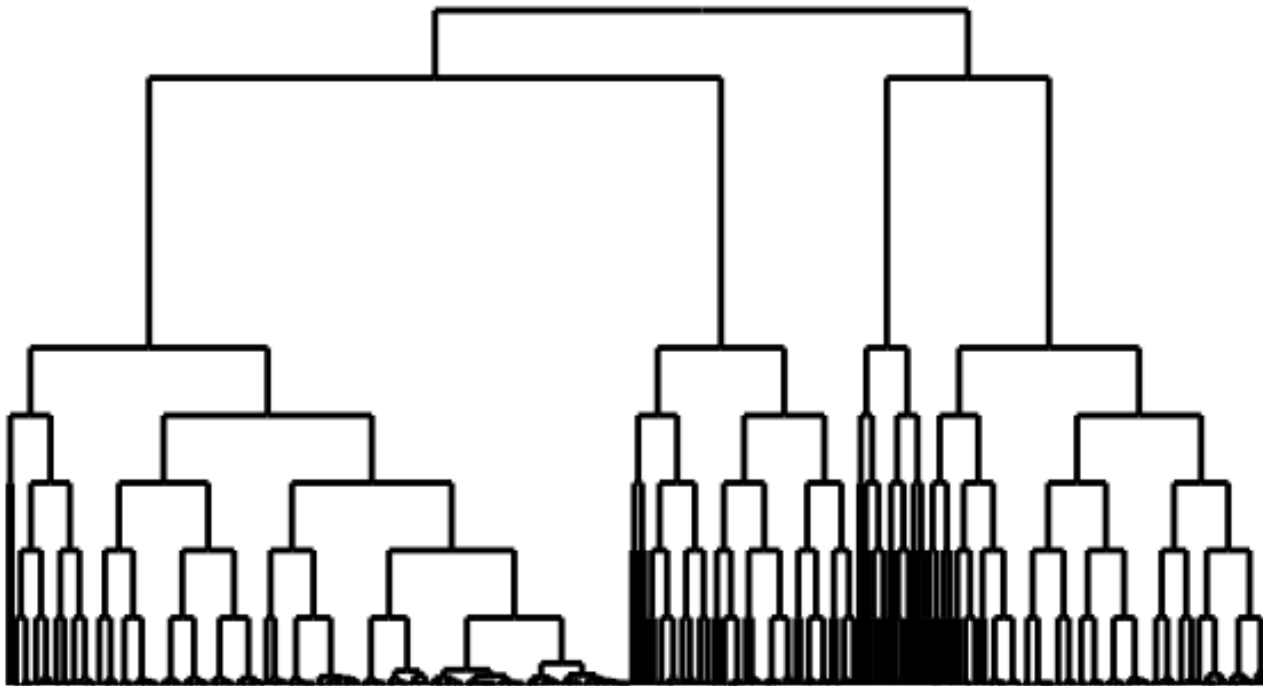
Dendrogramme adapté

⇒ Garder la structure hiérarchique en grappe



Dendrogramme adapté

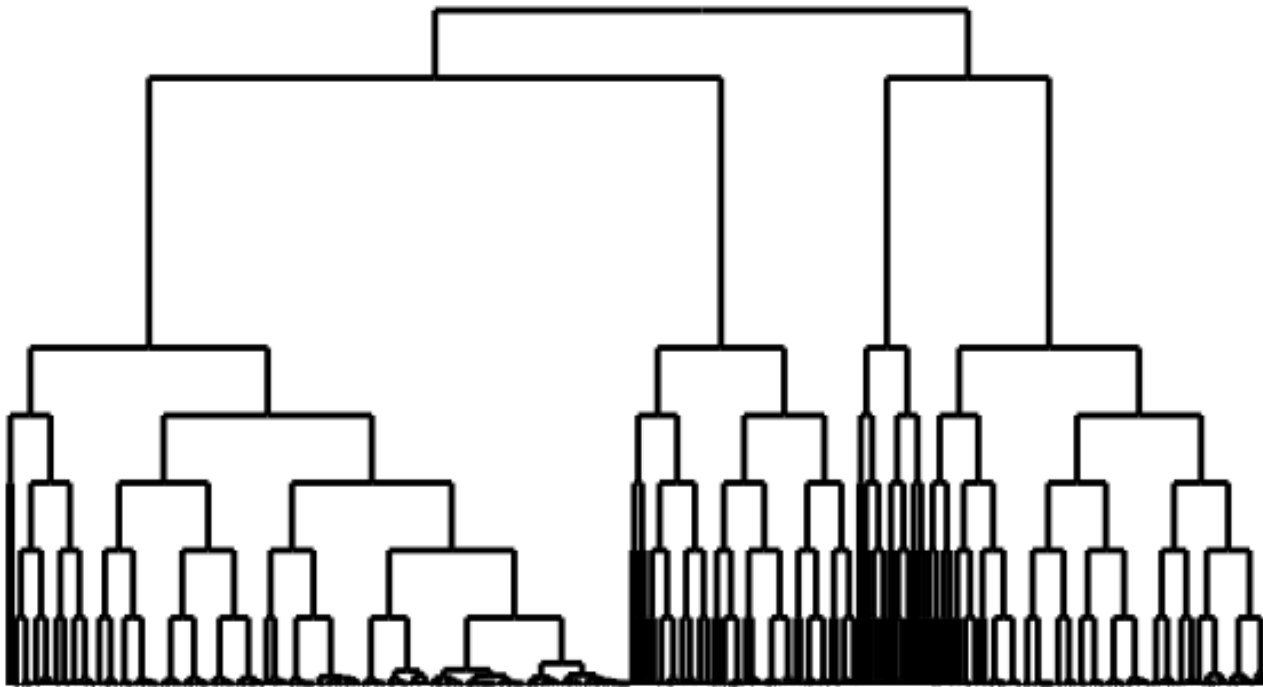
- ⇒ Garder la structure hiérarchique en grappe
- ⇒ Homogénéisation des niveaux



