

# Document de travail n° 40

Construction/Logement

## RECONSTITUTION DE LA VARIABLE CODE COMMUNE ABSORBÉ

Millésimes 1999 et 2001 de  
Filocom

**Auteurs : Nadège Couvert, Nina Guyon, Joyce Sultan Parraud**

# Sommaire

<b>1</b>	<b>- Introduction</b>	<b>4</b>
<b>2</b>	<b>- Signification du code commune absorbé et particularités des millésimes 1999 et 2001</b>	<b>6</b>
2.1	Signification du code commune absorbé	6
2.2	Particularités des millésimes 1999 et 2001	9
<b>3</b>	<b>- Correction de la variable dans le millésime 2001</b>	<b>9</b>
3.1	Étape 1 : appariement avec le millésime 2003	10
3.2	Étape 2 : utiliser l'identifiant de copropriété dans Filocom en 2001	11
3.3	Étape 3 : utiliser la propriété de panel de logements en regardant tous les millésimes de Filocom (2005 à 2015)	12
3.4	Étape 4 : utiliser l'attribution ordonnée de l'identifiant logement	12
<b>4</b>	<b>- Correction de la variable dans le millésime 1999</b>	<b>13</b>
4.1	Étape 1 : appariement avec le millésime 2001 corrigé	13
4.2	Étape 2 : cas des villes de Paris, Lyon, Marseille et Toulouse	14
4.3	Étape 3 : utiliser la propriété de panel de logements en regardant tous les millésimes de Filocom (2005 à 2015)	14
4.4	Étape 4 : utiliser l'attribution ordonnée de l'identifiant logement	15
<b>5</b>	<b>- Livrables</b>	<b>16</b>
5.1	Précisions sur l'identifiant logement	16
5.2	Tables corrigées	17
5.3	Programme	17

# 1 – Introduction

Dans le Fichier des logements à la commune (Filocom), les logements et les ménages fiscaux sont géolocalisés au niveau de la section cadastrale<sup>1</sup>. Le plan cadastral représente graphiquement le territoire communal. Il est subdivisé en sections et lieux-dits, et une section peut comporter une ou plusieurs parcelles. Historiquement, le cadastre de France a été réalisé dans la première moitié du XIX<sup>e</sup> siècle (on parle de cadastre napoléonien) ; la plupart des communes ayant rénové une ou plusieurs fois leur cadastre depuis. Ceci explique pourquoi il existe des différences dans la désignation des sections<sup>2</sup>. Le cadastre sert de base au calcul de la taxe foncière.

Dans le cadre de l'étude sur les effets du Programme national de rénovation urbaine (PNRU) menée par Nina Guyon<sup>3</sup> et dont l'objectif est de comparer la situation des ménages concernés par ce programme avec d'autres ménages aux caractéristiques semblables mais non concernés par ce programme, il a été nécessaire d'établir une correspondance entre « appartenance à une section cadastrale » et « appartenance à une Zone Urbaine Sensible (ZUS) ou à un autre Quartier de la Rénovation Urbaine (QRU) dans Filocom ».

Toutefois, la définition parfaite des ZUS et des QRU en termes de sections cadastrales n'est possible que pour le millésime 2009 de Filocom car le Commissariat général à l'égalité des territoires (CGET<sup>4</sup>) dispose d'une couche géographique des sections cadastrales millésimée 2009. En raison de la partielle instabilité des identifiants de section cadastrale dans le temps (la dénomination et la géographie des sections cadastrales pouvant évoluer au cours du temps), la correspondance établie en 2009 entre les QRU et les sections cadastrales n'est pas parfaitement assurée pour les autres millésimes. Le CGET a donc créé un panel de sections cadastrales afin de pouvoir déterminer les logements se trouvant dans un quartier de la rénovation urbaine pour chaque millésime (autre que 2009).

L'identifiant de section cadastrale est obtenu en concaténant le code commune Insee (variable « depcom »), le code commune absorbé (« comabs ») et le code de section cadastrale (« secca ») dans Filocom. Afin d'intégrer les millésimes 1999 et 2001 de Filocom au panel décrit plus haut, il a été nécessaire de reconstituer un code commune absorbé valide car la variable était manquante ou tronquée respectivement en 1999 et 2001.

---

<sup>1</sup> Le fichier Filocom est exhaustif pour la France métropolitaine. Il ne couvre cependant pas les logements des îles de Sein et Molène (Finistère) qui ne sont pas cadastrées.

