

# Le Fichier des LOgements par COmmunes - FILOCOM

<b>Objet</b>	Connaître les logements assujettis ou non à la taxe d'habitation et leurs occupants				
<b>Source</b>	Direction générale des finances publiques : rapprochement de fichiers fiscaux de perception d'impôts				
<b>Échelon géographique</b>	Infra-communal	Communal	SCoT/Intercommunal	Départemental/Régional	
<b>Couverture</b>	Totale		Partielle		Variable
<b>Accès</b>	Public		Institutionnels	État	Restreint
<b>Qualité</b>	+++		++	+	-
					--

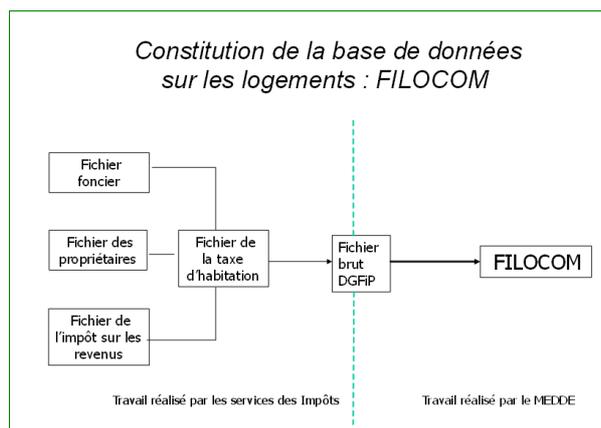
## Description de la base

Le fichier Filocom renseigne sur les logements et leur occupation. Il offre des informations sur le revenu des occupants des logements.

Cette base de données a été construite par la Direction Générale des Finances Publiques (DGFIP) pour les besoins du Ministère en charge du Logement. Elle a été fournie pour la première fois en 1996 à partir des différents fichiers la constituant au 1er janvier 1995.

### Caractéristiques de la source

Filocom est le résultat du traitement automatisé des données issues du fichier de la taxe d'habitation (TH), auquel sont rapprochés le fichier de la taxe foncière (des propriétés bâties), le fichier des propriétaires ainsi que le fichier de l'impôt sur les revenus des personnes physiques.



### Producteur et responsable de l'exploitation :

La Direction Générale des Finances Publiques (DGFIP) est propriétaire des fichiers fiscaux et du fichier brut produit.

L'obtention de Filocom et son exploitation ont nécessité une autorisation de la CNIL<sup>1</sup>, suivie d'une prise d'arrêt ministériel en date du 11 juillet 1994.

En août 1996, une convention a été passée entre la DGI (actuelle DGFIP) et la DAEI (actuel Commissariat Général au Développement Durable – CGDD – du Ministère de l'Écologie, du Développement Durable et de l'Énergie).

Elle définit la finalité du traitement Filocom et les services ou collectivités auxquels des informations peuvent être transmises par le Service de l'Observation et des Statistiques (SOeS) du CGDD.

Le SOeS a confié l'exploitation et l'anonymisation des données (fichier détail) de Filocom au CETE Nord-Picardie.

### Contenu

Le champ couvert est celui des locaux à usage d'habitation assujettis à la taxe d'habitation (logements occupés et imposés ou occupés et exonérés) ou non (logements vacants ou meublés soumis à la taxe professionnelle).

Filocom comporte principalement des logements occupés par des ménages à titre d'habitation, des résidences secondaires, des logements vacants, mais également et marginalement, des logements occupés à d'autres titres par des personnes morales.

Ne sont retenus que les locaux de superficie habitable comprise entre 5 et 1 000 m<sup>2</sup>, et comportant au moins 1 pièce d'habitation. Cette limitation sélectionne automatiquement les maisons et les appartements. Tous les autres locaux se trouvent éliminés : pièces indépendantes, garages, dépendances...

Ne sont pas imposables à la taxe d'habitation, et à ce titre ne figurent pas dans le fichier : <sup>2</sup>

- les locaux passibles de la cotisation foncière des entreprises qui ne font pas partie de l'habitation personnelle ;
- les bâtiments servant aux exploitations rurales ;
- les foyers d'hébergement ;
- les résidences universitaires gérées par les CROUS ;
- les locaux destinés au logement des élèves dans les écoles et pensionnats et au logement des enfants des colonies de vacances ;
- les caravanes dès lors qu'elles conservent en permanence des moyens de mobilité.

