

# Le projet OCTOPUSSE de nouvel Echantillon-Maître de l'INSEE.

*Marc CHRISTINE, Sébastien FAIVRE*  
*Insee, DG, Unité Méthodes Statistiques*

JMS 2009

24 mars 2009





***La communication présentée ici est le fruit d'un travail collectif auquel ont participé :***

***Vincent LOONIS, Edouard MAUGENDRE, Bruno BERLEMONT, Fabien GUGGEMOS, Emmanuel GROS (Insee).***



# OCTOPUSSE

**O**rganisation  
**C**oordonnée  
de **T**irages  
**O**ptimisés  
**P**our une **U**tilisation  
**S**tati**S**tique  
des **E**chantillons.



## PLAN de la PRESENTATION :

- I. Introduction : le contexte du nouveau recensement.
- II. Les nouvelles orientations retenues pour les futurs échantillons.
- III. La construction des Zones d'Action Enquêteurs (ZAE) de l'échantillon-maître.
- IV. Allocation et tirage des ZAE.
- V. Tirage des logements dans les ZAE : allocations et pondération.
- VI. Retour sur le tirage des ZAE : qualité, calage et pondérations.
- VII. Conclusion et perspectives.

# I. Introduction : le contexte du nouveau recensement.

- Depuis les années 60, l'Insee gère un système d'échantillonnage de logements ***fondé sur le recensement de la population*** et actualisé avec des logements « neufs » issus des fichiers de permis de construire.
  
- Depuis janvier 2004 : une méthodologie de recensement radicalement différente :
  - ❑ **Petites communes** (moins de 10 000 habitants) :
    - constitution aléatoire de 5 groupes de rotation de ***communes (à probabilités égales)***
    - recensement exhaustif annuel d'un des groupes de rotation.
  
  - ❑ **Grandes communes** (10 000 habitants ou plus) :
    - constitution aléatoire de 5 groupes de rotation ***d'adresses.***
    - tirage ***chaque année*** d'un échantillon de logements au sein d'un groupe donné (environ 8% des logements de la commune).
    - recensement de ces logements.

## II. Les nouvelles orientations retenues pour les futurs échantillons.

- ⇒ Le Nouveau Recensement « en continu » nécessite une redéfinition complète du système d'échantillonnage pour s'adapter au caractère ***rotatif et partiel*** du recensement.
- ⇒ **Un nouveau projet qui a commencé en 2003 et doit s'achever en 2009, pour le tirage des enquêtes à compter du 2ème semestre 2009.**



## II.1 Des constantes dans les choix méthodologiques ou organisationnels.

- Les constantes liées à l'organisation de la **collecte en face à face** subsistent.

=> Un système de type « Echantillon-maître » est toujours utile...

- ***... avec des ZAE (zones d'action enquêteurs) constituant des unités primaires construites une fois pour toutes...***
- ... dont un échantillon sera tiré aléatoirement...
- ... et au sein desquelles seront tirés les échantillons finaux de logements.
- ***Ceci assure une concentration géographique des échantillons et une limitation des coûts.***



## II.2 La nouveauté principale.

### **Profiter de la « fraîcheur » apportée par le RP :**

- ⇒ Utiliser comme base de sondage standard pour les enquêtes de la seconde moitié de l'année  $n + 1$  les logements recensés en  $n$ .
- ⇒ On tire donc chaque année dans une fraction du territoire (problème de base de sondage incomplète).

### **=> Avantages :**

Limiter les déchets dus aux transformations de logements, sources de déplacements inutiles des enquêteurs (fatigue, coûts...).

Sélectionner de manière plus efficace des échantillons sur population ciblée (de caractéristiques récentes connues).

Abandonner la Base de Sondage Logements Neufs :

- mécanique complexe et lourde
- basée sur une source administrative sur laquelle on n'a aucune prise
- trop coûteuse.



### III. La construction des ZAE de l'échantillon-maître.

***Défi : concilier principe d'interrogation dans le RP le plus récent avec constitution d'unités primaires fixes (ou ZAE : Zones d'Action Enquêteurs).***

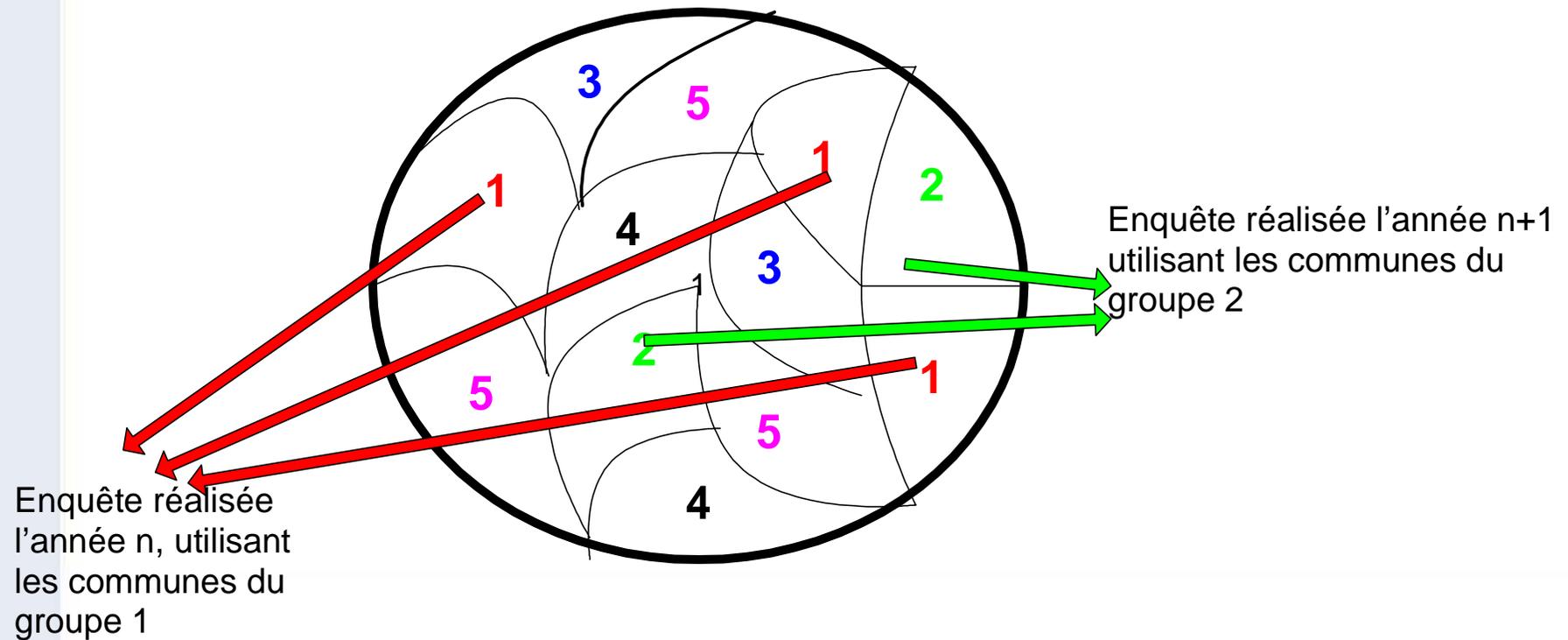
#### III. 1 Contraintes et objectifs.

⇒ Construire des unités primaires *au sein de chaque région, réalisant une partition du territoire :*

- ... comportant des communes des 5 groupes de rotation du RP...
- ... avec un nombre minimal de logements (300) dans chacune d'elles.



## Schéma idéal-type :





## III.2 La construction effective.

### A) Les grandes communes.

- **Une Grande commune constituera une ZAE à elle toute seule (les 5 groupes de rotation présents).**

### B) Les petites communes.

- Il s'agit de construire une agrégation optimale des communes :
  - Sous les contraintes de taille minimale (nombre de logements par groupe) et de respect des limites régionales.
  - Sous l'objectif de chercher à réduire l'étendue géographique des ZAE pour limiter les déplacements des enquêteurs et **garantir ainsi la qualité des enquêtes...**

=> Construction automatisée des ZAE à l'aide d'**algorithmes innovants conçus spécifiquement.**



## Formalisation mathématique :

Une ZAE *potentielle* construite autour d'une commune-pivot  $C^P$  est l'ensemble des communes  $C_i$

$$\text{solution du problème : } \underset{C_j}{\text{Min}} \sum_{j=1}^C \|C_j - C^P\|,$$

$$\text{sous les contraintes : } \forall g \in \{1, \dots, 5\} : \sum_{i=1}^C G_{i,g} N_i^C \geq 300$$

$N_i^C$  = nombre de logements principaux de la commune  $C_i$  .

$G_{i,g} = 1 \Leftrightarrow$  la commune  $C_i$  est affectée au groupe de rotation  $g$ ,



## Principe de l'algorithme de constitution des ZAE formées de petites communes

- › Algorithme itératif, démarrant avec la plus grande (en nombre de résidences principales) des petites communes de la région.
- › Constitution de la ZAE autour d'une **commune pivot** constituant le « centre » de la ZAE.
- › Une ZAE est constituée si parmi les communes de la région (non encore affectées) ***situées à une distance de la commune pivot inférieure à la distance maximale spécifiée*** se trouvent les communes nécessaires pour atteindre 300 résidences principales par groupe de rotation. Sinon la ZAE n'est pas constituée.
- › A chaque étape, la plus grande commune disponible non encore affectée est testée comme commune pivot potentielle d'une ZAE.