<sup>2</sup> Dans le cas des plans rénovés par voie de mise à jour, la désignation d'une section cadastrale se fait généralement par une lettre suivie d'un indice qui indique le numéro de la feuille. Dans le cas des plans renouvelés et refaits ou issus d'aménagements fonciers, une section est désignée par deux lettres. Ces règles souffrent toutefois d'exceptions, notamment pour les plans ayant fait l'objet des premiers renouvellements. Par ailleurs, les sections cadastrales des départements de Moselle, du Bas-Rhin et du Haut-Rhin sont désignées numériquement.

<sup>3</sup> Enseignante-chercheuse à la National University of Singapore (NUS) et chercheuse associée au Laboratoire Interdisciplinaire d'Evaluation des Politiques Publiques (LIEPP) de Sciences Po. Nina Guyon a accès à Filocom dans le cadre d'une convention tripartite entre le SDES, le CGET et le LIEPP.

<sup>4</sup> Ex SGCIV, Secrétariat général du Comité interministériel des villes.

Ce document a pour objectif de décrire précisément les procédures et les choix opérés pour reconstituer cette variable dans Filocom. La première partie rappelle brièvement la signification du code commune absorbé et les problèmes intrinsèques aux millésimes 1999 et 2001 relatifs à cette variable. La deuxième partie détaille les étapes mises en œuvre pour reconstituer un code commune absorbé.

Le travail méthodologique ayant conduit à la création du panel de sections cadastrales fera l'objet d'une publication spécifique au sein des collections du SDES.

## 2 – Signification du code commune absorbé et particularités des millésimes 1999 et 2001

Filocom résulte de l'appariement de quatre fichiers gérés par la Direction générale des finances publiques (DGFIP) : le fichier de la taxe d'habitation, le fichier foncier (des propriétés bâties), le fichier des propriétaires et le fichier de l'impôt sur les revenus des personnes physiques. Le champ couvert est celui des locaux à usage d'habitation, assujettis à la taxe d'habitation (logements occupés, qu'ils soient imposés ou exonérés) ou non (logements vacants ou meublés soumis à la cotisation foncière des entreprises<sup>5</sup>). C'est actuellement le seul fichier permettant de faire le lien entre les logements et leurs occupants. Par ailleurs, la présence d'un identifiant logement unique, et stable dans le temps, permet de reconstituer les parcours des logements et d'étudier partiellement les mobilités résidentielles. Enfin, la localisation des logements et des ménages est très précise<sup>6</sup>.

Pour les raisons évoquées ci-dessus et grâce à sa profondeur temporelle importante<sup>7</sup>, Filocom est largement utilisé pour l'évaluation de politiques publiques dans les domaines du logement et de la politique de la ville. La possibilité de reconstituer des zones géographiques stables dans le temps (par agrégation de sections cadastrales grâce au panel) pourra être utile pour des évaluations futures. Pour ce faire, la reconstitution du code commune absorbé, qui contribue à l'identification unique d'une section cadastrale au sein d'une commune est nécessaire. Dans une première partie, la signification de la variable est explicitée tandis que les particularités des millésimes 1999 et 2001 sont exposées dans une seconde partie.

### 2.1 Signification du code commune absorbé

Les communes utilisent les mêmes codes de sections et parcelles cadastrales. En cas de fusion/absorption, il n'est plus possible de les distinguer avec le même code commune.

La variable code commune absorbé correspond communément au préfixe de section cadastrale qui contribue à identifier de manière unique des parcelles provenant de communes fusionnées ou absorbées. Par défaut, cette variable est à blanc sauf dans le cas des communes disposant d'arrondissements (voir ci-après).

Dans le cas d'une absorption d'une commune par une autre, la commune absorbante conserve ses identifiants parcellaires en l'état. En revanche, la commune absorbée hérite du code Insee de la commune absorbante et conserve son ancien code Insee en tant que préfixe (code commune absorbé, initialement à blanc). Par exemple, si deux communes 153 et 267 ont été absorbées par la commune 018, les parcelles AB 25 existantes sur les trois communes seront respectivement désignées comme suit :

---

<sup>5</sup> Qui correspond à l'ancienne taxe professionnelle.

<sup>6</sup> La précision de la localisation dépend de la taille des secca

<sup>7</sup> Le premier millésime de Filocom date de 1995, mais la présence d'un identifiant logement signifiant et stable ne remonte qu'à 1999.

