

RECONSTITUER LES CONTOURS DE LA CARTE SCOLAIRE LORSQU'ILS SONT INCONNUS

Sylvain Maugis et Mustapha Touahir

PLAN

- CONTEXTE
- DONNÉES
- MÉTHODE
- RÉSULTATS
- PERSPECTIVES

PLAN

CONTEXTE

CONTEXTE

- En France, le lieu de résidence d'un élève détermine le collège où il est censé être scolarisé : on dit que l'élève est « sectorisé » sur ce collège.
- La carte scolaire est déterminée par les collectivités territoriales (les départements pour les collèges, les communes pour les écoles). Le collège de scolarisation n'est pas connu au niveau national.
- Objectif de notre article : proposer une méthode qui permette de retrouver le collège de secteur quand celui-ci est inconnu.

PLAN

DONNÉES

DONNÉES (1/2)

- Le « fichier géolocalisé 2015 » recense tous les élèves scolarisés en France à la rentrée 2015. (Le même fichier existe pour d'autres rentrées). Il contient des informations sur
 - L'établissement de scolarisation de chaque élève (type d'établissement, secteur, etc.),
 - Les caractéristiques scolaires de chaque élève (formation suivie, options choisies, etc.),
 - Les caractéristiques individuelles de chaque élève (sexe, PCS, âge, etc.),
 - Les coordonnées géographiques de chaque élève (x, y).
- Les données sur les vrais secteurs de collège sont connues pour les départements de la Gironde, du Puy-de-Dôme et de Paris.
 - Des initiatives locales ont permis de les obtenir pour la rentrée 2015.
- Le fait de connaître les secteurs réels dans les trois départements cités est utilisé pour juger la qualité de la reconstitution fictive.

DONNÉES (2/2)

- **Champ : élèves auxquels on souhaite attribuer un collège de secteur**
 - Présents dans le fichier géolocalisé 2015
 - Scolarisés dans un collège (public ou privé)
 - Avec des coordonnées géographiques de bonne qualité
 - Résidant à Paris, en Gironde ou dans le Puy-de-Dôme

Sélection des collégiens à sectoriser

	Gironde	Puy-de-Dôme	Paris	Total
Élèves à sectoriser	68 588	27 735	78 215	174 538
Dont scolarisés dans le public	56 799	22 496	54 331	133 626
Dont scolarisés dans le privé	11 789	5 239	23 884	40 912
% d'élèves scolarisés dans le privé	17,2%	18,9%	30,5%	23,4%

Source : fichier géolocalisé 2015, enrichi du vrai secteur de collège

Champ : Gironde, Puy-de-Dôme et Paris – collèges publics et privés

- **On va donc attribuer un collège fictif de secteur à 174 538 élèves.**

PLAN

MÉTHODE

MÉTHODE : ALGORITHME

■ Intuition

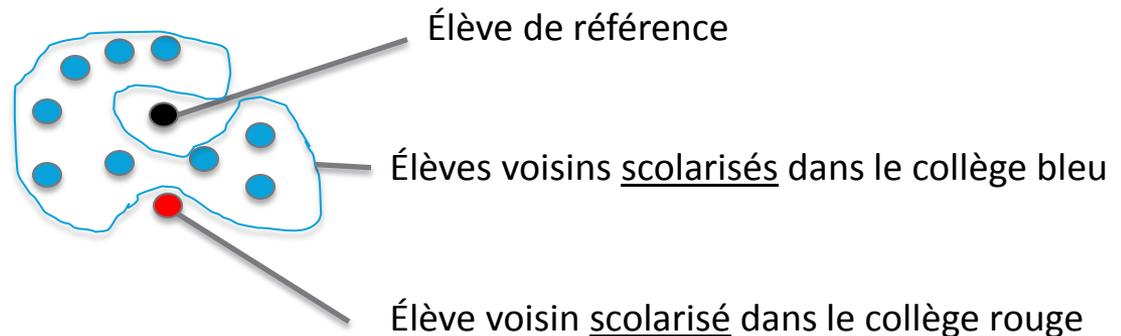
- A priori, un élève est scolarisé dans son secteur quand il fréquente le même collège que ses voisins.