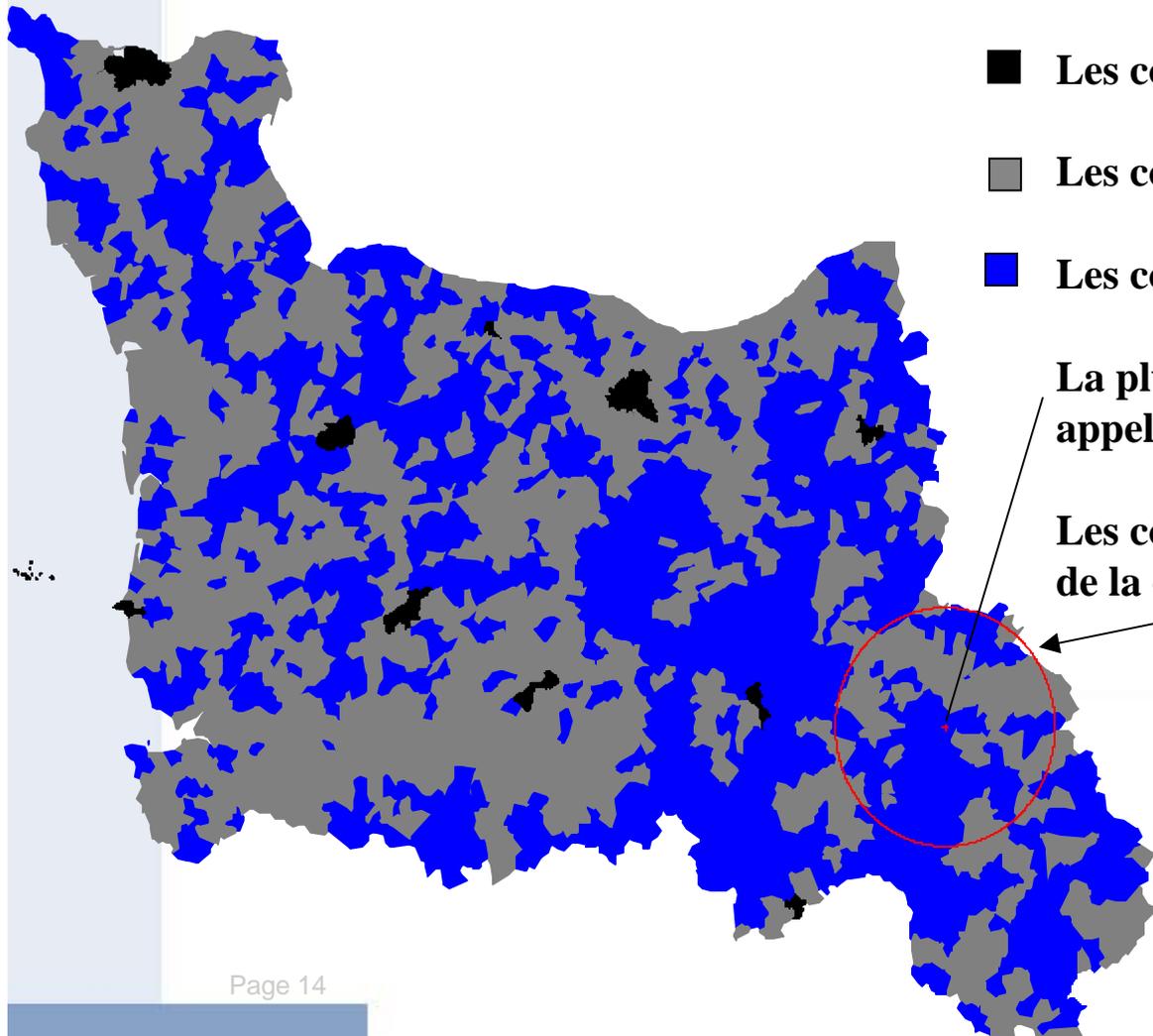
## Un principe de construction automatique des ZAE (cartographie réalisée par Vincent Loonis)

A une étape donnée,  
On distingue :

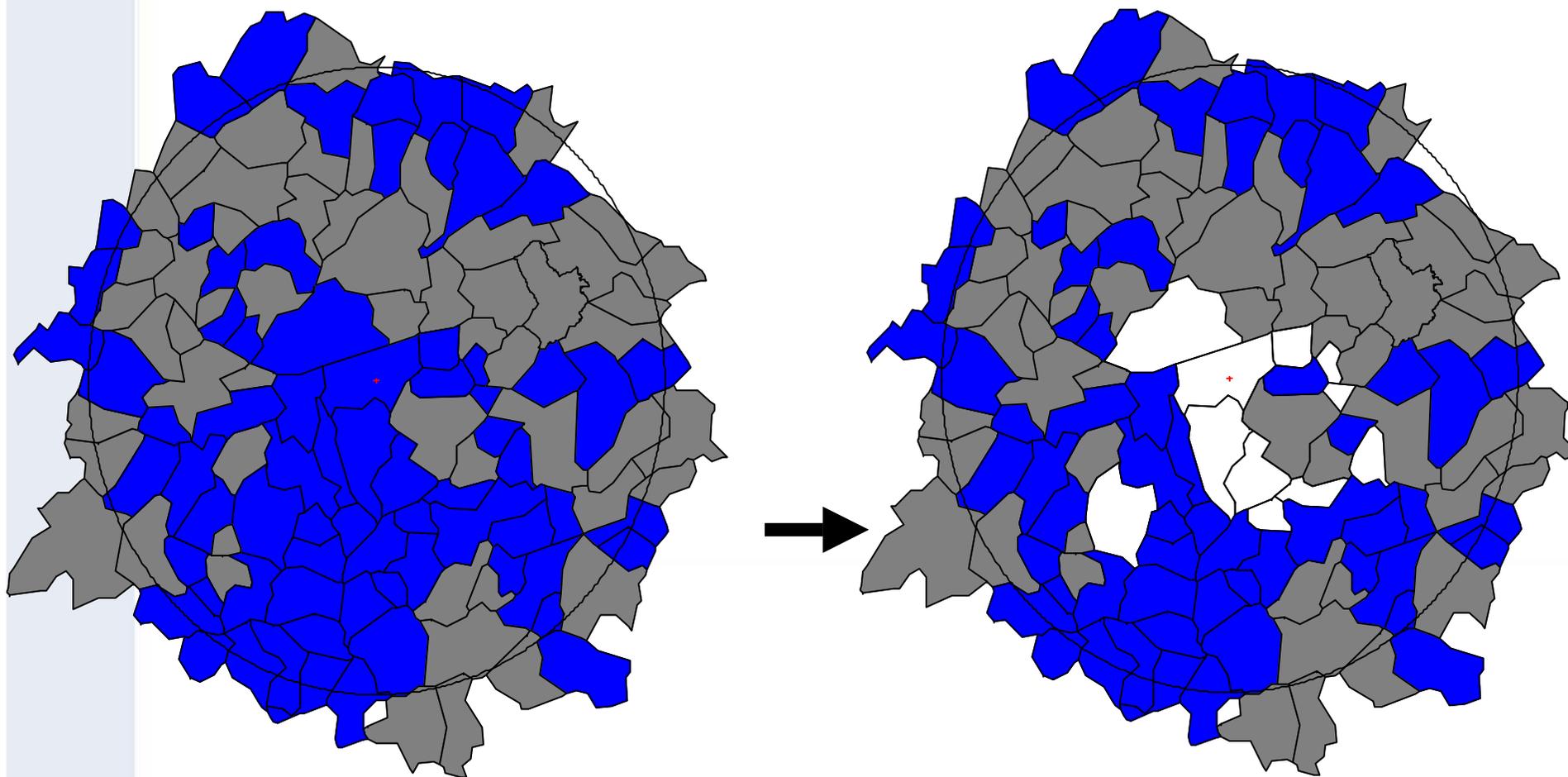
- Les communes de + de 10 000 habitants,
- Les communes déjà affectées,
- Les communes non encore affectées,

La plus grande des communes non affectées,  
appelée commune pivot de la ZAE.

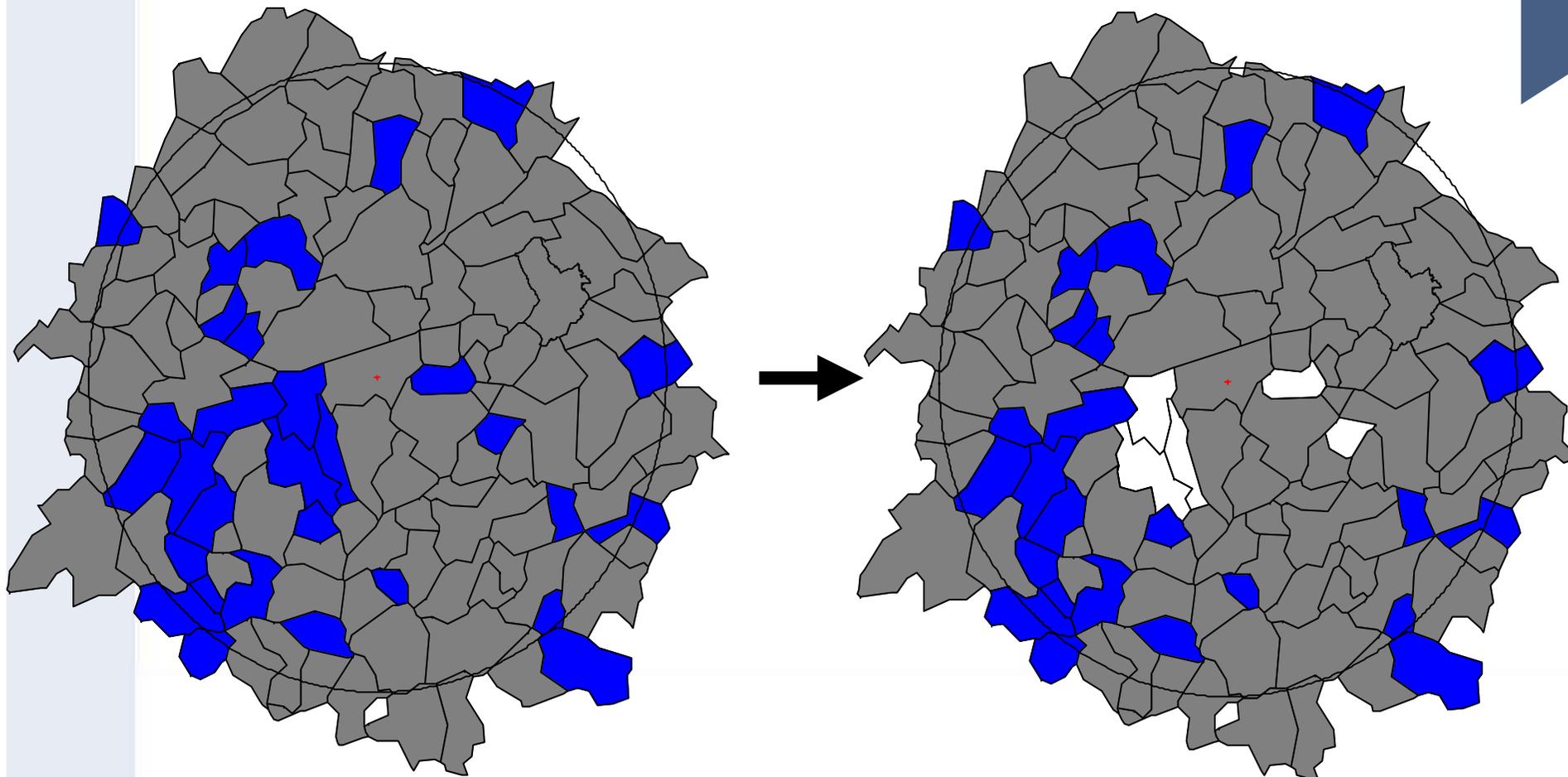
Les communes dans un rayon de 20 Km autour  
de la commune pivot.