**Tableau 1 : code commune absorbé dans le cas d'une absorption**

Code commune	Code commune absorbé	Code section cadastrale	Parcelle
018	153	AB	25
018	267	AB	25
018		AB	25

Dans le cas d'une fusion de deux communes dans une nouvelle commune, le code commune absorbé permet d'identifier les parcelles venant des communes fusionnées. Par exemple, les communes 153 et 267 ont été fusionnées pour donner la commune 018, alors les nouvelles parcelles de la commune 018 seront identifiées de la manière suivante :

**Tableau 2 : code commune absorbé dans le cas d'une fusion**

Code commune	Code commune absorbé	Code section cadastrale	Parcelle
018	153	AB	25
018	267	AB	25

Dans Filocom, la variable contribue donc à la construction de l'identifiant unique d'une section cadastrale par combinaison avec les variables « secca » (lettres de section) et « depcom » (code Insee de la commune où se situe le logement). Elle est présente dans le fichier à partir du millésime 2001. La variable a la particularité d'être codée sur quatre positions, d'un à quatre caractères, contrairement à d'autres bases (produites par la DFGiP : Plan cadastral informatisé, Fichiers fonciers etc., ou par l'Institut national de l'information géographique et forestière : BD PARCELLAIRE®) où elle est au maximum sur trois positions. Le codage sur quatre caractères ne vaut que pour la ville de Marseille.

Comme nous l'avons vu plus haut, la signification de la variable diffère selon les communes. Le tableau qui suit récapitule sa signification et les valeurs possibles pour les communes de Filocom :

**Tableau 3 : signification du comabs dans Filocom**

<b>Commune</b>	<b>Signification</b>	<b>Valeurs du comabs</b>
Paris	Code arrondissement	101 à 120
Lyon	Code arrondissement	381 à 389
Marseille	Concaténation du code arrondissement et d'un code quartier	Code arrondissement à deux chiffres (01 à 16) + un code quartier à 2 chiffres (de 01 à XX, nombre variable de quartiers par arrondissement, avec un maximum de 11 quartiers par arrondissement) <sup>8</sup> .
Toulouse	Code quartier	01 à 46
Communes ayant subi des fusions et ou des absorptions	Ancien code Insee de la commune absorbée	Code commune Insee à 3 chiffres
Autres communes (hors Paris, Lyon, Marseille et Toulouse et n'ayant subi ni fusion, ni absorption)	Sans objet	À blanc

NB : les 3 derniers chiffres de la variable depcom dans Filocom correspondent aux arrondissements dans les villes de Paris (75101 à 75120 pour les millésimes 99 à 2013 et 75001 à 75020 pour les millésimes 95 à 97), Lyon (69381 à 69389) et Marseille (13201 à 13216). Pour Toulouse, le depcom correspond au code commune Insee (31555).

<sup>8</sup> La documentation relative à la base Filocom est erronée à ce sujet (elle ne mentionne pas l'existence du code quartier propre à la ville de Marseille).

## 2.2 Particularités des millésimes 1999 et 2001

Dans les millésimes 2003 à 2015 de Filocom, le code commune absorbé est composé d'un à quatre caractères alors qu'il est tronqué à trois caractères à Marseille en 2001<sup>9</sup>. Les sections cadastrales portant le même code au sein d'un même arrondissement (mais qui seraient situées dans différents quartiers de ce même arrondissement) ne sont donc pas différenciables. La correction proposée ci-après pour le millésime 2001 ne concernera donc que Marseille. En revanche, dans le millésime 1999, la variable code commune absorbée est inexistante : il faut donc la recréer pour toutes les communes. Comme les fichiers sources permettant de créer Filocom n'existent plus, il n'est pas plus possible de recréer cette variable *ex post*, il faut l'imputer.

## 3 – Correction de la variable dans le millésime 2001

Il est possible d'imputer un code commune absorbé aux logements présents dans le millésime 2001 à partir du millésime 2003 à partir de la propriété de panel de logements<sup>10</sup>. La méthode proposée ici est simple : il suffit de récupérer pour 2001 le code commune absorbé des logements en 2003 grâce à la présence dans Filocom d'un identifiant logement significatif et stable dans le temps. Les fichiers reçus de la DGFIP servant à constituer Filocom contiennent en effet un identifiant pour chaque local (« idfilo ») créé au niveau de chaque direction fiscale. La concaténation du code de direction fiscale et de l'idfilo permet d'obtenir un identifiant logement stable dans le temps.