<sup>1</sup> Commission nationale informatique et libertés

<sup>2</sup> Liste non exhaustive

Principales variables - Les variables s'articulent autour de quatre grandes thématiques :

<b>Le logement</b>	<ul style="list-style-type: none"> <li>Type : individuel / collectif</li> <li>Nombre de pièces</li> <li>Superficie habitable</li> <li>Année de construction</li> <li>Classement cadastral<sup>3</sup> ....</li> </ul>
<b>L'occupation des logements</b>	<ul style="list-style-type: none"> <li>Mode d'occupation : résidences principales, secondaires, logements vacants</li> <li>Statut d'occupation</li> <li>Durée d'occupation ou de vacance,</li> <li>Sur-occupation ....</li> </ul>
<b>Les occupants des résidences principales</b>	<ul style="list-style-type: none"> <li>Nombre de ménages fiscaux</li> <li>Nombre de personnes</li> <li>Situation de famille de la personne de référence</li> <li>Âge de la personne de référence</li> <li>Revenus bruts</li> <li>Revenu imposable : oui / non</li> <li>Situation par rapport aux plafonds HLM</li> <li>Dégrévement de TH lié aux ressources ....</li> </ul>
<b>Les propriétaires du logement</b>	<ul style="list-style-type: none"> <li>Type (personnes physiques, HLM...)</li> <li>Âge du propriétaire</li> <li>Lieu de résidence</li> <li>Copropriété</li> <li>Date de mutation ....</li> </ul>

Voir le dictionnaire des variables – données de cadrage, du CETE Nord-Picardie – février 2012 (lien en fin de fiche)

## Couverture et disponibilité des données

L'ensemble du territoire métropolitain est couvert. Les résultats ne sont pas encore exploités pour les DOM.

Filocom est arrêté au 1er janvier de chaque année impaire. Sont disponibles les millésimes 1995, 1997, 1999, 2001, 2003, 2007, 2009 et 2011. Chaque millésime décrit la situation du logement au 1er janvier N avec les revenus de N-1.

Les données sont disponibles à la commune et donc pour toute agrégation de communes, dans le respect des règles d'éligibilité et de secret statistique (voir « accès aux données »).

Toutefois, la situation géographique du logement étant connue au mieux par sa section cadastrale, le CETE Nord-Picardie peut fournir sur demande, des extractions infra-communales : section cadastrale ou une agrégation comme les quartiers, dans la mesure où il est possible de faire coïncider des quartiers et des sections cadastrales ou des groupements de sections cadastrales. Pour certaines communes, les données peuvent être fournies directement à l'échelle de quartiers prédéterminés (existence d'un code quartier).

<sup>3</sup> Le classement cadastral est un indicateur de la qualité globale des logements. Il constitue la base de l'évaluation de la valeur locative. Il classe les propriétés en 8 catégories correspondant à l'impression d'ensemble dégagée par l'immeuble et le logement : les deux catégories les plus représentées sont les catégories 5 et 6 (plus des 2/3 du parc de France Métropolitaine). Le parc HLM est le plus souvent en catégorie 5 ou 6. Les catégories 1 à 4 représentent le haut de gamme, les catégories 7 et 8 le bas de gamme.

## Limites et précautions d'usage

### ✓ Contraintes liées au secret statistique :

L'origine fiscale des données implique de respecter le secret statistique. Toute extraction doit faire l'objet d'un examen conduisant à masquer les cases comprenant moins de 11 unités (logements ou ménages) – secret direct – et celles permettant de déduire leur contenu – secret induit (voir page suivante : accès aux données).

Les exploitations fournies par le CETE gèrent le secret un peu différemment : si une section cadastrale comporte moins de 11 unités, ces dernières sont regroupées dans une section cadastrale fictive (ne possédant pas d'identifiant). De même, lorsqu'une commune comporte moins de 11 unités, ces dernières sont regroupées dans une commune fictive ne comportant pas de code commune INSEE.

### ✓ Précautions sur les variables (non exhaustives)<sup>4</sup> :

- L'origine de ce fichier est la perception de l'impôt, principalement de la taxe d'habitation ; la fiabilité des variables et leur mise à jour sont donc meilleures sur les variables qui permettent ce calcul de l'impôt.

Des réserves sont à émettre quant à la comparaison avec des données similaires provenant d'autres sources.

Il faudra donc manier les comparaisons avec précaution, une même variable pouvant avoir plusieurs définitions selon ses différentes sources<sup>5</sup>.

#### Exemple : Le nombre de ménages fiscaux est différent du nombre de résidences principales

Au sens statistique, un ménage désigne l'ensemble des occupants d'un même logement. Selon cette définition utilisée dans le recensement de l'INSEE et dans les enquêtes auprès des ménages, le nombre de ménages coïncide avec le nombre de résidences principales.