Parmi les communes autour de Sainte-Gauburge, on crée une ZAE avec les communes les plus proches non affectées ayant 300 résidences principales pour chaque vague de recensement.

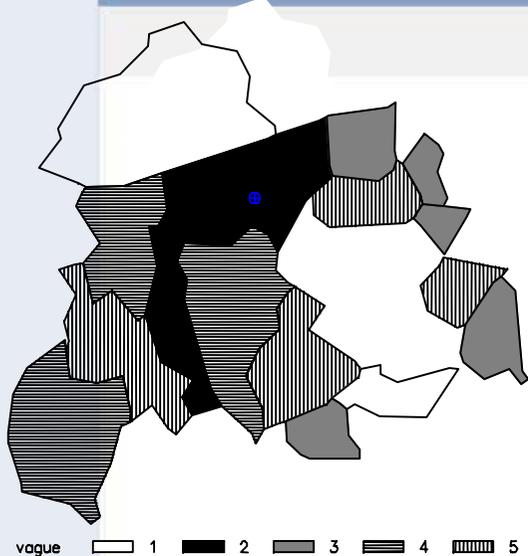


On affecte à la ZAE de Sainte Gauburge, les communes non affectées les plus proches (dans la Limite de 20 Km). C'est la phase d'atterrissage.

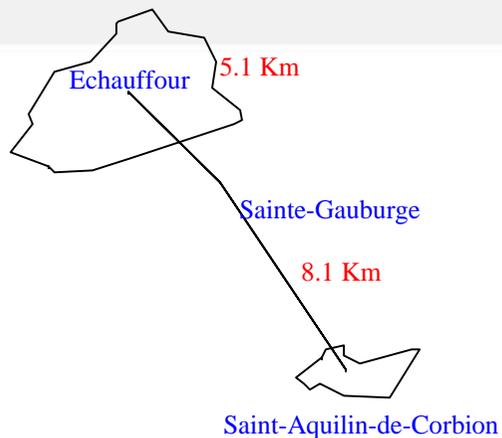


A la fin il reste toujours quelques communes non affectées parce que trop éloignées d'une commune pivot ou dans un environnement très peu dense...