Dendrogramme adapté

⇒ Garder la structure hiérarchique en grappe

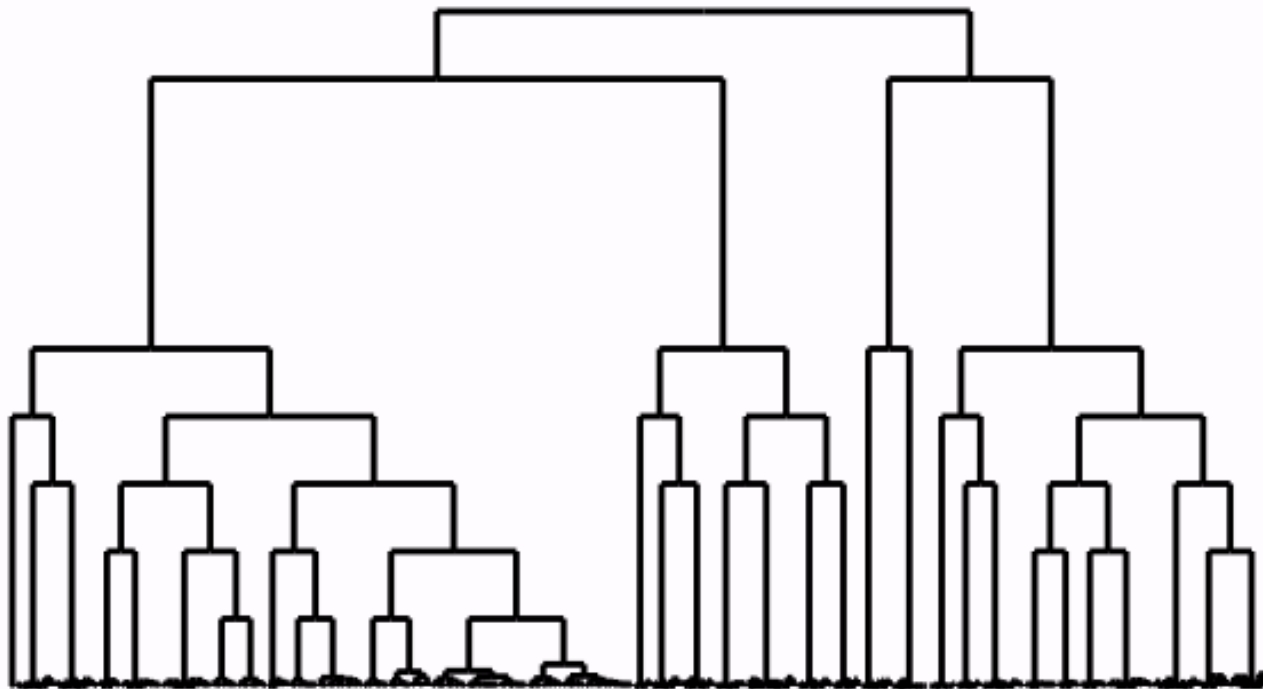
⇒ Pour chaque branche, à partir du niveau 2 : Si la somme des cluster dans les branches inférieures est inférieure à un seuil ⇒ fusion de ces branches



Dendrogramme adapté

⇒ Garder la structure hiérarchique en grappe

⇒ Pour chaque branche, à partir du niveau 2 : Si la somme des cluster dans les branches inférieures est inférieure à un seuil ⇒ fusion de ces branches