■ Principe général de l'algorithme :

- On choisit un élève à sectoriser (« l'élève de référence »).
- On détermine ses 10 voisins les plus proches (« les élèves voisins »).
- Si, parmi les « élèves voisins », un collège de scolarisation **majoritaire** se dégage, on décide qu'il s'agit du collège de secteur de notre « élève de référence ».

Exemple

Dans ce cas, on dit que l'élève de référence est sectorisé sur le collège bleu



MÉTHODE : ALGORITHME

■ Hypothèses sur les voisins :

■ L'algorithme est mis en œuvre après avoir retiré des voisins

- Les élèves du secteur privé
- Ceux dont la probabilité qu'ils soient hors secteur est élevée
 - scolarisés dans un autre département ou très loin de leur domicile,
 - ayant changé de collège au cours de leur scolarité,
 - inscrits dans des formations atypiques (Segpa*, UPE2A**, classes à horaires aménagés, etc.).

Part de voisins scolarisés hors de leur collège de secteur

Département	Avant sélection des voisins	Après sélection des voisins
Gironde	15,7%	9,7%
Puy-de-Dôme	16,8%	9,5%
Paris	27,1%	15,7%
Total	20,5%	12,0%

Source : fichier géolocalisé 2015, enrichi du vrai secteur de collège

Champ : Gironde, Puy-de-Dôme et Paris – collèges publics

*Sections d'enseignement général et professionnel adapté

**Unités pédagogiques pour élèves allophones arrivants

MÉTHODE : ALGORITHME

■ Hypothèses sur l'algorithme :

- Les distances sont calculées à vol d'oiseau.
- Les 10 premiers voisins sont pris en compte (chiffre apparaissant comme optimal à l'issue de nombreux tests).
- Si le onzième voisin est situé à la même distance de l'élève de référence que le dixième voisin, il est alors inclus dans le calcul (même raisonnement pour le douzième voisin, le treizième, etc.)
- Tout voisin résidant dans la même commune que l'élève de référence est considéré comme plus proche que tout voisin résidant dans une autre commune.
- Quand un collège majoritaire se dégage, la majorité peut être absolue ou relative : soit elle est strictement supérieure à 50% ; soit elle est inférieure ou égale à 50%, les proportions pour les autres collèges de scolarisation étant toutes inférieures.

■ Traitement des cas où aucune majorité ne se dégage

- Si aucune majorité ne se dégage parmi les voisins d'un élève de référence, on attribue à celui-ci le collège majoritaire de l'élève de référence le plus proche de lui.

PLAN

RÉSULTATS

RÉSULTATS : TAUX D'ERREUR

■ DÉFINITION

$$\text{■ Taux d'erreur} = \frac{\text{Nombre d'élèves à qui on a attribué un collège de secteur erroné}}{\text{Nombre d'élèves à qui on a attribué un collège de secteur}}$$

■ RÉSULTATS

- Taux d'erreur global = 6,1%
- Un taux d'erreur variable selon l'ampleur de la majorité qui se dégage