Ce n'est pas le cas dans Filocom où l'occupation des résidences principales est analysée par « ménage fiscal » ; dans ce cas un ménage fiscal est constitué par le regroupement des foyers fiscaux répertoriés dans un même logement, que les personnes qui composent ce ménage habitent ou non le logement. Ainsi, les personnes qui ne sont pas indépendantes fiscalement sont comptabilisées dans le ménage où elles sont déclarées à charge et non dans leur résidence principale.

Source : documentation sur l'univers Filocom non secrétisé  
1- Méthodologie de la source Filocom

Il existe donc des logements identifiés comme résidences principales sans information sur leurs occupants, notamment des étudiants rattachés au foyer fiscal de leurs parents (2,3 % des résidences principales en 2011 en Basse-Normandie par exemple).

- Nombre de personnes par logement : il est sur-évalué pour certains logements (voir ci-dessus le cas des étudiants rattachés au foyer de leurs parents).
- Tranche d'âge « moins de 25 ans » : difficilement exploitable pour les mêmes raisons de rattachement.

<sup>4</sup> Voir en fin de fiche « pour en savoir plus... » source Filocom – méthodologie et définition  
– 2 : documentation sur les variables

<sup>5</sup> Voir en fin de fiche « pour en savoir plus... » comparaison RGP / FILOCOM

- Du fait de l'enrichissement progressif du fichier par de nouvelles variables ou de nouvelles modalités, des différences existent entre les versions successives. Ces différences sont indiquées dans les documents téléchargeables.
- Durée d'occupation : remise à zéro lors d'un changement de situation matrimoniale (décès, veuvage, mariage...) de personnes par logement : il est sur-évalué pour certains logements (voir ci-dessus le cas des étudiants rattachés au foyer de leurs parents).

Environ 7,8% du nombre de logement ayant changé d'occupants entre 2007 et 2009 dans Filocom, sont en réalité des changements de situation matrimoniale sans déménagement.

- La vacance est, pour des raisons fiscales, sur-estimée ; un logement vacant dans Filocom, est un logement qui n'est pas habité au 1er janvier.
- Nombre de pièces : variable peu fiable, car peu mise à jour. Privilégier la surface habitable pour évaluer la sur-occupation.
- La variable « classement cadastral » est à utiliser avec précaution en raison d'une actualisation de l'information consécutive à des travaux d'amélioration perfectible. En effet, contrairement aux constructions nouvelles, aux additions de construction et aux changements d'affectation, les travaux d'amélioration intérieure ne sont pas soumis à déclaration obligatoire.

Le phénomène s'applique également à la variable « confort » qui est un regroupement des modalités du classement cadastral.

- Statut d'occupation et concept de logement social : des logements des HLM/SEM peuvent être comptabilisés dans la modalité « locataire d'une collectivité territoriale », et des logements conventionnés (n'appartenant pas au parc HLM/SEM) comptabilisés dans le parc privé.
- Logements neufs : le délai d'intégration des logements neufs rend peu exploitable l'information de l'année n-1 (données CETE). La dernière tranche de l'année de construction est ainsi un peu sous-estimée (données DREAL / DDT-M).

## Accès aux données

**Seuls le SOeS, les DREAL et le CETE Nord Picardie ont la possibilité de traiter les demandes de données Filocom.**

La diffusion doit respecter 3 conditions :

- **Le secret statistique :**
  - aucune valeur publiée ne peut porter sur moins de 11 unités (logement ou foyers fiscaux) ;
  - aucune unité statistique ne doit concourir à plus de 85 % du total d'une variable donnée ;
  - Lors de ventilations multidimensionnelles, si le contenu d'une cellule est occulté du fait que sa valeur repose sur moins de 11 unités statistiques, la reconstitution de la donnée manquante ne doit pas être possible.

Exemple :

	Résidence principale	Résidence secondaire	Vacant	total
Nbre logements	50	5	15	70

	Résidence principale	Résidence secondaire	Vacant	total
Nbre logements	50	sd	15	70

	Résidence principale	Résidence secondaire	Vacant	total
Nbre logements	50	sd	si	70

secret direct

secret induit

- **La finalité des traitements** qui se limite à :
  - l'aide à la définition des politiques locales de l'habitat ;
  - l'aide à la programmation du logement social ;
  - l'aide à l'observation et à la mise en œuvre des politiques locales de l'habitat ;
  - l'évaluation des politiques publiques.
- **La nature des destinataires des données :**

Les données statistiques agrégées confidentialisées peuvent être transmises aux :

- Directions régionales de l'environnement de l'aménagement et du logement ;
- Directions départementales des territoires (et de la mer) ;
- Services de l'État compétents en matière d'habitat de la politique de la ville et d'aménagement du territoire ;
- Collectivités locales et leurs regroupements dotés d'une fiscalité propre sous réserve de signature d'un acte d'engagement et de destruction des données à l'issue de l'étude ;
- Bureaux d'études ou agences mandatés par les destinataires ci-dessus, sous réserve de signature d'un acte d'engagement et de destruction des données à l'issue de l'étude.