La méthode proposée appelle néanmoins trois commentaires :

1. Elle suppose tout d'abord que la propriété de panel de logements est bien vérifiée dans Filocom, c'est-à-dire que la réattribution d'un identifiant logement à un logement nouvellement apparu dans le fichier n'est pas possible. Empiriquement, il arrive que des logements disparaissent temporairement d'un millésime (c'est-à-dire que des identifiants de logements disparus d'un millésime donnés sont retrouvés dans des millésimes qui suivent, après leur disparition temporaire). Or, d'après des discussions avec la DGFIP et le Cerema Nord-Picardie, les identifiants des logements détruits ne sont jamais réattribués à des logements nouvellement construits. En revanche, il est possible que des locaux assujettis à la taxe d'habitation ne le soient plus ponctuellement et qu'ils disparaissent du fichier Filocom. Par exemple, il est possible qu'un logement (précisément un local à usage d'habitation) devienne temporairement un local commercial ou à usage professionnel.
2. Par ailleurs, l'attribution d'un code commune absorbé en N à partir du millésime N+2 suppose que les découpages entre arrondissements et quartiers sont restés fixes entre N et N+2. Cette hypothèse n'est pas vérifiée.

---

<sup>9</sup> Ceci est probablement dû à un problème d'import des fichiers bruts qui ont permis la constitution du millésime 2001. Ces fichiers ont été détruits.

<sup>10</sup> Il n'est pas possible d'utiliser en premier lieu le millésime 1999 car la variable comabs n'existe qu'à partir de 2001.

3. Enfin, pour les logements détruits entre 2001 et 2003, il n'est pas possible de récupérer un code commune absorbé en 2001 à partir du millésime 2003. C'est pourquoi la correction du code commune absorbé nécessite des étapes supplémentaires pour les logements qui auraient été détruits entre deux millésimes.

Il est ensuite possible de récupérer un code commune absorbé pour le millésime 1999 grâce à l'identifiant logement, mais ce n'est pas le cas pour les millésimes les plus anciens de Filocom (1995 et 1997) qui ne contiennent pas l'identifiant logement.

### **3.1 Étape 1 : appariement avec le millésime 2003**

Pour rappel, la troncature du code commune absorbé ne concerne que la ville de Marseille en 2001. La variable `comabs` est donc corrigée uniquement pour cette ville. Par ailleurs, il est rappelé que le code commune absorbé permet de retrouver l'arrondissement et le quartier à Marseille (sa présence ne traduit pas l'absorption d'une commune par une autre).

L'appariement des millésimes 2001 et 2003 à Marseille donne les résultats suivants :

- 404 612 logements sont présents dans les deux millésimes : pour ces logements, il est possible de réattribuer le code commune absorbé de 2003 en 2001.
- 7 353 logements apparaissent en 2003 (n'existent pas en 2001) : ces logements ne posent aucun problème particulier. La plupart de ces logements correspondent à des constructions nouvelles mais il peut aussi s'agir de logements ayant changé de destination.
- Enfin, 3 031 logements disparaissent en 2003 alors qu'ils étaient présents dans le millésime 2001 : il est impossible pour ces logements de récupérer un code commune absorbé valide en utilisant seulement la propriété de panel de logements de Filocom.

Pour les 404 612 logements présents en 2001 et 2003, il est possible de récupérer un code commune absorbé valide si deux conditions sont vérifiées. La première est que les logements présents dans les deux millésimes sont bien situés dans les mêmes arrondissements. Cette condition est satisfaite (elle est vérifiée grâce à la variable « `depcom` » qui est précise au niveau de l'arrondissement à Marseille).

La seconde consiste à vérifier que le code commune absorbé en 2001 (tronqué aux trois premiers chiffres) correspond bien aux 3 premiers chiffres du code commune absorbé corrigé (récupéré dans le millésime 2003). C'est bien le cas pour 404 583 logements. En revanche, 29 logements ont un code commune absorbé en 2001 différent des trois premiers caractères du code commune absorbé corrigé. Ces cas correspondent à des problèmes de valeurs manquantes :

1. Le code commune absorbé est manquant en 2001 alors qu'il ne l'est pas en 2003, soient 20 logements. Deux comabs concernés :
  - comabs 0906 en 2003 (une seule section : 13209\_0906\_M) ;
  - comabs 1004 en 2003 (une seule section : 13210\_1004\_U) ;
  - Donc comme pour les secca des 404 583 autres logements : on peut récupérer le comabs de 2003.
2. Le code commune absorbé est manquant en 2003 mais ne l'est pas en 2001 pour 9 logements. Une seule section est concernée en 2001 : la section 13216\_160\_B qui devient en 2003 la section « 13216\_\_ » (c'est-à-dire qu'elle perd son code section en 2003, car elle passe sous le seuil de diffusion). Pour ce cas, il a été décidé de ne pas conserver le code commune absorbé de 2001, qui est de toute façon tronqué à 3 caractères.