**ZAE de Sainte Gauburge**  
Communes enquêtées sur 2009-2013



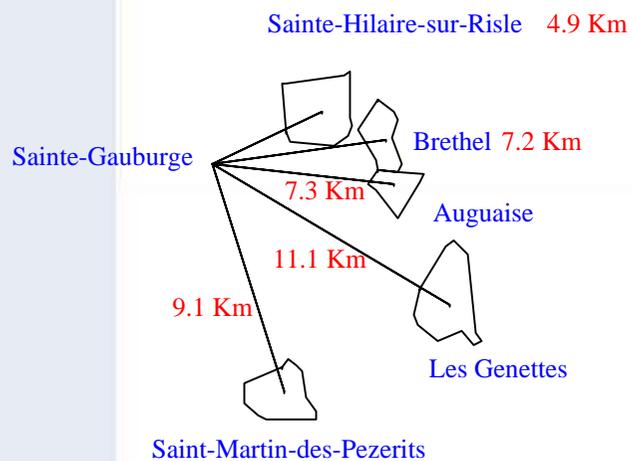
**Communes enquêtées en 2009**



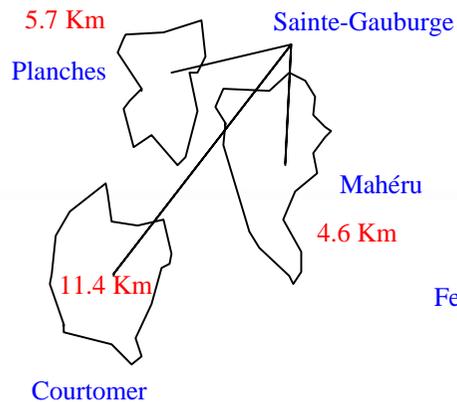
**Communes enquêtées en 2010**



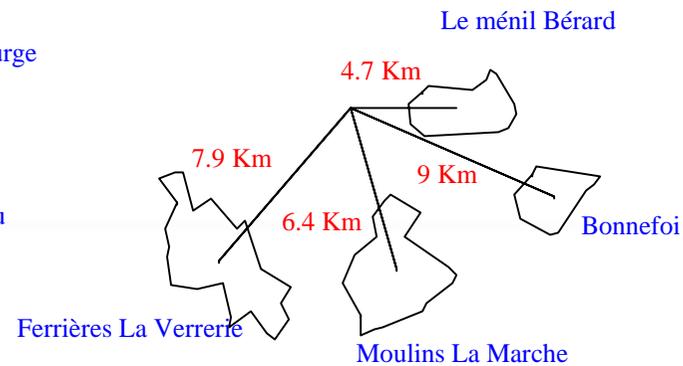
**Communes enquêtées en 2011**



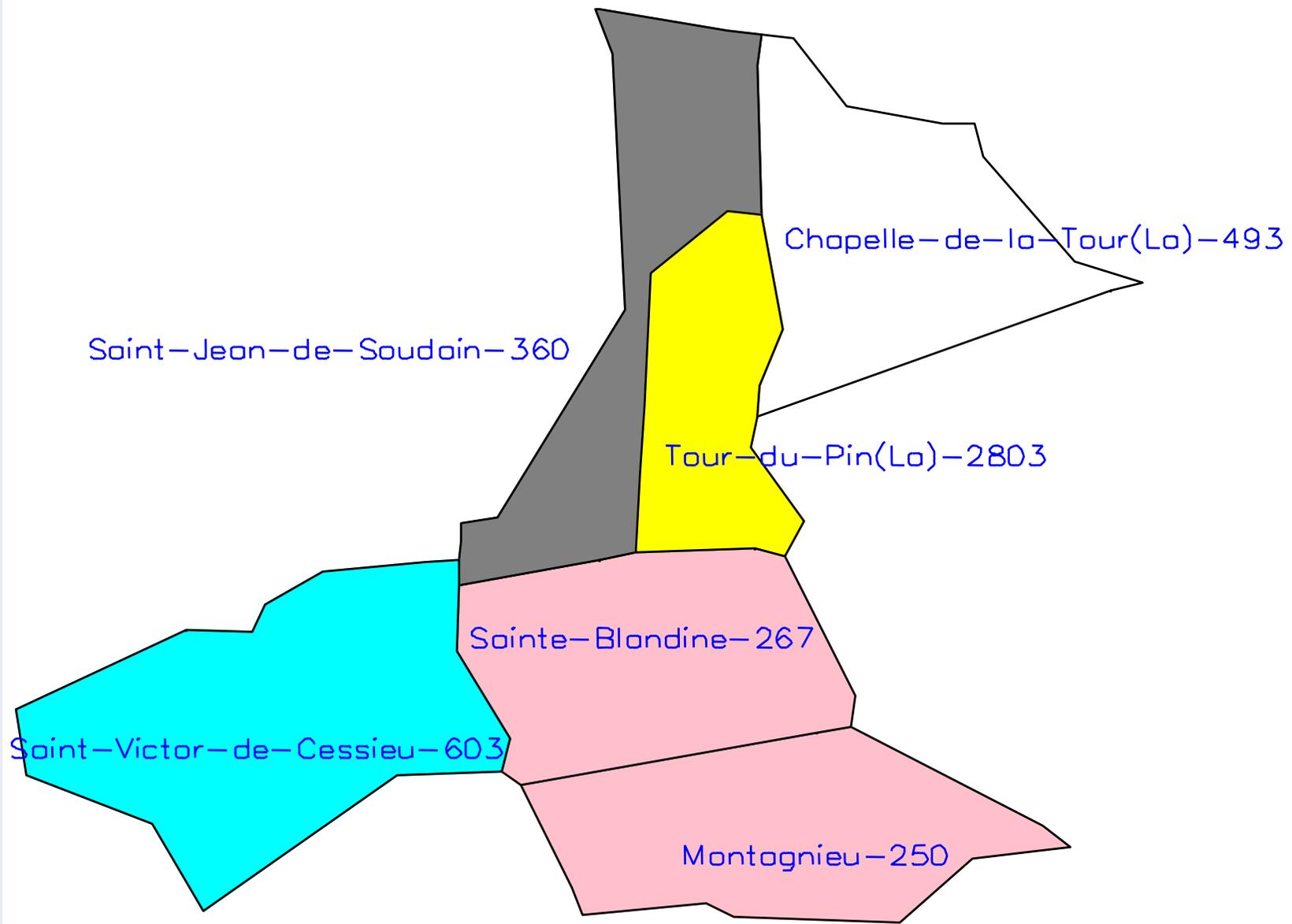
**Communes enquêtées en 2012**



**Communes enquêtées en 2013**



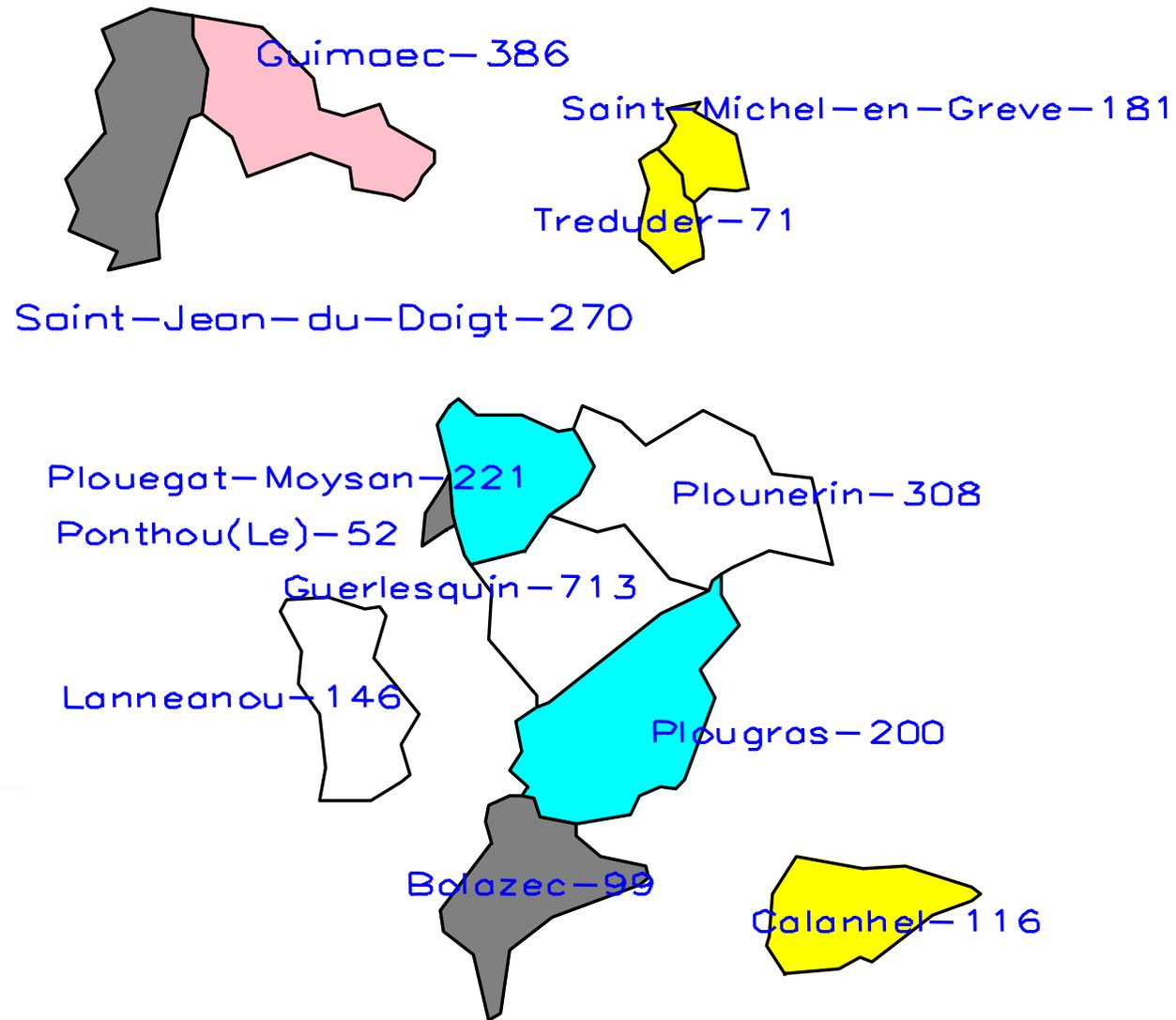
# ZAE de la Tour du Pin (Z38509) Isère



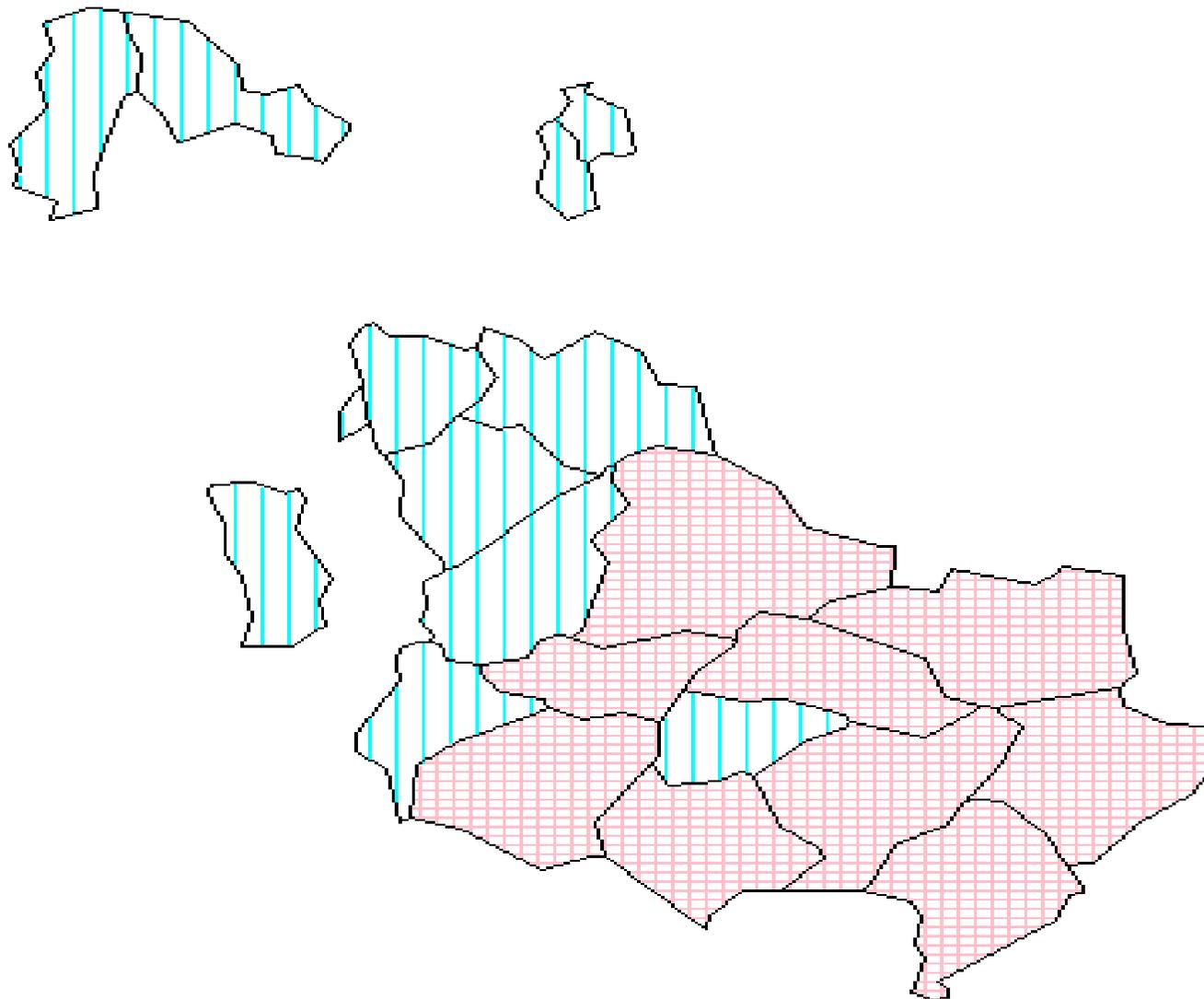
vague 1 2 3 4 5



# ZAE de Plounérin (Z22227) Bretagne



# Imbrication des ZAE de Plounerin (bleu) et de Callac



P

**IdZae**



**Z22025**

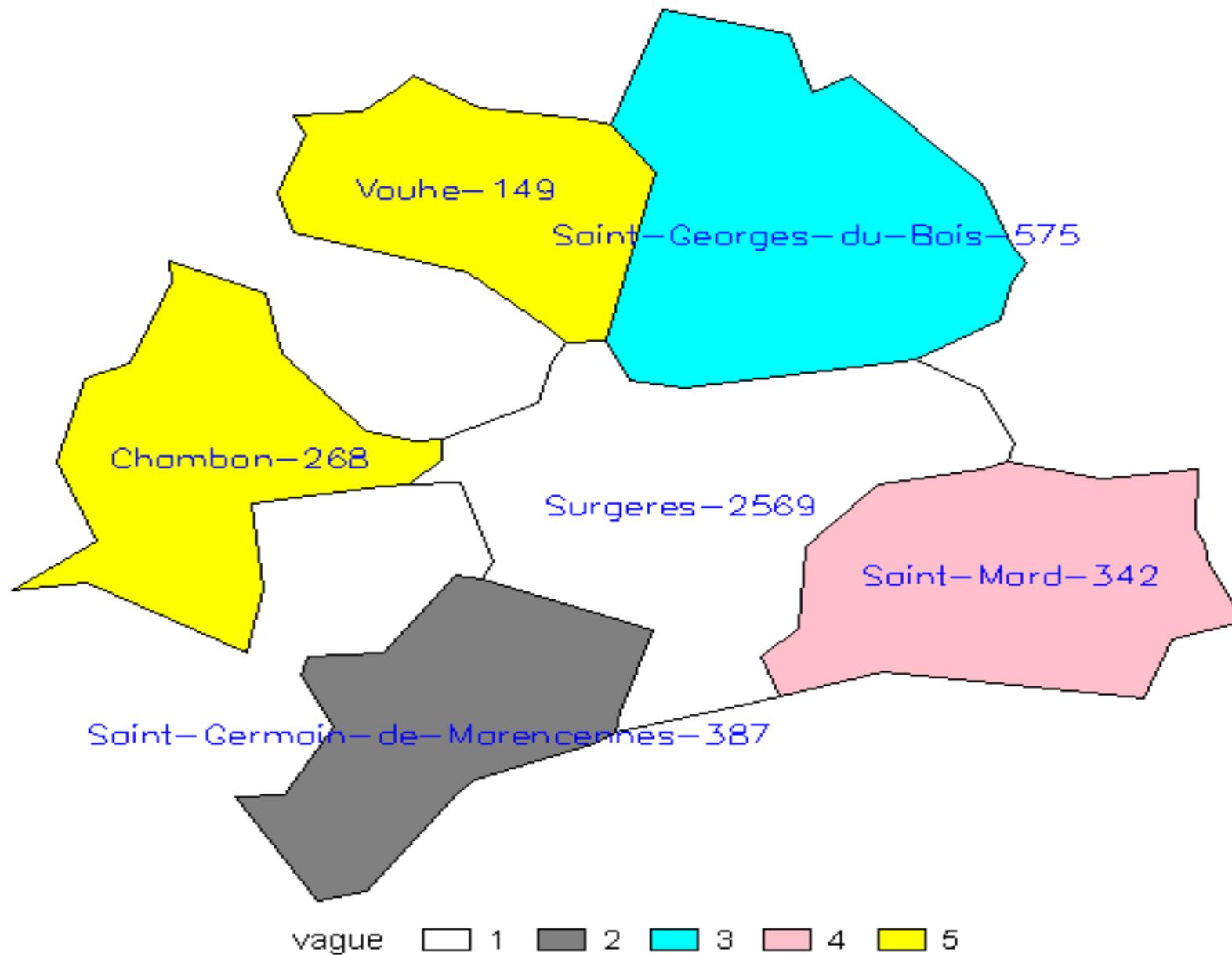


**Z22227**



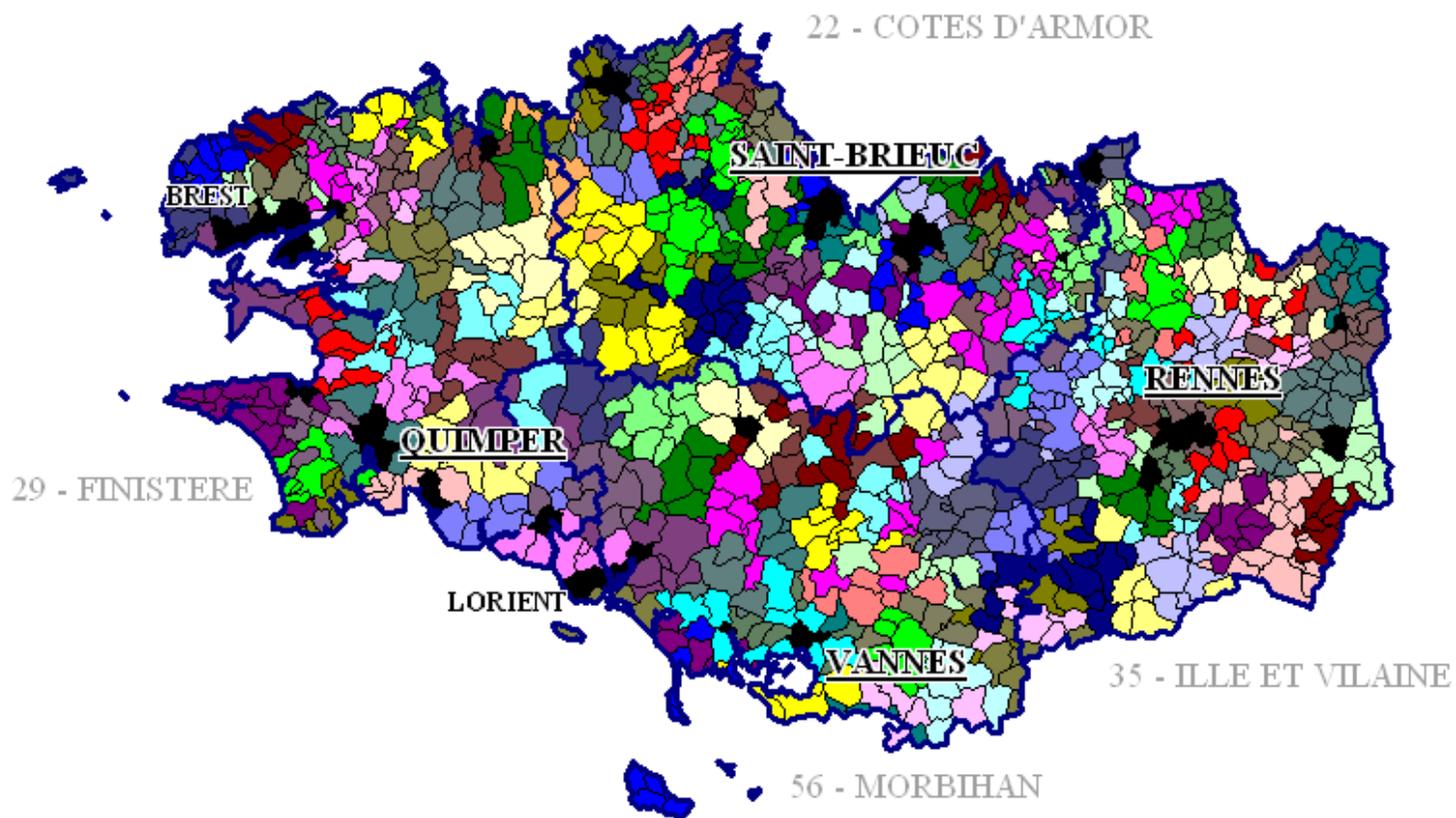
JMS 2009 – 24 mars 2009

# ZAE hétérogène en tailles : ZAE Z17434





## ZAE - REGION 53 - BRETAGNE

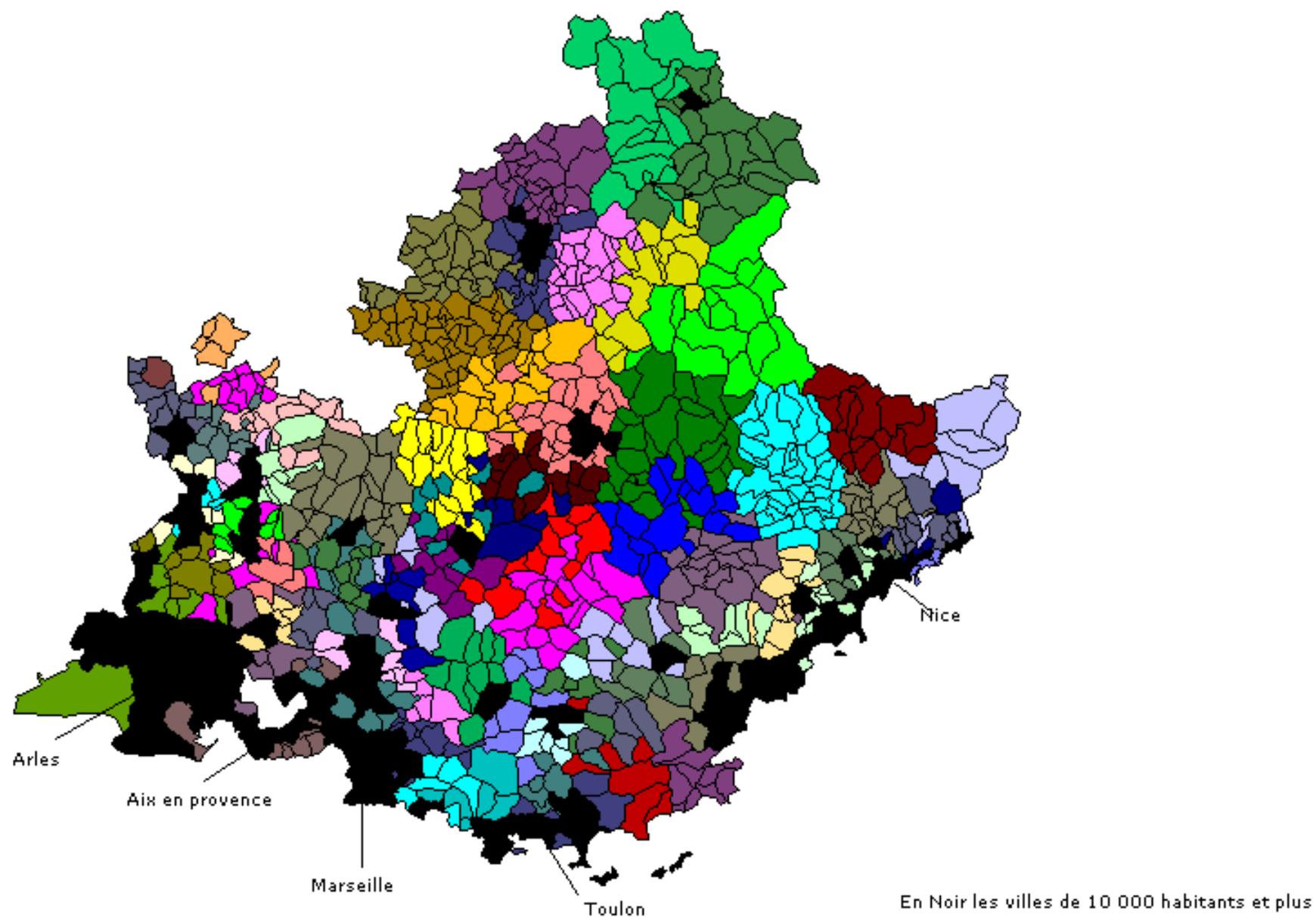


En noir, les villes de 10 000 habitants et plus

Source : INSEE



## LES ZAE EN REGION PACA



En Noir les villes de 10 000 habitants et plus

Source : Insee



### III.3 Le résultat de la construction des ZAE.

- 2893 ZAE Petites Communes.
- 892 ZAE Grandes Communes (y compris arrondissements PLM)...
- ... soit 850 communes stricto sensu.
  
- TOTAL = **3743 ZAE.**
  
- L'algorithme de construction est déterministe mais l'affectation initiale des communes en groupes de rotation est aléatoire

***=> les ZAE sont des « objets aléatoires ».***

## IV. Allocation et tirage des ZAE.

### IV. 1 Calcul de l'allocation.

#### Hypothèses de base :

- Des ZAE qui seront tirées proportionnellement à leur taille (nombre de logements principaux)
- Dont certaines retenues d'office (« exhaustives »).

#### Paramètres choisis :

- Pour une taille standard d'échantillons au taux de sondage  **$TAU = 1/2000$**  (un peu moins de 12.000 logements principaux).
- Affectation moyenne de  **$e = 20$  Fiches-adresses** par ZAE non exhaustive (**1 ZAE = 1 enquêteur**).