## Échelon géographique des données

Les données sont disponibles à l'échelle communale ou regroupement de communes. Les demandes sont à faire auprès :

- des statisticiens régionaux qui jugeront de la faisabilité de la commande. Ils sont habilités à exploiter l'extrait du fichier détail de Filocom diffusé dans **Géokit3**<sup>6</sup> : « Filocom non secrétisé ». Les indicateurs, axes d'analyses et documents prédéfinis disponibles dans cet univers sont présentés dans une documentation réalisée par la division base de données Géokit (*voir le lien vers leur page intranet en fin de fiche*) ;
- du CETE Nord-Picardie, pour l'échelon infra-communal.

Le secret statistique est géré par leurs soins avant transmission des exploitations à des tiers.

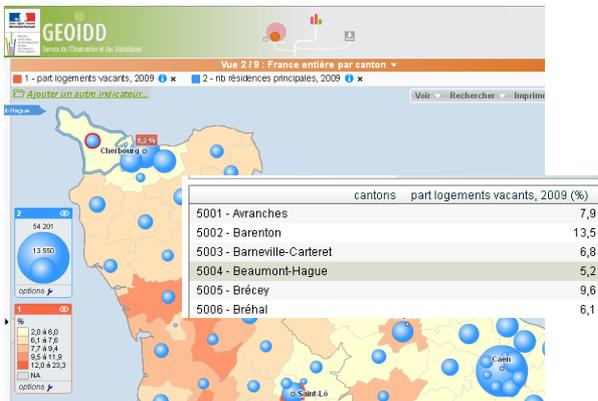
Les fichiers agrégés à la commune avec la **base Beyond 20/20**<sup>7</sup>

Une inscription préalable auprès du SOEs est nécessaire pour pouvoir accéder à la base.

Le secret statistique est géré en amont.

Le Site internet du SOEs propose en outre :

**L'interface de Géoidd**<sup>8</sup> : avec l'« outil de visualisation statistique », quelques données générales cartographiées sont consultables et filtrables à volonté. Elles peuvent être exportées sous forme de tableaux.

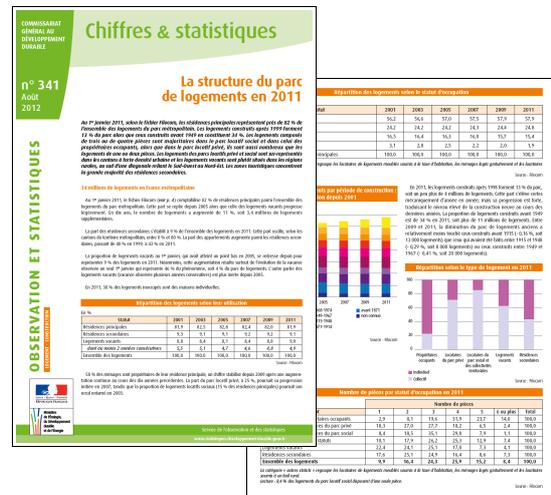


Échelles : communale, cantonale, départementale, régionale

## La publication « chiffres et statistiques » :

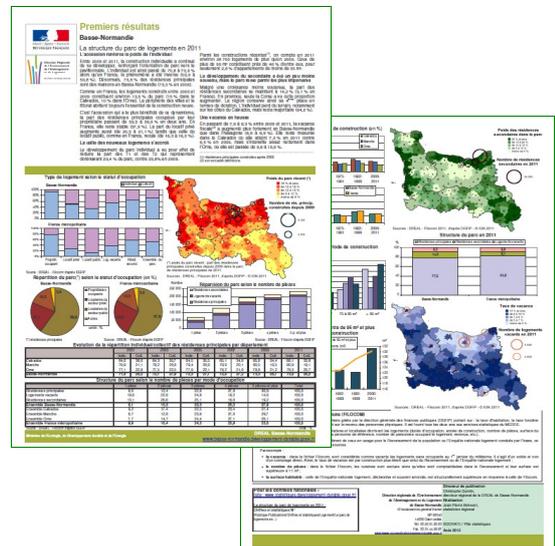
3 numéros sont publiés à l'occasion de la mise à disposition d'un millésime :

- la structure du parc de logement
- les occupants des logements
- les conditions d'occupation des logements



## Les sites internet des DREAL

Les DREAL gèrent leurs sites internet de manière autonome. Les mises à disposition de données sont donc multiples, mais en général, elles proposent des publications périodiques.