Après cette première étape, il est possible de corriger le code commune absorbé de 404 459 logements (parmi les 407 643 logements présents à Marseille en 2001).

### 3.2 Étape 2 : utiliser l'identifiant de copropriété dans Filocom en 2001<sup>11</sup>

Cette deuxième étape consiste à utiliser les informations géographiques relatives aux copropriétés. *A priori*, les logements d'une même copropriété doivent en effet avoir le même code commune absorbé : s'il est possible de récupérer le code commune absorbé d'un logement d'une copropriété, il est alors possible de réattribuer ce même code commune absorbé à l'ensemble des logements de la copropriété.

À noter qu'il existe deux types de copropriétés : les copropriétés dites « verticales » et les copropriétés dites « horizontales ». Ces deux types de copropriétés sont régis par les mêmes textes. Leur différence réside dans le bâti et la configuration : une copropriété verticale correspond à un immeuble collectif divisé en plusieurs appartements et une copropriété horizontale correspond à une ou plusieurs constructions sur un terrain commun.

Il est possible qu'une copropriété horizontale repose sur plusieurs sections cadastrales. Et il est également possible (mais très rare) d'avoir des logements situés sur des sections différentes au sein d'un même immeuble collectif. Enfin, il peut y avoir des copropriétés mixtes (verticale et horizontale). Par souci de simplicité, seules les copropriétés situées sur une seule et même section cadastrale (un seul identifiant de section, un seul code commune absorbé) sont prises en compte ici.

Au total, 245 358 logements en copropriété à Marseille en 2001 (soient 7 161 copropriétés au total) sont décomptés. Parmi ces logements, 244 369 ont un code commune absorbé corrigé après l'étape 1 (ils sont répartis dans

---

<sup>11</sup> Attention, dans Filocom, l'identifiant de copropriété n'est pas stable dans le temps : constituer un panel de copropriétés n'est pas faisable.

7 136 copropriétés). Il reste alors 25 copropriétés sans code commune absorbé après la première étape : pour les logements de ces copropriétés il sera donc impossible de déduire un code commune absorbé valide. 509 copropriétés sont situées sur plus d'une section cadastrale parmi les 7 136 copropriétés avec un code commune absorbé corrigé : il n'est pas possible de récupérer un code commune absorbé pour les logements de ces copropriétés. Au total, 6 627 copropriétés sont situées sur une seule et même section, ce qui conduit à récupérer une variable valide pour 574 logements supplémentaires grâce à l'identifiant de copropriété en 2001.

Après les deux premières étapes, il est possible de corriger le code commune absorbé de 405 033 logements (parmi les 407 643 logements présents à Marseille en 2001).

### **3.3 Étape 3 : utiliser la propriété de panel de logements en regardant tous les millésimes de Filocom (2005 à 2015)**

Comme énoncé plus haut, il arrive que des logements disparaissent temporairement d'un millésime de Filocom. Ainsi, 347 logements disparus en 2003 existent dans des millésimes postérieurs à cette date (c'est-à-dire qu'ils sont retrouvés en 2005, et/ou en 2007, et/ou en 2009 etc.). Parmi ces logements, 119 ont un code commune absorbé à blanc dans les millésimes où on les retrouve.

Les 228 logements restants ont un comabs renseigné (comabs qu'ils conservent en l'état à partir du millésime où ces logements sont retrouvés et dans les millésimes qui suivent). Il est raisonnablement possible de leur réattribuer ce code commune absorbé en 2001.

Après les trois premières étapes, 405 261 logements (parmi les 407 643 logements présents à Marseille en 2001) ont désormais un code commune absorbé corrigé.

### **3.4 Étape 4 : utiliser l'attribution ordonnée de l'identifiant logement**

À ce stade, 2 382 logements n'ont pas de code commune absorbé corrigé à Marseille en 2001. Pour ces logements, il est possible d'utiliser en dernier recours, la manière dont sont construits les identifiants logements dans Filocom.