Taille de l'échantillon :  $n = \tau N = \tau N^{exh} + k e$

Taille de la strate exhaustive :  $N^{exh} = \sum_{i / N_i \geq S} N_i$

Nombre de ZAE à tirer :  $k = \frac{\tau(N - N^{exh})}{e}$

Probabilité de tirage d'une ZAE :  $\pi_i = k \frac{N_i}{N - N^{exh}} = \frac{\tau N_i}{e}$

Seuil d'exhaustivité :  $S = \frac{e}{\tau}$



## Résultats :

- Seuil d'exhaustivité de 40.000 logements principaux.
  
- ⇒ **37** grandes communes exhaustives affectées à plusieurs enquêteurs.
  
- ⇒ **488** ZAE non exhaustives tirées dont :
  - 286 ZAE-PC
  - 202 ZAE-GC non exhaustives.



## IV. 2. Tirage des ZAE.

- Stratifié *par région* (cas particulier Ile de France : Séparation petite / grande couronne).
- Equilibré sur des *totaux régionaux* :
  - Il est nécessaire d'équilibrer non seulement au niveau de l'ensemble de la ZAE ***mais aussi de chacun des 5 groupes de rotation...***
  - ... de manière à avoir chaque année une base de sondage « représentative ».
  - Cela multiplie le nombre de contraintes d'équilibrage et réduit d'autant le nombre de variables indépendantes à introduire.



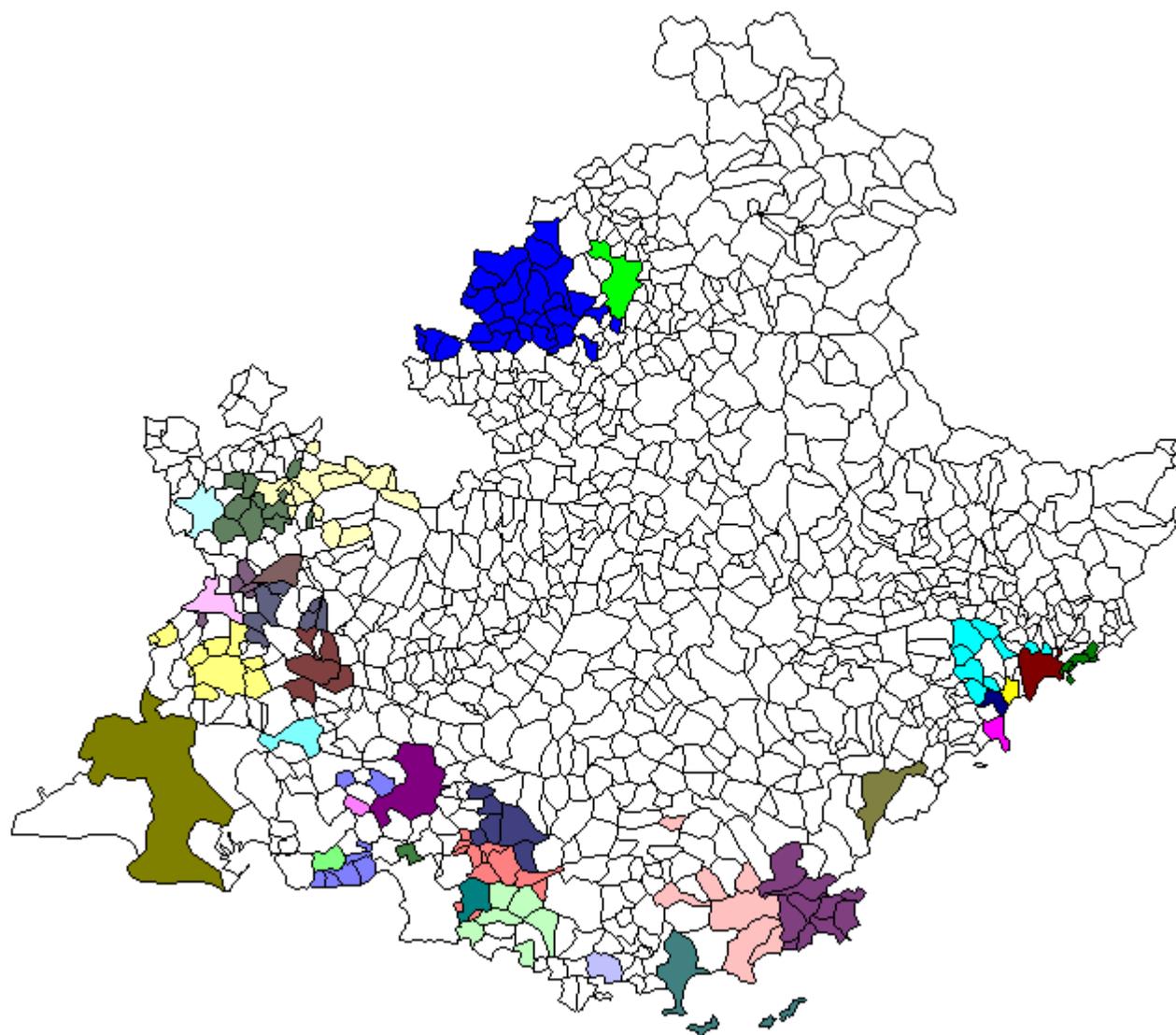
De plus...

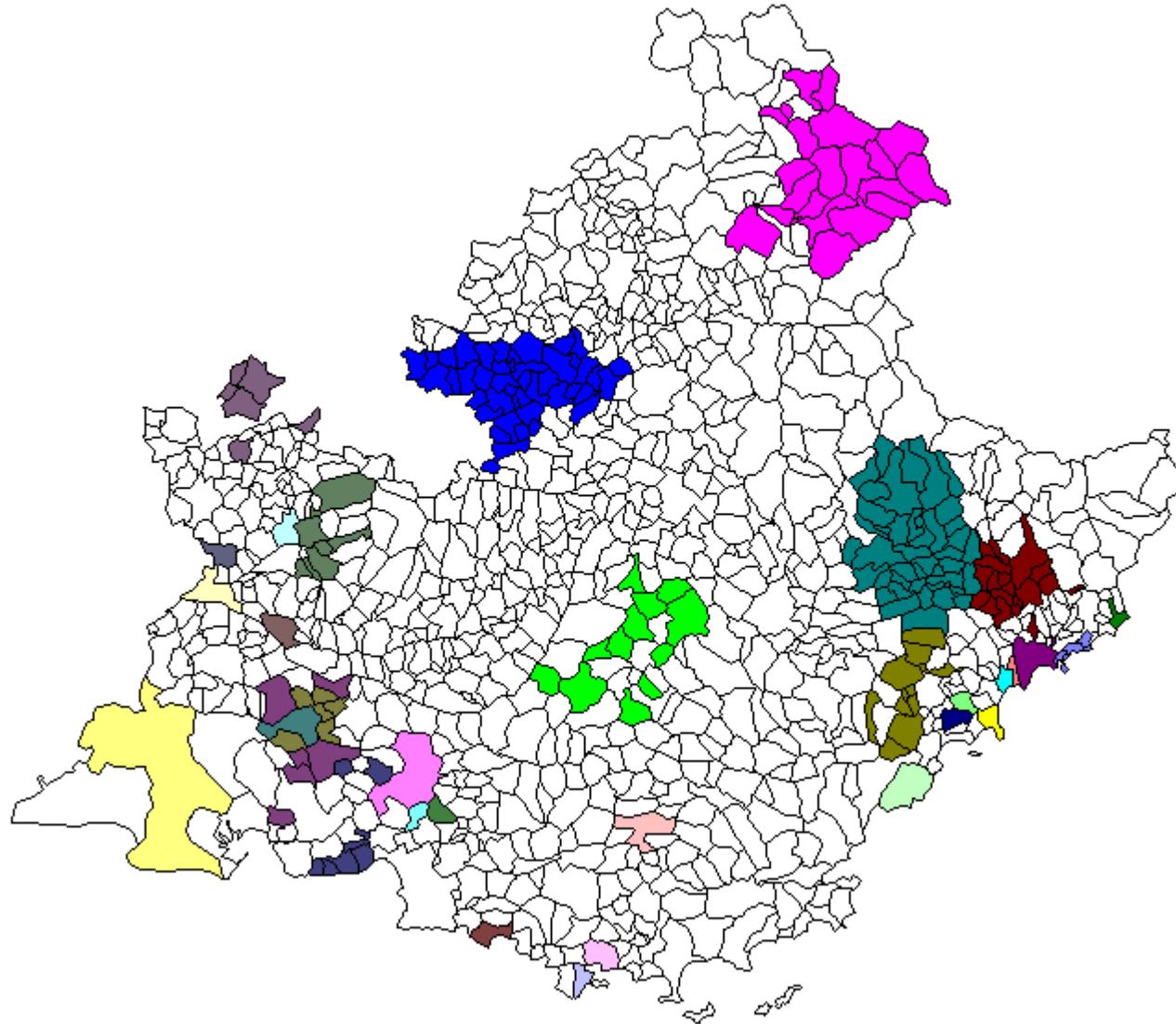
- La base de sondage annuelle est équilibrée sur le total des communes du groupe de rotation impacté....
  - ... mais pas sur le total France métropolitaine (groupes de rotation pas équivalents).
- ⇒ Source de biais pour les estimations (utilisation, pour une année donnée, d'une base de sondage « incomplète »).



## Variables d'équilibrage utilisées.

- **Nombre de résidences principales** des communes de la ZAE pour chacun des cinq groupes de rotation.
- **Revenu fiscal total** des communes de la ZAE pour chacun des cinq groupes de rotation.
- Nombre total de **logements de la ZAE en zone péri-urbaine**, en **zone rurale** et, par complément, en **zone urbaine (ensemble de la ZAE)**.
- Des variables d'équilibrage supplémentaires en **Ile de France** (âge, type de ménage, caractéristiques de l'habitat : collectif, HLM, statut d'occupation, nombre d'étrangers...).