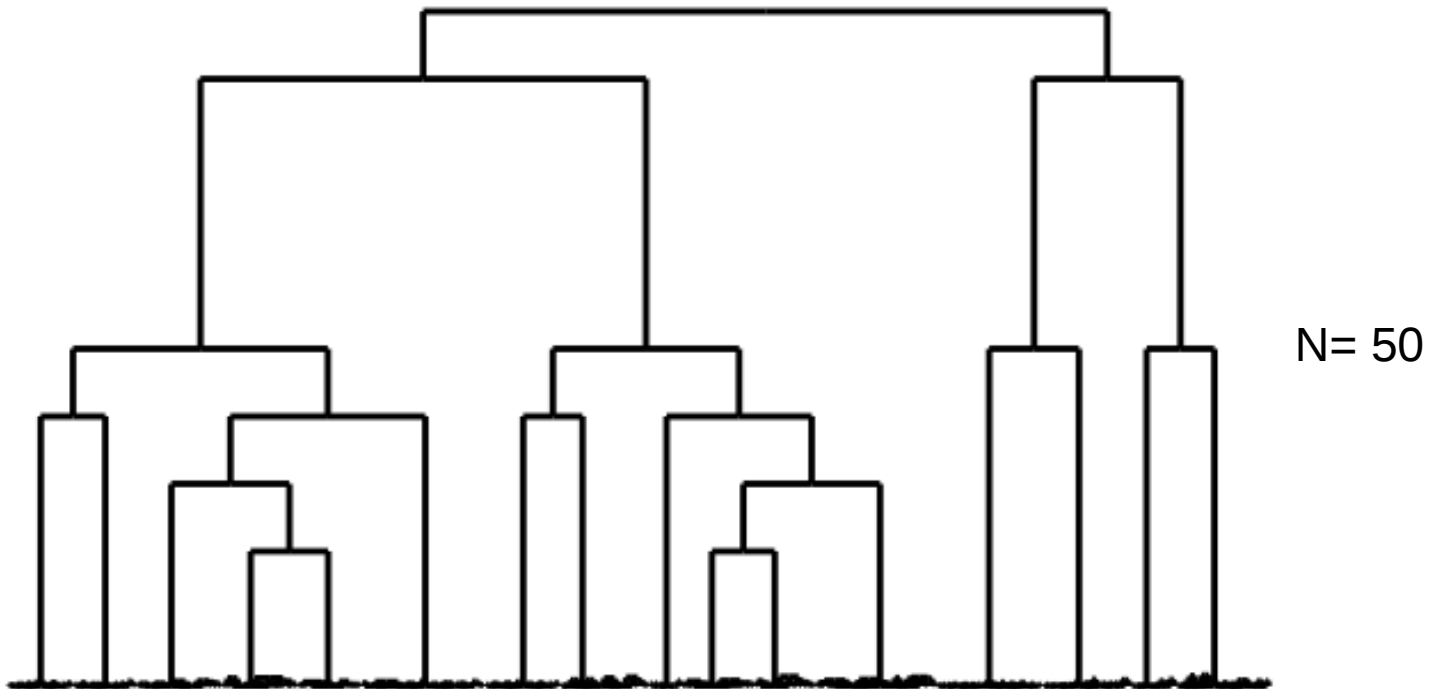


N= 20

Dendrogramme adapté

⇒ Garder la structure hiérarchique en grappe

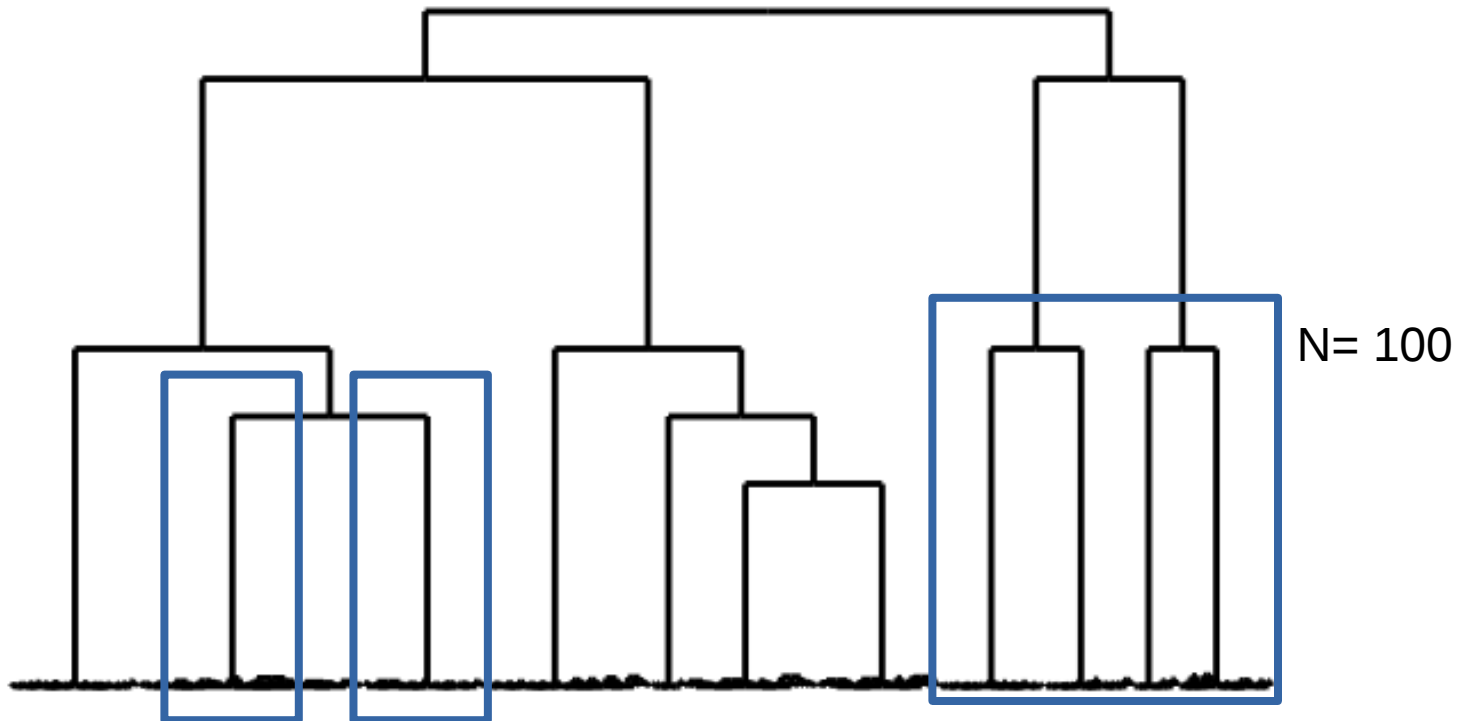
⇒ Pour chaque branche, à partir du niveau 2 : Si la somme des cluster dans les branches inférieures est inférieure à un seuil ⇒ fusion de ces branches



Dendrogramme adapté

⇒ Garder la structure hiérarchique en grappe

⇒ Pour chaque branche, à partir du niveau 2 : Si la somme des cluster dans les branches inférieures est inférieure à un seuil ⇒ fusion de ces branches



Dendrogramme adapté

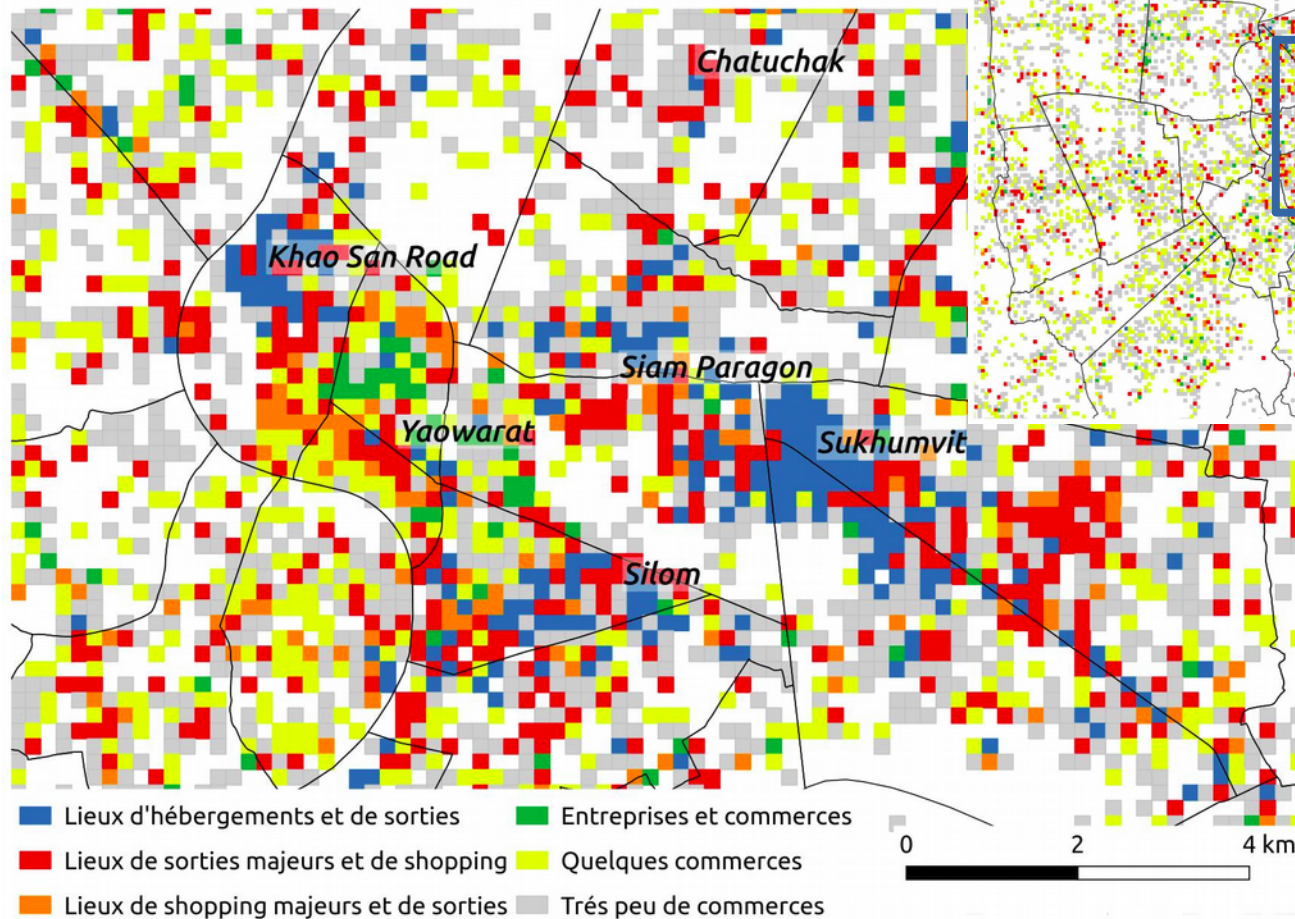
Méthode supervisée – plutôt expérimentale

(-) Perte de l'information sur l'inertie

(+) Garde les classes avec peu d'individus & segmente celles avec beaucoup d'individus

Classification des zones commerciales

⇒ Regroupement thématique en 6 classes



Bonne cohérence
avec la connaissance
du terrain

Classification des zones commerciales

(+) Résultats cohérent avec les connaissances du terrains

(+)Discrimine assez bien les différentes entités et fait ressortir les principaux quartiers

(-) Pas d'études de sensibilité ; notamment sur l'effet de la taille de la maille

(-) Dépend des catégories initiales en entrée

(-) Utilisation *ad hoc* \Rightarrow pas comparable avec d'autres typologies d'utilisation du sol

Données Google & Open Street Map

Ecoles
Universités
Religion & lieux de cultes
Hôpitaux / santé
Administrations
Loisirs