L'identifiant logement Filocom (variable « idfilo ») a été attribué séquentiellement<sup>12</sup> (probablement en 1999) au sein de chaque direction fiscale. Cet identifiant est donc un simple compteur. Pour les nouveaux locaux (construits après 1999), chaque nouvel idfilo est attribué séquentiellement à partir du dernier attribué en 1999. **Cela signifie d'une part que la numérotation initiale (celle du millésime 1999) est cohérente**

---

<sup>12</sup> D'après le dictionnaire des variables Filocom et des discussions avec le Cerema Nord-Picardie et la DGFIP.

**avec le découpage entre arrondissements et quartiers et d'autre part qu'il est possible de déduire le comabs d'un logement en fonction de deux paramètres : sa position dans le fichier Filocom et par le comabs des logements qui l'encadrent.**

Par exemple, à Marseille : le logement 0104510 non présent en 2003 est encadré par le logement 0104509 et le logement 0104511, tous les deux dans le premier quartier du premier arrondissement (comabs « 0101 ») en 2003 (et donc dans le 0101 en 2001, puisque le comabs de 2001 a été déduit à partir de 2003 principalement). L'hypothèse selon laquelle le logement 0104510 est aussi situé dans le premier quartier du premier arrondissement de Marseille (comabs « 0101 ») est raisonnable.

En revanche, il est impossible de déduire le code commune absorbé d'un logement qui est encadré par des logements ayant un comabs différent avec cette méthode. Par exemple, toujours à Marseille, le logement 0144097 est encadré par les logements 0144096 (avec un code commune absorbé à blanc) et 0144098 avec un comabs égal à 0204 et il est impossible de trancher entre ces deux options.

La méthode basée sur l'attribution ordonnée de l'identifiant logement conduit à retrouver le code commune absorbé de 2 079 logements supplémentaires.

Après avoir mis en place les quatre étapes décrites ci-dessus, la quasi-totalité des logements marseillais ont un code commune absorbé non vide, à l'exception de 303 logements parmi les 407 643 logements que compte la ville en 2001.

## **4 – Correction de la variable dans le millésime 1999**

L'imputation d'un code commune absorbé valide pour les logements du millésime 1999 nécessite la mise en œuvre des mêmes étapes que pour la correction de la variable dans le millésime 2001. Cependant, l'identifiant de copropriété n'existe pas dans ce millésime (et les fichiers bruts ont été détruits), l'étape 2 est donc caduque. Par ailleurs, la variable comabs est totalement absente du millésime 1999, contrairement au millésime 2001 dans lequel elle pose problème uniquement à Marseille.

### **4.1 Étape 1 : appariement avec le millésime 2001 corrigé**

29 801 759 logements sont décomptés dans le millésime 1999 de Filocom, dont 308 403 n'existant pas dans le millésime 2001. Pour les 29 493 356 logements présents en 2001, il est possible de récupérer un code commune absorbé valide (mais qui peut être à blanc) à condition que l'absorption soit bien antérieure à 1999. En pratique, si l'absorption a lieu après 1999, les codes commune (variable « depcom ») seront différents

entre les millésimes 1999 et 2001. C'est le cas de 20 467<sup>13</sup> logements retrouvés en 1999 et 2001 et ayant un code commune absorbé non nul en 2001.

Pour les 308 403 logements qui disparaissent, il faut distinguer les logements parisiens, lyonnais, marseillais et toulousains des logements des autres villes.

## **4.2 Étape 2 : cas des villes de Paris, Lyon, Marseille et Toulouse**

Pour les villes de Paris, Lyon, Marseille et Toulouse, le code commune absorbé a une signification particulière (*voir le tableau 3*). Au total, ces villes totalisent 2 284 380 logements dans le millésime 1999 de Filocom dont 34 338 qui disparaissent en 2001 et 744 qui ont un code commune absorbé vide après appariement avec le millésime 1999.

Pour Paris et Lyon<sup>14</sup>, le code commune absorbé est imputé directement respectivement à partir des deux derniers/du dernier caractère(s) de la variable « depcom ».

À Toulouse, 3 136 logements ont un code commune absorbé vide après appariement avec le millésime 2001 (232 750 logements au total). À Toulouse, la méthode basée sur l'attribution ordonnée de l'identifiant logement est utilisée et un comabs valide est récupéré pour l'ensemble des logements, sauf 10 d'entre eux.