### IV.3 Le problème du tirage simultané des ZAE EM et EMEX.

- Plusieurs jeux de ZAE à tirer.
- L'EMEX regroupe les ZAE tirées pour l'EM et les ZAE additionnelles tirées pour les **extensions régionales**.
- Tirage simultané EM – EMEX dans chaque région.
- Un EMEX standard sera mobilisé pour des extensions de taille modérée (on a retenu au total le **double** de l'effectif des ZAE–EM).
- Si nécessaire, pour les grandes extensions : EMEX élargi.

Page 33



## V. Tirage des logements dans les ZAE : allocations et pondération.

Au sein de chaque ZAE tirée, les unités secondaires (logements) sont tirées par sondage aléatoire simple, dans le groupe de rotation à impacter.

### **MAIS :**

1ère difficulté en petites communes :

- La constitution des ZAE est dépendante de l'affectation aléatoire des communes en groupes de rotation (phase RP) => **c'est une partition aléatoire du territoire.**
- Les probabilités d'inclusion des communes dans l'échantillon des ZAE sont **conditionnelles** à cette affectation.
- La probabilité de tirage d'un logement dans OCTOPUSSE (via le tirage de la ZAE à laquelle il appartient) n'a de sens que conditionnellement au tirage initial des groupes de rotation.



## 2ème difficulté en grandes communes :

- 1ère phase RP : affectation inégale et « semi-déterministe » des adresses dans les différents groupes de rotation (notamment les grandes adresses) (et non à probabilités égales 1/5)  
=> nécessité de reconstitution d'une « pseudo-probabilité » d'affectation.
  - 2ème phase RP : surreprésentation des « adresses neuves » et des « grandes adresses » dans la phase de sélection des adresses à recenser par le RP (recensées exhaustivement dans chaque groupe de rotation d'adresses).
- => nécessité d'un ***rééchantillonnage*** des logements au sein de la base de sondage pour disposer d'une base effective de logements à poids identiques.

## Répartition des logements en grandes adresses par groupe de rotation (RIL2006)

	Ensemble des GC	GC ayant au plus 49 GA	GC ayant au plus 9 GA	GC ayant au plus 4 GA	GC ayant une GA
Nombre Communes	775	671	375	227	89
<b>Log GR1</b>	<b>203 595</b>	<b>115 072</b>	<b>26 530</b>	<b>9 747</b>	<b>2 925</b>
Log GR 2	195 835	110 640	26 070	9 287	1 625
Log GR 3	195 245	110 356	27 073	11 052	1 914
Log GR 4	183 420	102 500	20 237	5 776	437
<b>Log GR 5</b>	<b>184 694</b>	<b>98 821</b>	<b>16 982</b>	<b>3 239</b>	<b>68</b>
TOTAL Log	962 789	537 389	116 892	39 101	6 969



## Principe du rééchantillonnage en grandes communes

- › Objectif: avoir une base de logements dans la ZAEGC « à probabilités égales », sans surreprésentation de certaines catégories d'adresses, pour enchaîner des tirages à probabilités égales.
- › Pour une strate donnée: « Probabilité de première phase » (probabilité d'appartenance d'un logement de la strate au groupe de rotation recensé) estimée empiriquement à partir du RIL comme la part des logements de la commune appartenant au groupe de rotation recensé.



- › Probabilité de première phase calculée dans chaque grande commune :
- › -Pour les logements en grandes adresses:

$$P1P\_GA = \frac{Nb \log\_GA\_GRi}{Nb \log\_GA}$$

-Pour les logements en petites adresses:

$$P1P\_PA = \frac{Nb \log\_PA\_GRi}{Nb \log\_PA}$$



- › Le regroupement des autres adresse et des adresses neuves vient du fait la définition des adresses neuves n'est pas homogène suivant les groupes:
  - dans le groupe qui vient d'être recensé, les logements en adresses neuves correspondent aux logements du groupe construits il y a moins d'un an
  - alors que dans le groupe qui va être recensé ils correspondent aux logements construits il y a moins de cinq ans.



- › On estime ensuite pour chaque strate la probabilité que le logement soit recensé comme le produit de la probabilité de première estimée par la probabilité de seconde phase fournie par le Recensement (valant 1 pour les logements en grandes adresses et en adresses neuves).
- › Sauf cas exceptionnels, les logements en autres adresses sont ceux qui ont la probabilité de recensement la plus faible.



- › On garde alors pour constituer la base de sondage l'ensemble des logements en autres adresses et une fraction des logements en grandes adresses et en adresses neuves, de sorte à ce que tous les logements aient la même probabilité de présence finale dans OCTOPUSSE (produit de la probabilité de recensement par le taux de rééchantillonnage).
- › En pratique, les logement conservés en grandes adresses et en adresses neuves sont sélectionnés par tirage systématique à probabilités égales.



## Exemple théorique

- › Exemple théorique d'une commune où les groupes de rotation sont équilibrés et où 1/3 des logements en autres adresses du groupe de rotation sont recensés:

Strate	« Probabilité » de 1ère phase	Probabilité de 2nde phase	Part des logements recensé	Taux de rééchantillonnage
GA	0,2	1	20%	1 logement sur 3
AN	0,2	1	20%	1 logement sur 3
AA		0,3333	6,66%	Tous les logements conservés

# En résumé : schéma général de tirage.

	Petites communes	Grandes communes
1 <sup>ère</sup> phase : échantillonnage RP	Constitution aléatoire de 5 groupes de rotation équiprobables de <b>communes</b> .	Constitution aléatoire de 5 groupes de rotation d' <b>adresses</b> dans chaque commune.
		Echantillon annuel d'adresses formant des grappes de logements (taux moyen = 8 % des logements de chaque commune).
		Les grandes et les nouvelles adresses forment une strate exhaustive.
Construction des Unités Primaires (ZAE)	Stratification géographique (région)	
	Construction des UP par agrégation de communes, <b>conditionnellement au résultat de la 1<sup>ère</sup> phase RP.</b>	
		Chaque grande commune forme une UP à elle toute seule.
2 <sup>ème</sup> phase : échantillonnage des logements.		Seuils d'exhaustivité des grandes communes.
	1 <sup>er</sup> degré : tirage équilibré des UP dans chaque région (trage en 2 phases EM-EMEX). <b>UP fixes pour 5 ou 10ans.</b>	
	Base de sondage annuelle = intersection des UP tirées et des logements appartenant au groupe de rotation de l'année.	
		Rééchantillonnage des logements appartenant aux grandes adresses, pour assurer des probabilités égales pour tout logement appartenant à la base de sondage.
	2 <sup>nd</sup> degré : ( <b>pour chaque enquête</b> ) tirage des logements dans chaque UP sélectionnée au 1 <sup>er</sup> degré	

## V.1 Estimations.

« *Estimateur en expansion* »

(analogue à HORWITZ-THOMSON) :

$$\hat{T}_e(Y) = 5 \sum_{k \in \text{ZAEPC}} \frac{N_{k,t}}{\pi_k} \bar{y}_{k,t} + \sum_{k \in \text{ZAEGC}} \frac{\hat{N}_k}{\pi_k} \bar{y}_{k,t}$$

$N_k$  = nombre de logements dans la ZAE  $k$

$N_{k,t}$  = nombre de logements dans le groupe  $t$  de la ZAE  $k$

$\pi_k$  = probabilité d'inclusion de la ZAE  $k$

$\bar{y}_{k,t}$  = moyenne empirique d'une variable d'intérêt sur l'échantillon de logements tiré dans le groupe  $t$  de la ZAE  $k$



## Pondérations logements résultantes :

➤ Pour un logement tiré au 2<sup>nd</sup> degré dans une commune du groupe de rotation  $t$  de la ZAE  $k$  :

• Dans une ZAEPC :

$$\omega_l = 5 \frac{1}{\pi_k} \frac{N_{k,t}}{n_{k,t}}$$

• Dans une ZAEGC :

$$\omega_l = \frac{1}{\pi_k} \frac{\hat{N}_k}{n_{k,t}}$$

$n_{k,t}$  = taille de l'échantillon de logements  
dans le groupe  $t$  de la ZAE  $k$ .



## V.2 Calcul des allocations.

- Dans les ZAEGC exhaustives, l'allocation calculée est proportionnelle à la taille de la ZAEGC en nombre de résidences principales
- Dans les ZAE non exhaustives, l'équiallocation est un objectif (charges de travail uniformes par enquêteur).
- Si elle ne peut être satisfaite rigoureusement, les calculs d'allocation cherchent à **minimiser la dispersion des pondérations finales logements**, sous des contraintes :
  - de taille totale d'échantillon fixée
  - de charge minimale et maximale par enquêteur (c'est-à-dire par ZAE non exhaustive) : dans la pratique, fourchette 20-40.



## VI. Retour sur le tirage des ZAE : qualité, calage et pondérations.

- On a cherché à analyser la qualité du tirage des ZAE en comparant :
  - l'estimation (à partir de l'échantillon de ZAE) du total « France entière » de différentes variables auxiliaires (obtenue à partir des totaux observés sur les ZAE tirées, supposés connus, i.e. avant tirage du second degré)
  - avec le vrai total France entière (connu par des sources exhaustives, RP 99).



➤ On constate des erreurs relatives

- plus ou moins importantes
- variables d'un groupe de rotation à l'autre,
- quel que soit le type d'estimateur retenu,
- notamment pour :

la segmentation par type d'espace (rural / périurbain / urbain)  
la répartition du nombre de personnes employées par secteur

(variables imparfaitement ou pas prises en compte dans l'équilibrage lors du tirage).

=> pénalisantes en particulier pour les enquêtes annuelles, ***mesurant des évolutions.***



## Erreur relative des bases de sondage annuelles sur la variable « nombre de résidences principales dans l'espace rural ».

Groupe de rotation	Erreur relative (estimateur en expansion)
GR 1	+3,4%
GR 2	-3,3%
GR 3	-7,9%
GR 4	-8,1%
GR 5	-9,4%



## Une solution pour remédier à ce problème : le calage des ZAE.

- Effectuer *chaque année* un (re)calage des ZAE pour obtenir une base de sondage annuelle « représentative ».

### Principe du calage des ZAE.

$$\forall t \in \{1, \dots, 5\} : \sum_{k \in s} \omega_k T_{k,t}(Z) = T(Z)$$

$T_{k,t}(Z)$  = total d'une variable d'intérêt Z (de calage)

sur les communes du groupe de rotation  $t$  dans la ZAE  $k$ .