Google : points
OSM : polygones

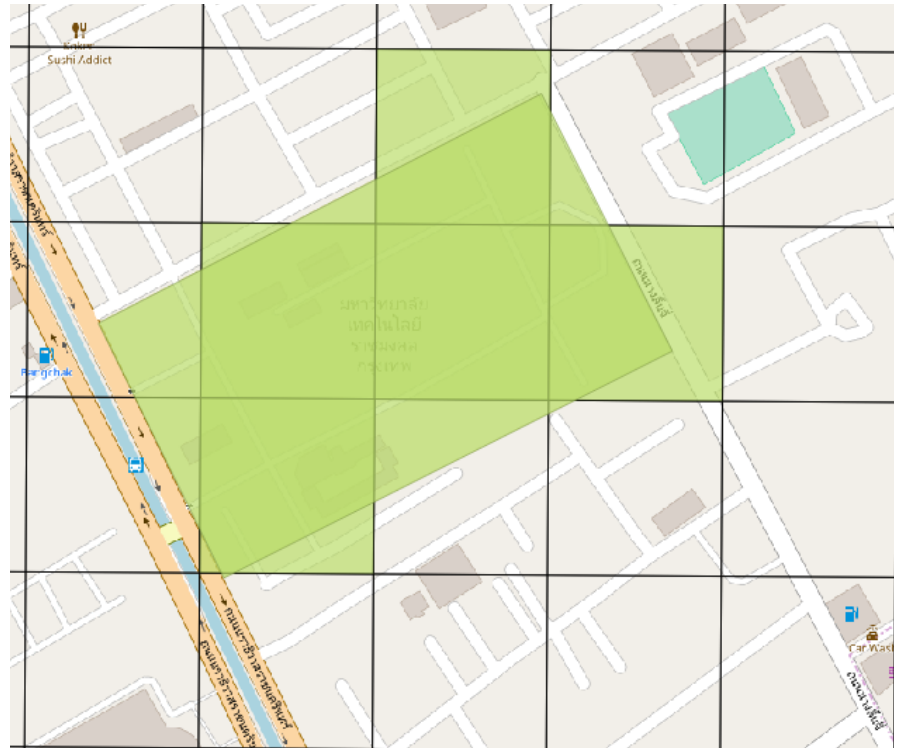


Données Google & Open Street Map

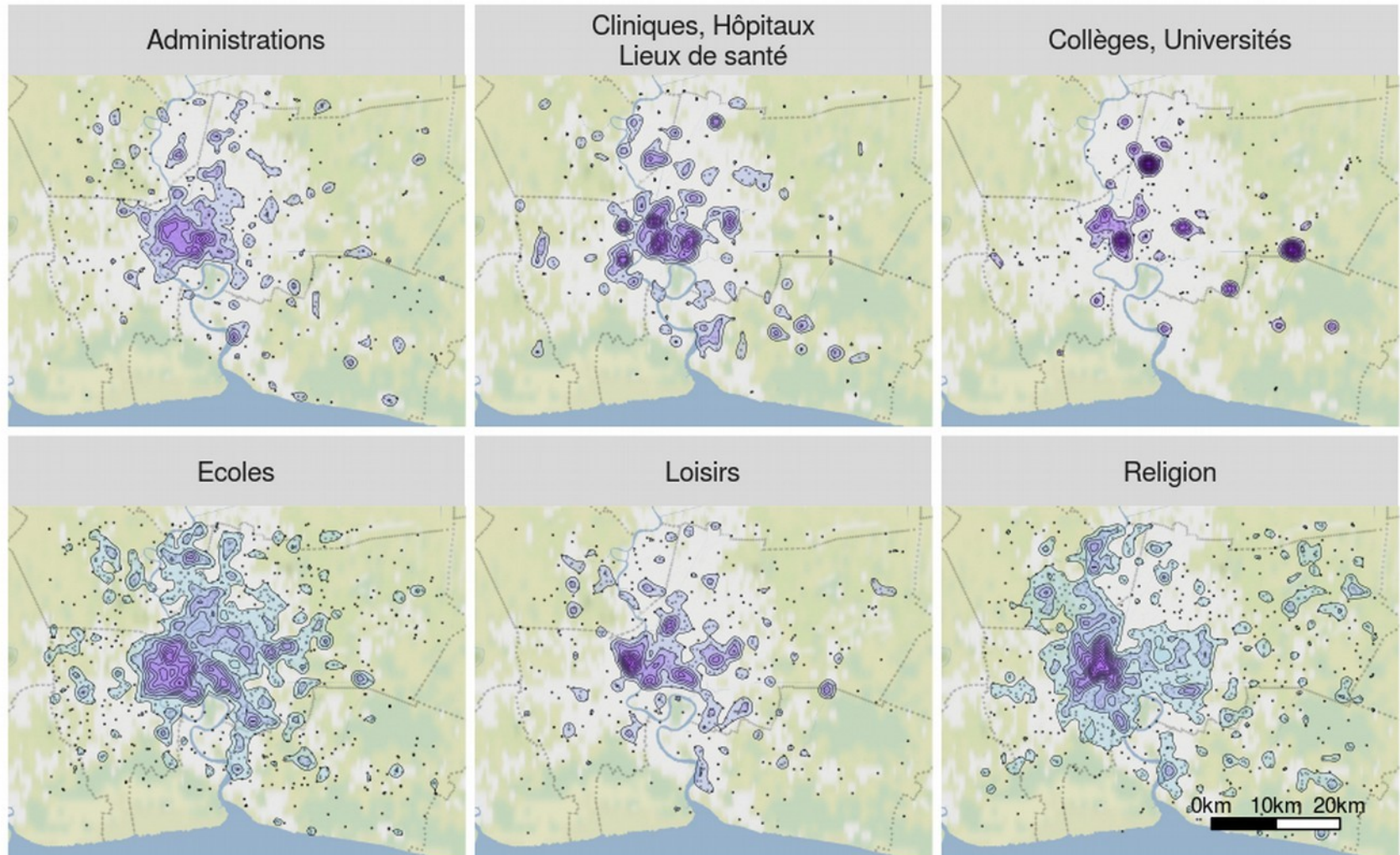
Ecoles
Universités
Religion & lieux de cultes
Hôpitaux / santé
Administrations
Loisirs