Enfin, pour Marseille (402 957 logements en 1999), la méthode basée sur l'attribution ordonnée de l'identifiant logement est également mise en place. 249 logements demeurent avec un comabs à blanc parmi les 3 616 logements qui avaient un comabs vie après l'appariement avec le millésime 2001.

Après les étapes 1 et 2, 29 527 459 logements ont un comabs corrigé en 1999.

## **4.3 Étape 3 : utiliser la propriété de panel de logements en regardant tous les millésimes de Filocom (2005 à 2015)**

Parmi les 274 300 logements sans code commune absorbé valide après les étapes 1 et 2, 14 810 sont retrouvés dans l'un des millésimes de 2005 à 2015, alors qu'ils avaient disparus en 2001.

Parmi ces logements, 244 gardent le même code commune absorbé à partir du millésime où ils sont retrouvés et dans les millésimes qui suivent. L'absorption est postérieure à 1999 pour 21 logements. Pour ceux-là, il n'est pas possible de réattribuer un code commune absorbé en 1999 alors qu'il est raisonnablement possible de le faire pour les 223 restants.

Après ces trois premières étapes, il est possible de corriger le code commune absorbé de 29 527 702 en 1999.

---

<sup>13</sup> Précisément, 39 630 logements sont retrouvés en 1999 et 2001 dans des communes différentes (la variable depcom diffère entre les deux millésimes). Parmi ces logements, 20 467 ont un code commune absorbé non nul en 2001, témoignant d'une absorption postérieure à 1999.

<sup>14</sup> Après appariement avec le millésime 2001, respectivement 23 641 et 4 689 logements sont dénombrés avec un comabs vide à Paris (1 385 245 logements au total) et à Lyon (263 428 logements au total).

## 4.4 Étape 4 : utiliser l'attribution ordonnée de l'identifiant logement

À ce stade, 969 communes avec un code commune absorbé non vide sont décomptées en 1999 et ces 969 communes abritent 4 937 239 logements.

Il reste 274 078 logements pour lesquels il est possible d'imputer une variable cohérente. Ces logements n'existent pas en 2001 (l'appariement avec le millésime 2001 corrigé n'a pas abouti) et ils ont un code commune absorbé vide même après les étapes 2 et 3.

Ces logements sont répartis dans 24 094 communes. Parmi eux, seuls 30 039 sont situés dans des communes avec un code commune absorbé non vide en 1999.

La méthode basée sur l'attribution ordonnée de l'identifiant logement permet de retrouver un code commune absorbé pour 3 896 logements supplémentaires<sup>15</sup>.

Après avoir mis en place les quatre étapes décrites ci-dessus, on décompte 2 455 220 logements avec un code commune absorbé renseigné (soit 8,2 % des logements du millésime 1999, voir le tableau 4).

Tableau 4 : nombre de logements avec un code commune absorbé renseigné dans Filocom

Millésime	Proportion de logements avec un code commune absorbé renseigné	Nombre de logements avec un code commune absorbé renseigné	Nombre total de logements (France métropolitaine)
1999	8.24	2 455 220	29 801 759
2001	8.20	2 490 401	30 368 608
2003	8.09	2 502 709	30 920 514
2005	7.99	2 520 465	31 532 330
2007	7.87	2 537 650	32 239 460
2009	7.75	2 562 451	33 078 985
2010	7.69	2 575 454	33 473 723
2011	7.66	2 591 555	33 820 701
2013	7.67	2 648 703	34 539 969
2015	7.61	2 676 054	35 174 102

Source : base de données MTES-CGDD-Filocom (millésimes 1999 à 2015) d'après la DGFIP

<sup>15</sup> La méthode basée sur l'attribution ordonnée des idfilo pourrait permettre d'attribuer un code commune absorbé non vide pour tous les logements d'une commune dès lors qu'il y a des codes communes absorbés renseignés pour certains logements de ladite commune. Cette attribution forcée n'est pas souhaitable. Elle conduirait à attribuer un code commune absorbé non vide à des logements qui existent en 2001 et pour lesquels le code commune absorbé demeure vide.