Taux d'erreur selon l'ampleur de la majorité

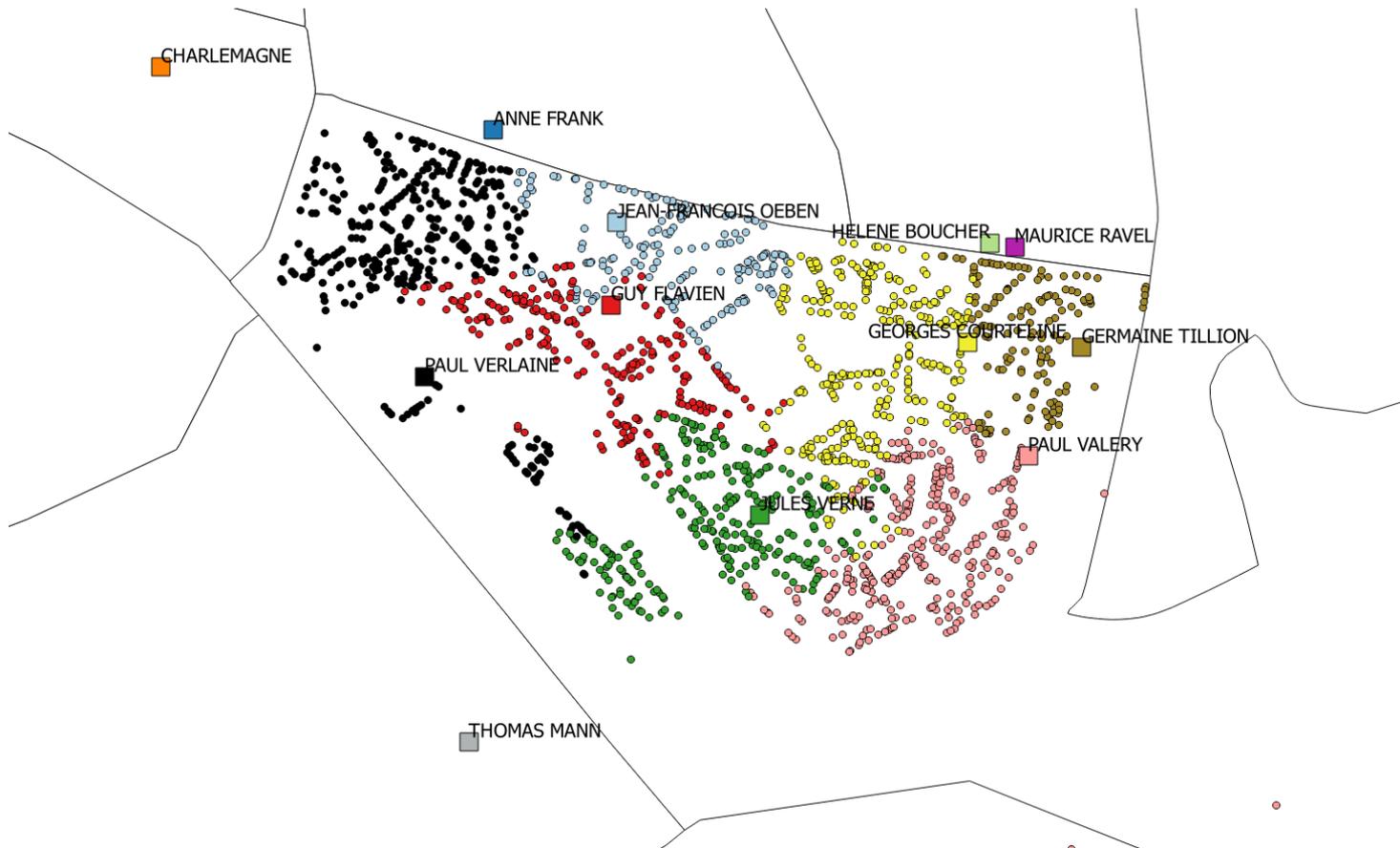
	Majorité absolue	Majorité relative	Absence de majorité – attribution du collège attribué au voisin le plus proche	Total
Erreur	9 034	890	711	10 635
Pas d'erreur	161 751	979	1 173	163 903
Taux d'erreur	5,3%	47,6%	37,7%	6,1%