Google : points
OSM : polygones

Regrouper les POI et
intersection OSM / grille



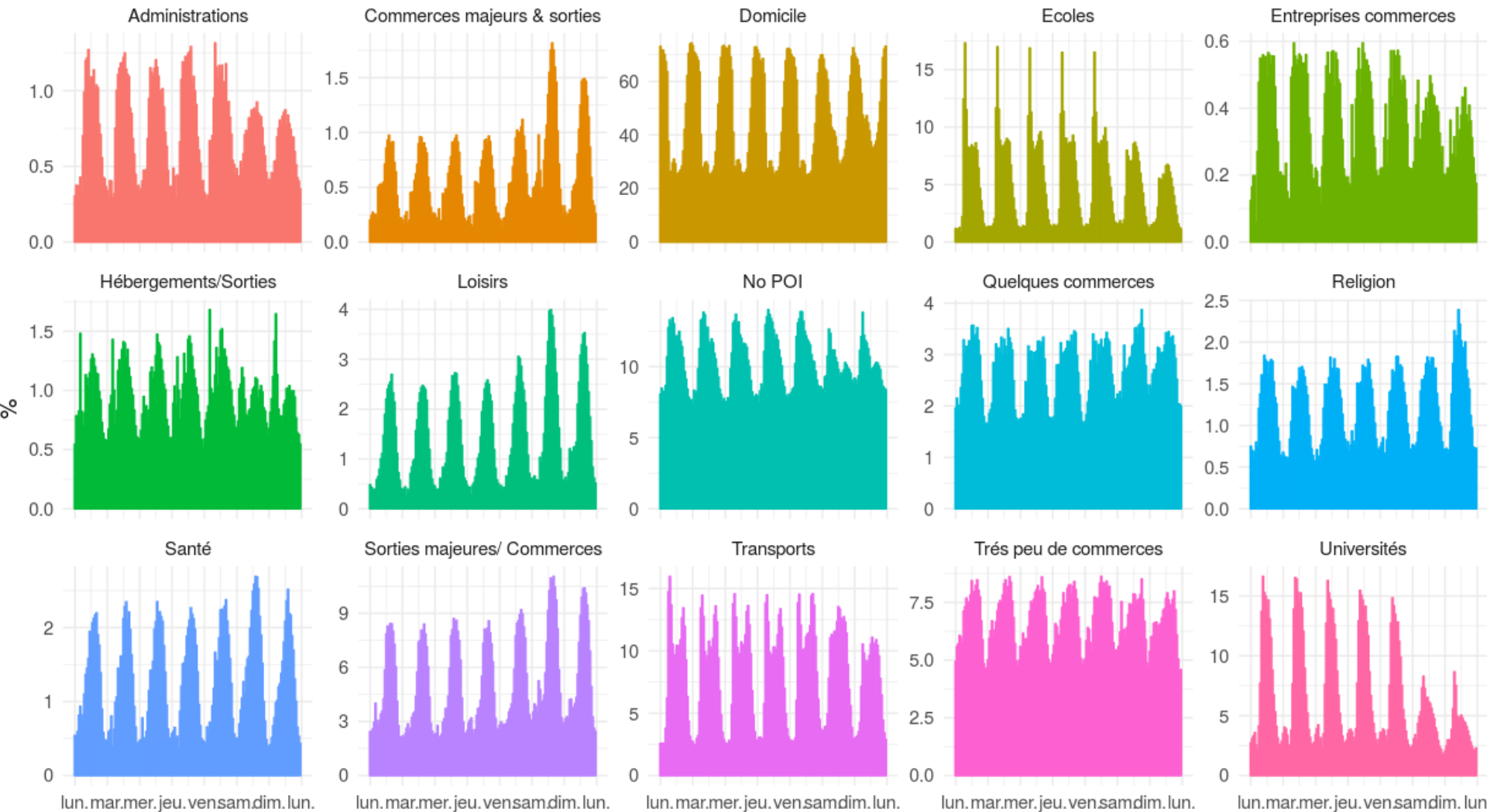
Données Google & OSM



Densité

max

Temporalités des différentes activités



Un ABM pour les mobilités individuelles

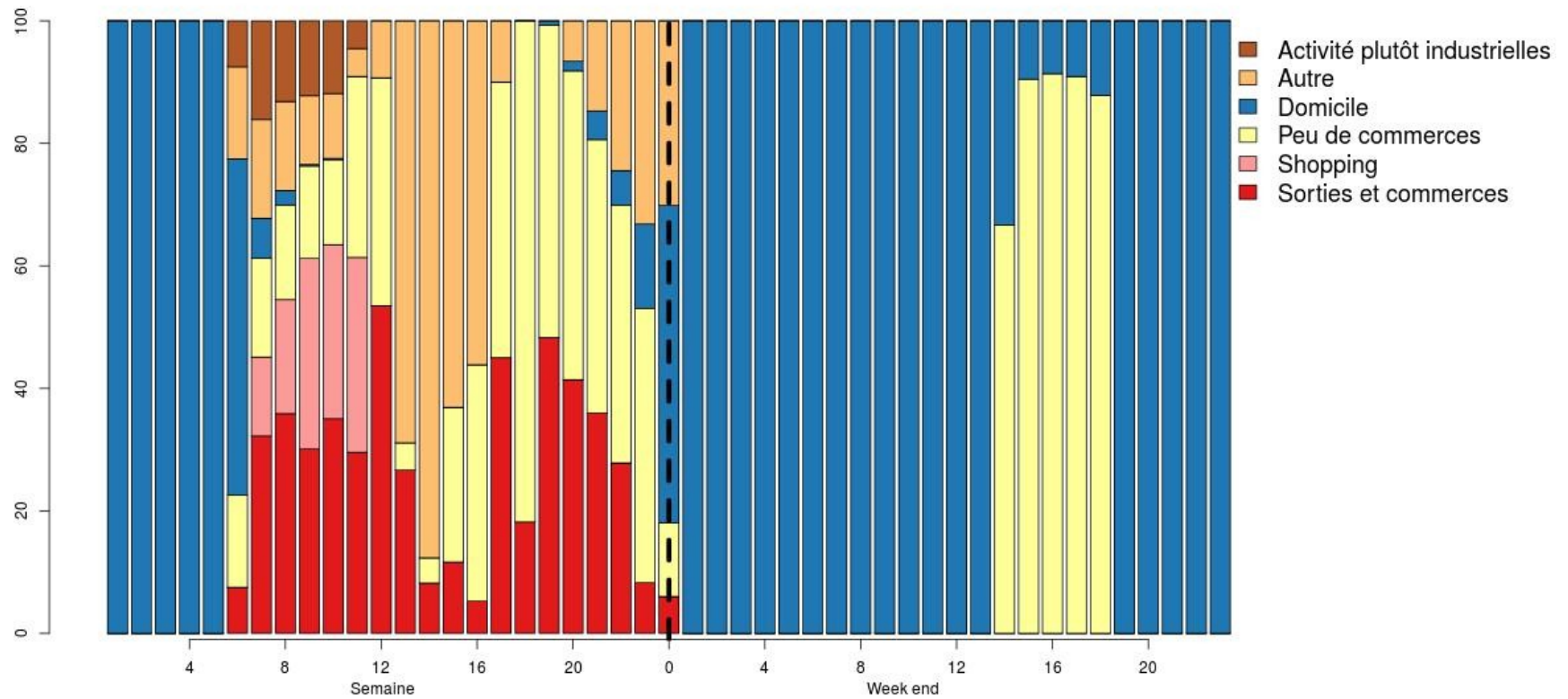
Par tranche d'âge :

<15 ans : règles de mobilités simples (Domicile / Ecole / Loisirs)