$T(Z)$  = vrai total de la variable d'intérêt Z dans l'univers.

$\omega_k$  = nouveau poids de la ZAE  $k$  à l'issue du calage.



- On part de poids initiaux donnés par l'estimateur en expansion
- Le calage est effectué nationalement (sauf en cas d'extension régionale) et pour chaque groupe de rotation séparément => **plusieurs jeux de poids calés.**



## Variables de calage utilisées :

- les **variables d'équilibrage lors du tirage des ZAE** (nombre de résidences principales, revenu fiscal total, âge en trois tranches, nombre de résidences principales dans les espaces urbain/périurbain/rural)
- les variables **d'emploi par secteur**
- la répartition par **tranches de taille d'unité urbaine.**



## Résultats du calage en termes d'erreurs relatives

Erreur relative égale à zéro sur les variables de calage et qui n'augmente pas pour les autres variables d'intérêt : on n'observe plus d'erreurs relatives importantes.



## Exemple d'impact du calage sur la variable « demandes d'emploi en fin de mois » (ANPE).

variable	Groupe	Expansion Non calé UP 99	Valeur Estimateur "Expansion Non calé"	Erreur relative scenario 1 "Expansion Non calé"	Erreur relative "Expansi on calé "
<b>DEFM 31/12/2007</b>	1		2 105 629,5	0,2	-0,5
	2		2 085 597,4	-0,7	-0,2
	3		2 101 086,9	0,0	0,8
	4		2 074 397,7	-1,2	-0,6
	5		2 111 643,7	0,5	0,0
		Valeur réelle	2 100 415,0		

# Impact du calage sur la dispersion du poids des logements pour un échantillon de 20000 logements

Rapports inter-fractiles	Groupe	D90/D10	D95/D5	D97,5/D2,5
Scénario sans calage des ZAE	1	1,6	2,3	3,4
	2	1,4	2,1	3,5
	3	1,5	2,1	3,1
	4	1,5	2,4	3,7
	5	1,5	2,3	3,4
Scénario avec calage des ZAE	1	1,7	2,3	3,2
	2	1,7	2,3	3,5
	3	1,5	2,0	3,1
	4	1,6	2,7	4,0
	5	1,6	2,4	3,5



## Conséquences du calage.

- Les nouveaux poids calés des ZAE (*recalculés chaque année*) sont réinjectés dans le calcul des allocations de logements par ZAE...
  - .. lesquelles *dépendent donc de ces nouveaux poids aléatoires* et non plus directement des poids initiaux des ZAE.
- En théorie, le calage permet d'augmenter les allocations dans les zones sous-représentées et de les diminuer dans les zones surreprésentées.
- En pratique, cette propriété est affaiblie par les contraintes de nombre minimum et maximum de fiches-adresses à tirer par ZAE.



## Conclusion sur le calage des ZAE

- › Une méthodologie innovante de calage des unités primaires.
- › Une validation empirique de cette méthode sur la base de l'échantillon de ZAE tiré.
- › Permettra d'incorporer une **information récente** en calant sur les données du nouveau RP, disponibles début 2009 (alors que la construction et le tirage des ZAE ont utilisé des données du RP 1999).



## VII. Conclusion et perspectives.

- Le recours à un RP annuel permet un gain substantiel en qualité de la base de sondage (possibilité de cibler sur des caractéristiques fraîches).
- Mais complexification importante du processus d'échantillonnage.
- Un problème majeur à résoudre: le calcul de précision



*Merci de votre attention !*



# ANNEXES

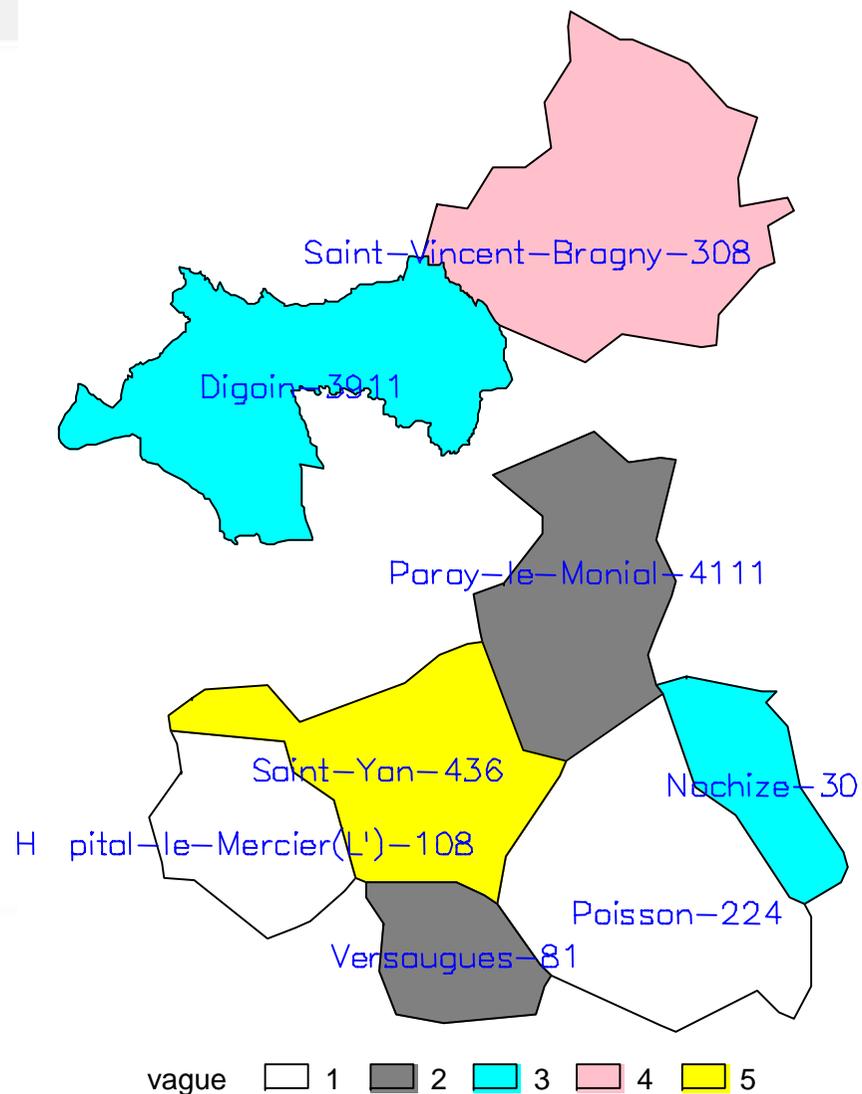




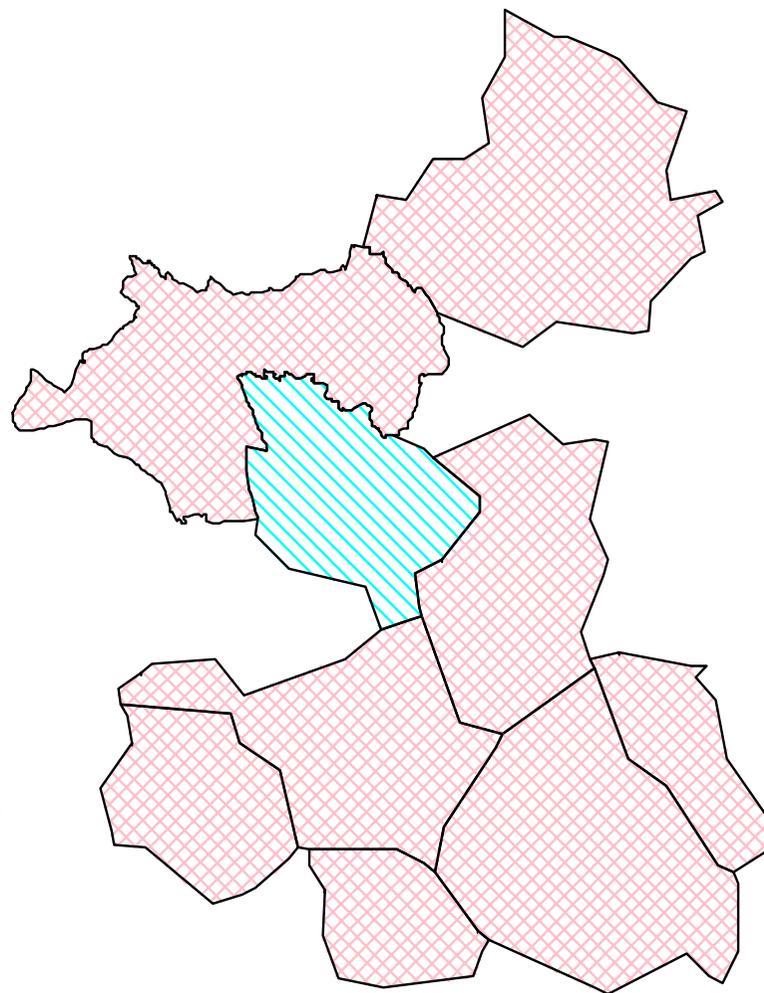
## Impact du tirage des groupes de rotation sur la constitution des ZAE

- › Exemple: cas des communes de Paray-le-Monial et Vitry en Charollais
- › Paray-le-Monial: plus grande petite commune de la région Bourgogne (4111 résidences principales)
- › => testée en premier en tant que commune pivot dans l'algorithme
- › ZAE constituée avec 9209 résidences principales

# ZAE de Paray-le-Monial (Z71342)



# ZAE de Paray-le-Monial et commune de Vitry-en-Charollais



# Impact du tirage des groupes de rotation sur les ZAE

- › Vitry-en-Charollais (349 résidences principales): petite commune la plus proche de Paray-le-Monial
- › Appartient comme Paray-le-Monial au groupe 2, donc ne fait pas partie de la même ZAE
- › Aurait fait partie de la ZAE de Paray-le-Monial si elle avait été dans un autre groupe de rotation (donc d'une ZAE contenant au moins 5360 résidences principales)
- › Appartient à une ZAE de taille plus réduite (1770 résidences principales), donc ayant une probabilité de tirage plus faible