## 5 – Livrables

### 5.1 Précisions sur l'identifiant logement

L'identifiant logement présent dans Filocom (variable « idfilo ») est un identifiant créé par chaque direction fiscale. Ces directions fiscales correspondent aux départements (sauf dans les départements 13, 59, 75 et 92 où coexistent plusieurs directions fiscales). L'identifiant idfilo est donc unique au sein d'une même direction fiscale. Un même identifiant idfilo peut en conséquence apparaître plusieurs fois dans un même millésime de Filocom parce qu'il est utilisé simultanément dans différentes directions fiscales.

Les différentes versions du dictionnaire des variables de Filocom suggèrent de concaténer la variable « depcom » (code commune Insee) à l'idfilo pour obtenir un identifiant unique. Cependant, comme l'idfilo est géré et n'a de sens qu'au niveau de chaque direction fiscale, la concaténation de la direction fiscale et de l'idfilo a été choisie comme identifiant logement. Cela implique cependant que des logements peuvent être perdus dans le temps, quand ces derniers "passent" d'une direction fiscale à l'autre (échange de parcelles entre communes ou fusion de communes, quand les communes appartiennent initialement à des directions fiscales différentes). Les identifiants « direction fiscale » + « idfilo » de ces logements changent en conséquence, ce qui rend impossible l'appariement de deux millésimes successifs de Filocom sur cet identifiant. En revanche, avec la concaténation de « depcom » et de l'idfilo, beaucoup plus de logements sont perdus pour le suivi dans le temps. En effet, le problème est toujours le même : des communes voisines pouvant échanger des parcelles, voire fusionner, le code commune des logements impliqués est donc modifié, ce qui modifie leur identifiant en conséquence. Mais le fait de changer de code commune a beaucoup plus de chance de se produire que celui de changer de direction fiscale.

L'exemple suivant permet d'illustrer le propos :

Au 01/11/2010, les communes de Chanteloup-en-Brie (77 085) et Montévrain (77 307) ont échangé des parcelles<sup>16</sup>. Entre les millésimes 2010 et 2011 de Filocom, une partie des logements (n=7) de la section OC de la commune de Montévrain est effectivement transférée dans une section également appelée OC de la commune de Chanteloup-en-Brie. Il y a également des échanges dans l'autre sens (de Chanteloup-en-Brie vers Montévrain) mais ils n'apparaissent peut être pas tels quels parce que les sections impliquées sont sous le seuil de diffusion ou bien que les parcelles impliquées dans l'échange étaient non habitées. Pour ces 7 logements, l'identifiant logement passe ainsi entre 2010 et 2011 :

- pour depcom+idfilo : de 77 307\_idfilo à 77 085\_idfilo ;
- depcom+comabs+idfilo: de 77 307\_000\_idfilo à 77 085\_000\_idfilo (dans le cas d'une fusion, le comabs aurait également été modifié).

Si les millésimes 2010 et 2011 de filocom sont appariés sur cet identifiant, ces 7 logements sont perdus.

---

<sup>16</sup>

<http://www.insee.fr/fr/methodes/nomenclatures/cog/fichecommunale.asp?codedep=77&codecom=307>

En revanche, quand l'identifiant direction fiscale+idfilo est utilisé, ces 7 logements ne sont pas perdus car ils continuent d'évoluer au sein de la même direction fiscale (770) et leur identifiant n'est pas modifié.

## 5.2 Tables corrigées

Deux tables corrigées sont mises à dispositions, « fracorr01 » et « fracorr99 », contenant les colonnes suivantes :

- Idfilo : l'identifiant logement propre à chaque direction fiscale, variable brute filocom ;
- Depcom : code commune Insee, variable très légèrement redressée (voir ci-après, programme sur demande) ;
- Secca : code de section cadastrale, variable brute filocom ;
- Comabs : la variable donnant le code commune absorbé corrigé ;
- Idlog : l'identifiant logement issu de la concaténation du code de direction fiscale et de l'idfilo.

Il est donc possible pour les personnes qui le souhaitent d'apparier ces tables sur l'identifiant issu de la concaténation de la « direction fiscale » et de « l'idfilo » ou sur celui issu de la concaténation du « depcom » et de « l'idfilo ».

## 5.3 Programme

Le programme corrigeant la variable code commune absorbé a été écrit en SAS et est disponible sur demande.

**Ministère de la Transition écologique et solidaire**  
Commissariat général au Développement durable  
Service de la donnée et des études statistiques  
Tour Séquoia  
92055 La Défense cedex  
Courriel : [diffusion.sdes.cgdd@developpement-durable.gouv.fr](mailto:diffusion.sdes.cgdd@developpement-durable.gouv.fr)