Source : fichier géolocalisé 2015, enrichi du vrai secteur de collège

Champ : Gironde, Puy-de-Dôme et Paris – collèges publics et privés

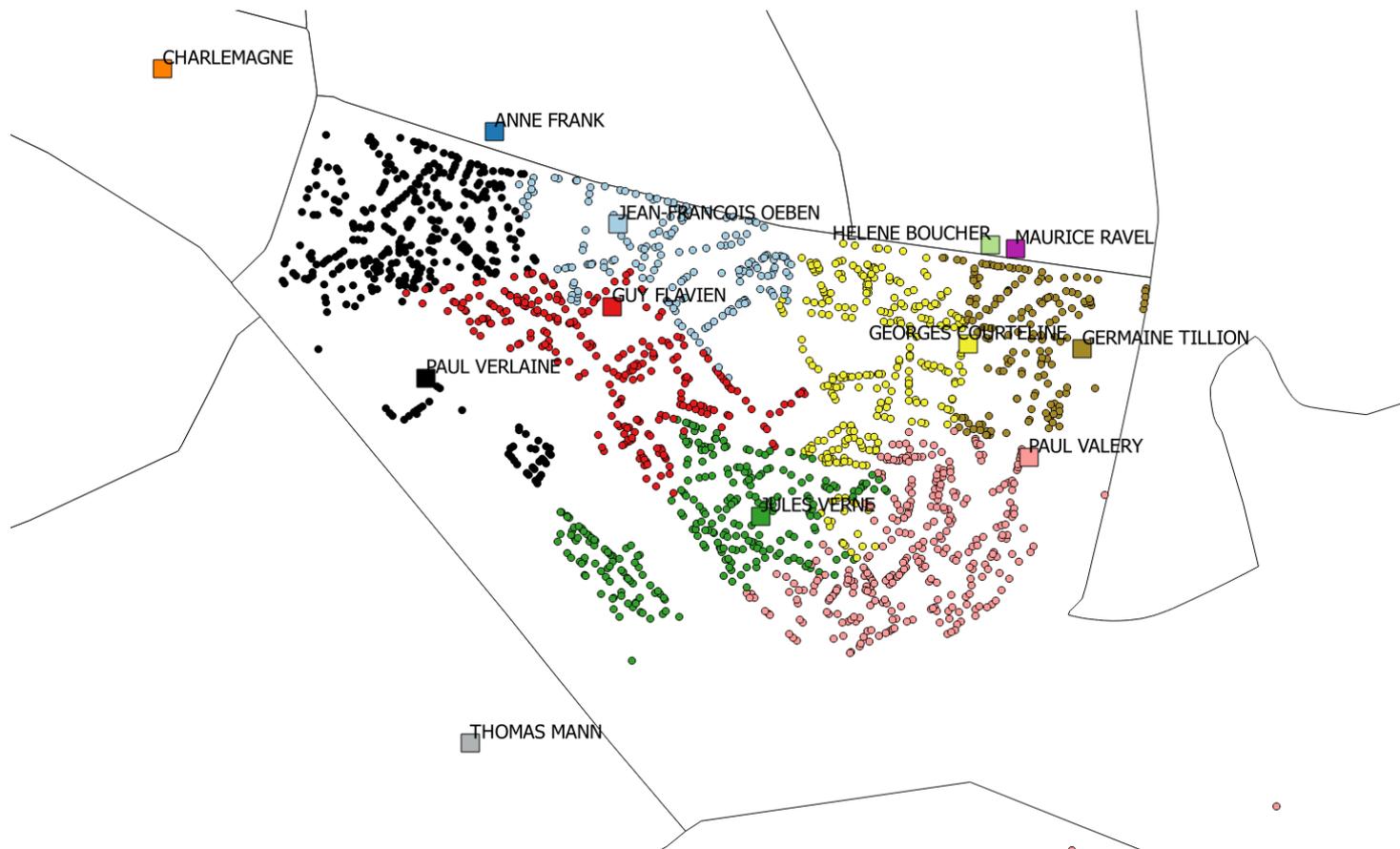
RÉSULTATS : ILLUSTRATION – ELEVES RÉSIDANT DANS PARIS 12^E

■ Secteurs reconstitués



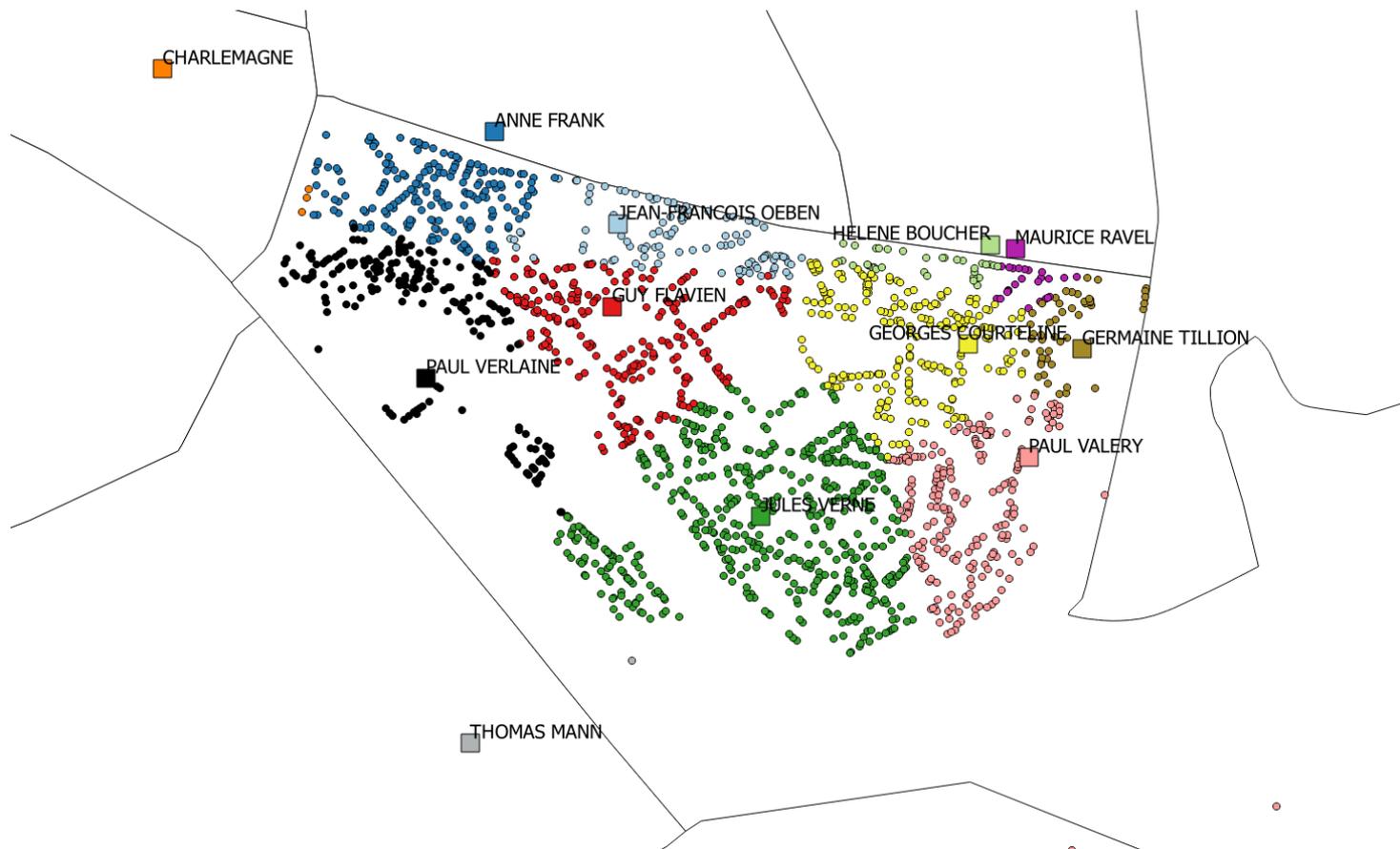
RÉSULTATS : ILLUSTRATION – ELEVES RÉSIDANT DANS PARIS 12^E

■ Secteurs réels



RÉSULTATS : ILLUSTRATION – ELEVES RÉSIDANT DANS PARIS 12^E

- Secteurs reconstitués en considérant que le collège le plus proche est le collège de secteur



RÉSULTATS : ESTIMER LA PART D'ÉLÈVES SCOLARISÉS HORS SECTEUR

CADRE THÉORIQUE

- Les élèves à sectoriser peuvent se classer en deux groupes :
 - ceux dont le collège attribué est quasi certain (groupe 1) : majorité absolue.
 - les autres (groupe 2) : majorité relative ou absence de majorité.
- Dans le cadre d'une estimation statistique de la part d'élèves scolarisés hors secteur, nous supposons que seul le groupe 2 est susceptible de biaiser les résultats.
- Pour contrôler ce biais, nous élaborons une approche en « situations extrêmes » : les élèves du groupe 2 sont
 - soit tous scolarisés dans leur collège de secteur
 - soit tous scolarisés hors secteur.