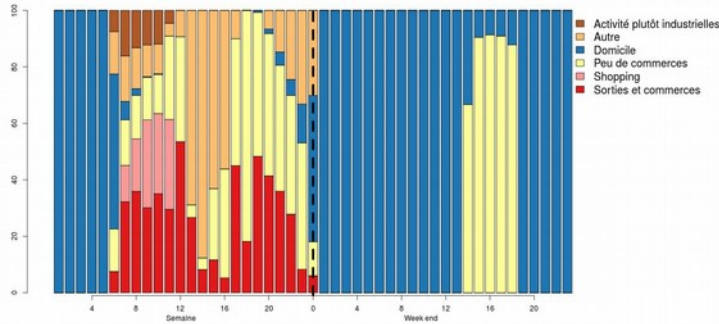
15 - 35/40 : Inspiré de Twitter ; calibré avec le Time use census

+35 : Time use & temps de transports

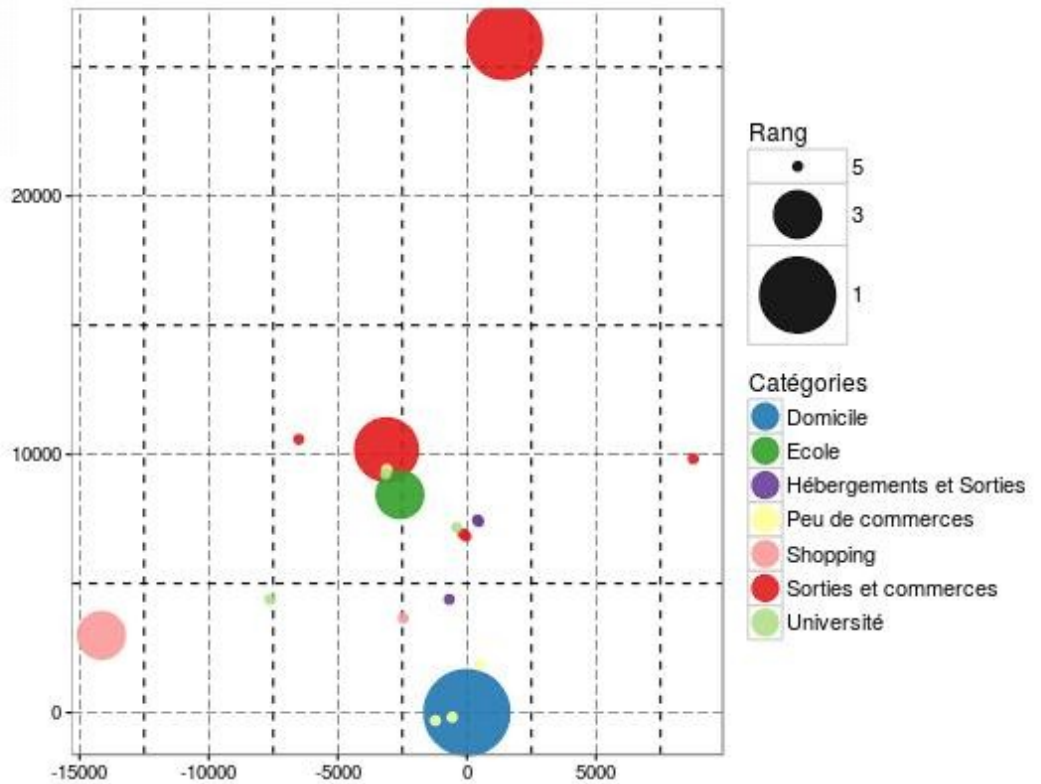
Agendas individuels



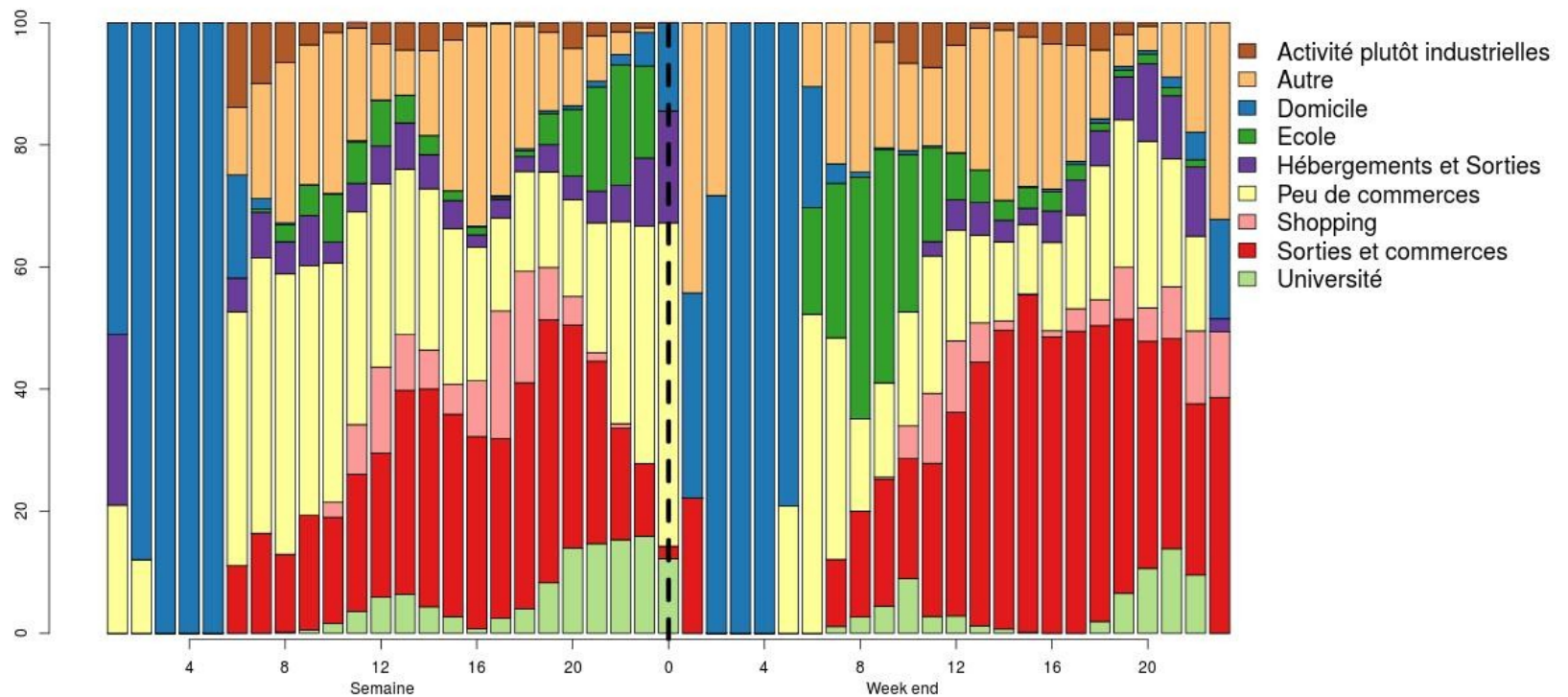
Agendas individuels



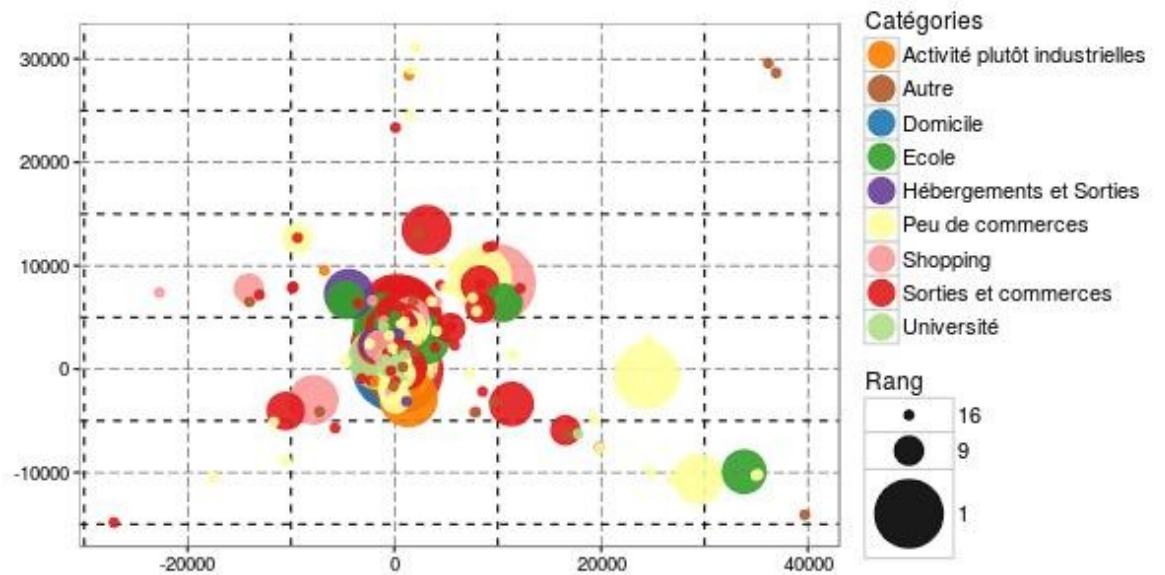
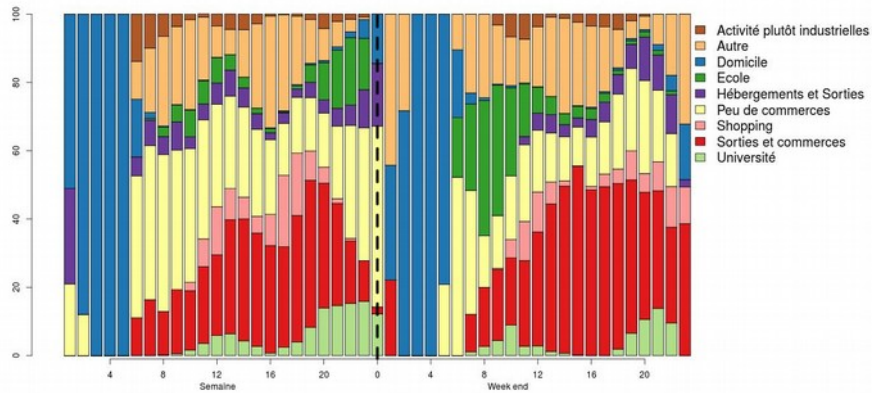
Huraux et al.
Théoquant, 2017



Agendas individuels



Agendas individuels



Implémentation dans MO³

Simulation: mo3_model:MO3model - MO3/models/mo3.gaml - Gama

File Edit Search Experiment Agents Views Help

Experiment: 59 cycles elapsed

Model mo3 / Experiment MO3model Models


- Library models (225)
 - Features (113 models)
 - Syntax (4 models)
 - Toy Models (78 models)
 - Tutorials (30 models)
- Plugin models (54)
- User models (12)

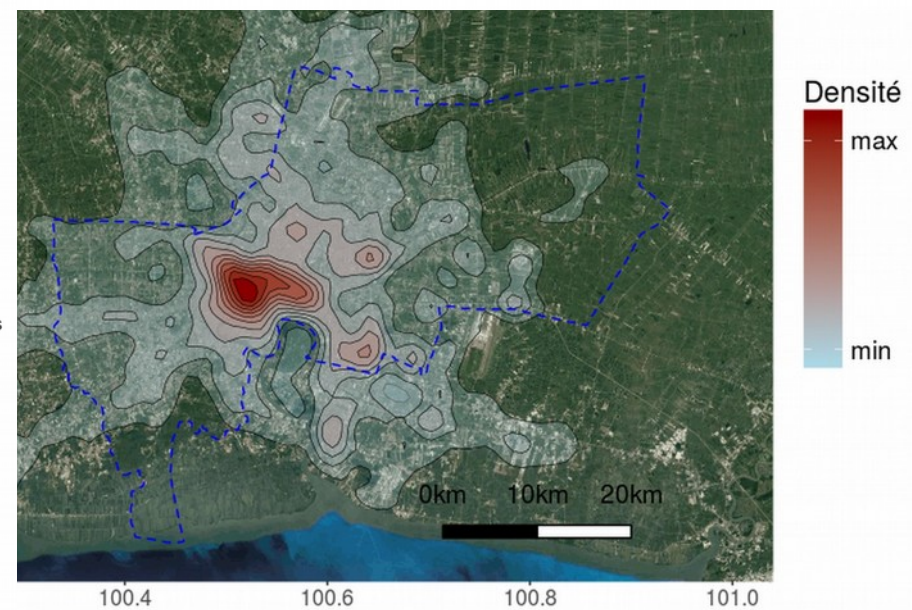
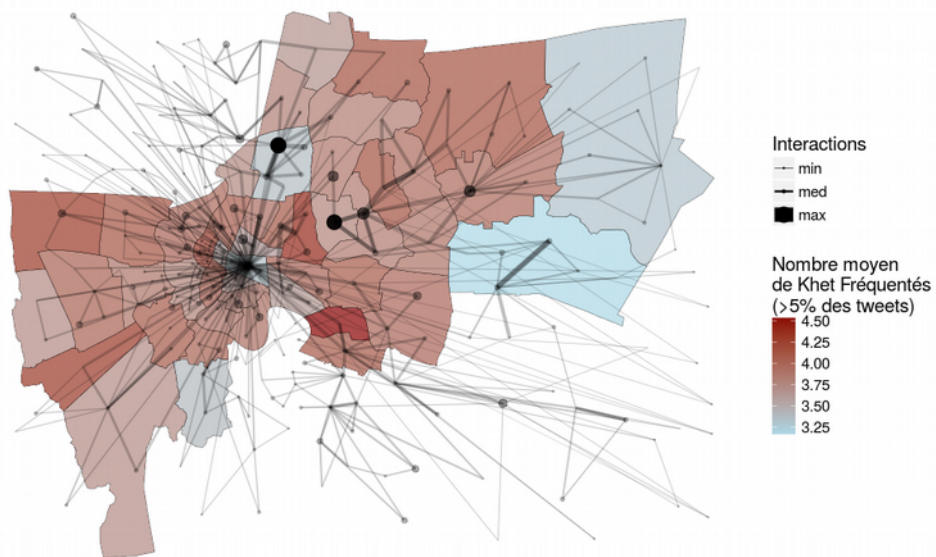
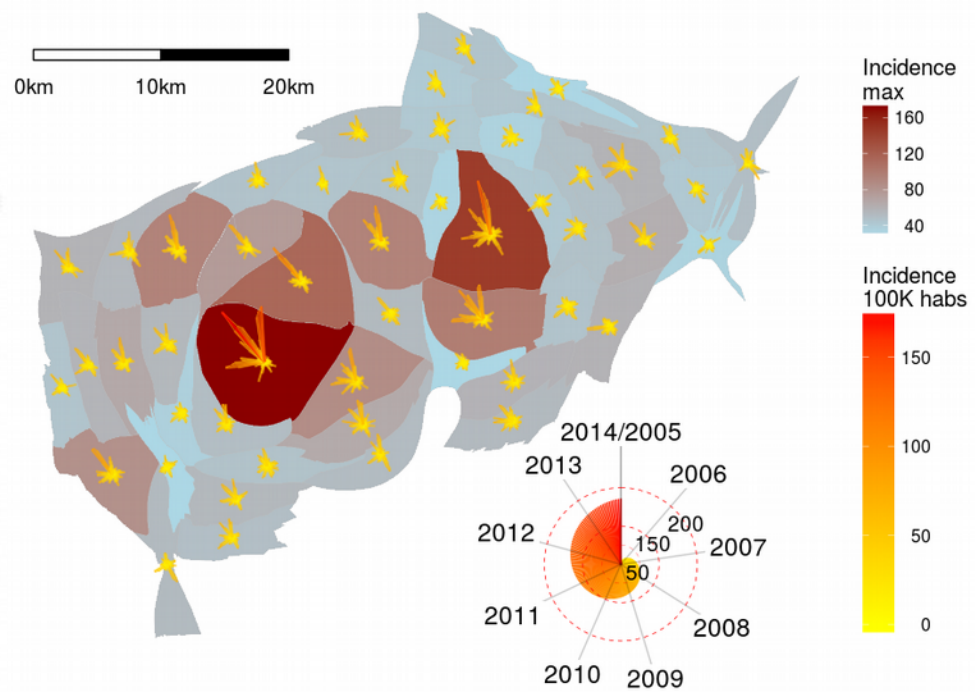
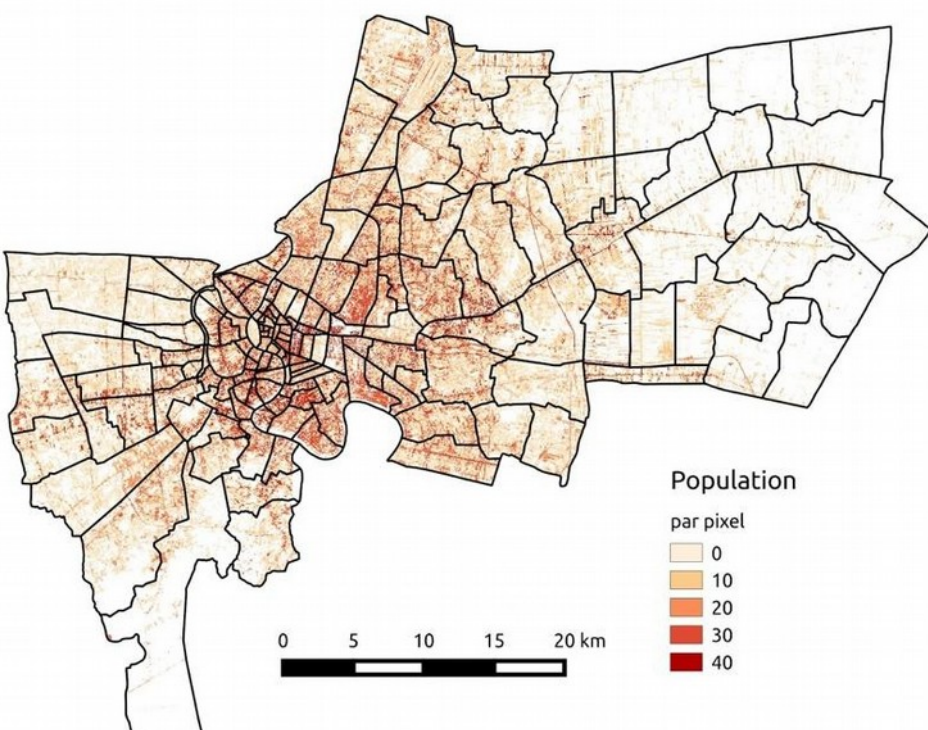
Interactive console mo3.gaml

MO3model Used in...

```
1121 {
1122   fNbNewP <- NOMINAL_DAILYSURV_PUPAE_RATE * f
1123   fListStkPtoA[i] <- (fListStkPtoA[i] + fNbNew
1124   fListStkPtoA[i + 1] <- 0; //vider stock of the
1125 }
1126
1127 if (fListStkLtoP[0] != 0)
1128 {
1129   addP <- (PUPATION_SURV * fListStkLtoP[0]);
1130   fListStkPtoA[dayNeedEmerg] <- (fListStkPtoA[
1131   fListStkLtoP[0] <- 0;
1132 }
1133
1134 return addP;
1135 }
1136 * This action is called by the reflex update stock to
1137
1138 { //---- ADULT: temperature dependency -----
1139   if (fSpatObjAirAvgTempC > ADULT_DAILYSURV_LOW_TI
1140
1141   {
1142     fAdultSurvTempRate <- cal_temp_surv_rate(fSp
1143   } else if (fSpatObjAirAvgTempC > ADULT_DAILYSURV
1144
1145   {
1146     fAdultSurvTempRate <- cal_temp_surv_rate(fSp
1147   } else if (fSpatObjAirAvgTempC >= ADULT_DAILYSUR
1148   {
1149     fAdultSurvTempRate <- 1.0;
1150   } else if (fSpatObjAirAvgTempC <= ADULT_DAILYSUR
1151   {
1152     fAdultSurvTempRate <- ADULT_DAILYSURV_LOW_TI
1153   } else if (fSpatObjAirAvgTempC >= ADULT_DAILYSUR
1154 }
```

Day 38





Limites et perspectives

Limites de l'étude

Utilisation du sol

Sensibilité – taille de la maille

Validation \Rightarrow comparer avec d'autres données (Facebook / Foursquare)

Mobilité :

\Rightarrow Twitter : intentionnalité du tweet :

\Rightarrow traces numériques convergent vers mobilité réelle ?

\Rightarrow Représentativité vis-à-vis de la population « Twitter Inspirée »?

\Rightarrow Mais on se focalise a priori sur la population cible de la maladie

A l'avenir : Améliorer le modèle de mobilités

Mobiliser d'autres données :

Check'in de Facebook \Rightarrow Attractivité

Time use census \Rightarrow améliorer les agendas (durée d'une activité + temps de transport)

Du Terrain ?

Merci

Contact:

a.cebeillac@free.fr / alexandre.cebeillac@etu-univ-rouen.fr

Equipe :

Eric Daudé

Thomas Hureau

Somsakun Maneerat

Renaud Misslin

Alain Vaguet