RÉSULTATS : ESTIMER LA PART D'ÉLÈVES SCOLARISÉS HORS SECTEUR

CADRE THÉORIQUE

■ Soient :

- N_1 le nombre d'élèves du groupe 1
- N_2 le nombre d'élèves du groupe 2
- p_1 la proportion d'élèves scolarisés hors secteur dans le groupe 1
- p_2 la proportion d'élèves scolarisés hors secteur dans le groupe 2

■ La proportion p d'élèves scolarisés hors secteur s'écrit ainsi :

$$p = \frac{(N_1 p_1 + N_2 p_2)}{(N_1 + N_2)}$$

■ Les situations extrêmes sont obtenues dans les deux cas suivants :

- tous les élèves du groupe 2 sont scolarisés dans leur collège de secteur ($p_2 = 0$)
- tous les élèves du groupe 2 sont scolarisés en dehors de leur secteur ($p_2 = 1$)

■ On a ainsi : $p_{min} = \frac{N_1 p_1}{N_1 + N_2}$ et $p_{max} = \frac{N_1 p_1 + N_2}{N_1 + N_2}$

RÉSULTATS : ESTIMER LA PART D'ÉLÈVES SCOLARISÉS HORS SECTEUR

■ PART D'ÉLÈVES SCOLARISÉS HORS SECTEUR PAR DÉPARTEMENT

Département	% d'élèves scolarisés dans un collège public hors secteur	% <u>estimé</u> d'élèves scolarisés dans un collège public hors secteur	
		Minimum	Maximum
Gironde	11,3	10,3	11,1
Puy-de-Dôme	12,2	11,0	12,7
Paris	17,7	16,6	18,4

Source : fichier géolocalisé 2015, enrichi du vrai secteur de collège

Champ : Gironde, Puy-de-Dôme et Paris – collèges publics et privés

■ PART D'ÉLÈVES SCOLARISÉS HORS SECTEUR SELON LE SEXE ET L'APPARTENANCE DE L'ÉLÈVE A UN QUARTIER POLITIQUE DE LA VILLE

	% d'élèves scolarisés dans un collège public hors secteur	% <u>estimé</u> d'élèves scolarisés dans un collège public hors secteur	
		Minimum	Maximum
Sexe			
Garçons	14,9	13,7	15,2
Filles	13,8	12,7	14,2
Appartenance à un QPV			
Oui	21,8	20,3	21,8
Non	13,9	12,8	14,2

Source : fichier géolocalisé 2015, enrichi du vrai secteur de collège

Champ : Gironde, Puy-de-Dôme et Paris – collèges publics et privés

PLAN

PERSPECTIVES

PERSPECTIVES

■ TRAVAUX ULTÉRIEURS

- Estimation de la part d'élèves hors secteur pour d'autres départements et/ou d'autres millésimes (2009, 2011 et 2013).
- Vérification de la donnée « collège de secteur » collectée dans les panels de la DEPP sur déclaration des familles.

■ DIFFUSION DE RÉSULTATS AU NIVEAU ÉTABLISSEMENT

- Plus compliqué à mettre en œuvre : pour un collège, le champ des élèves à considérer est soumis à incertitude.
- Résultats en cours d'expertise (fiabes pour l'estimation des élèves qui recourent au privé).
- Diffusion envisageable au cas par cas.

■ APPROCHE EN « SITUATIONS PROBABLES »

- L'idée consiste à lire dans le collège de leurs voisins l'établissement dont dépendent les élèves.
- En particulier pour traiter les cas où le repérage des situations extrêmes est difficile. Exemple : impact du recours au privé et au public hors secteur sur la ségrégation entre établissements.



MERCI DE VOTRE ATTENTION