6 Géokit3 est un outil interne du Ministère de diffusion des données produites par le SOEs, par l'INSEE ou sources externes (DGFiP, PERVAL...). Géokit 3 est exploité avec un requêteur de type « Business Object ». Géokit3 comprend une application spécifique « zonages » développée en complément pour permettre la gestion de zonages géographiques locaux.

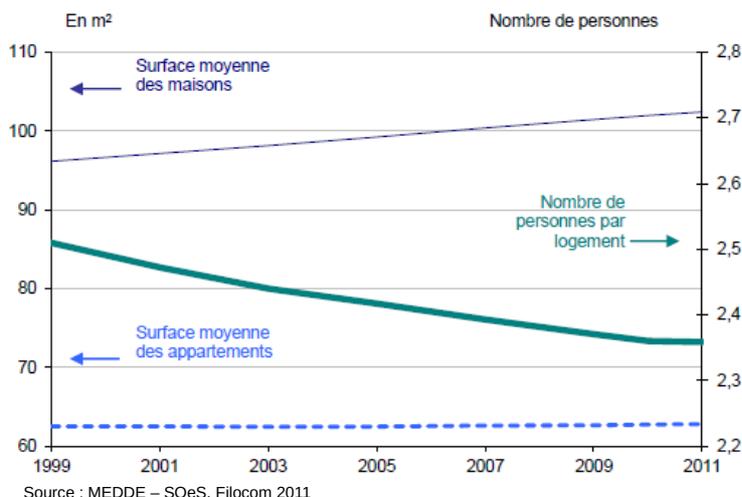
7 Beyond 20/20 est un logiciel permettant d'afficher des données sous forme de tableaux multidimensionnels et de graphiques. Il offre également la possibilité de télécharger des bases de données au format txt.

8 Géoidd (Géographique et indicateurs liés au développement durable) est l'outil de cartographie interactive du Service de l'Observation et des Statistiques (SOEs)

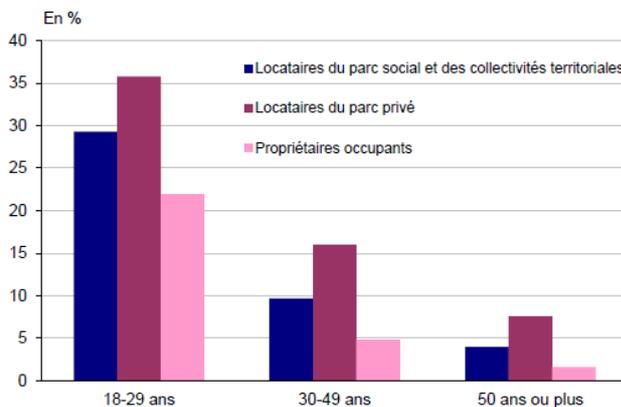
## Exemples d'exploitation

- Références – juin 2013 / CGDD – L'état du logement en 2011

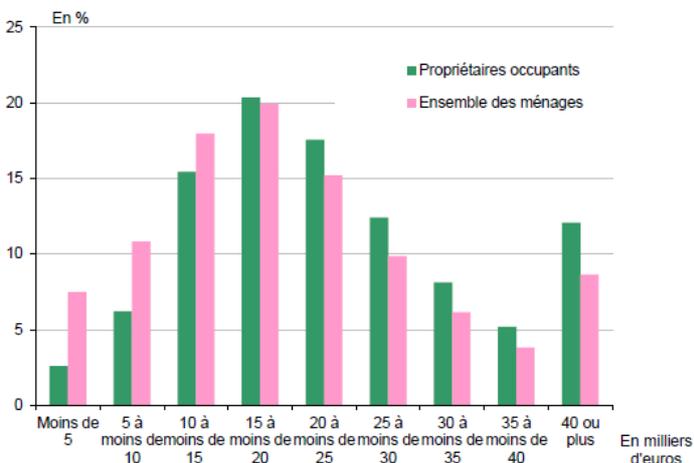
### Des logements de plus en plus spacieux



### Taux moyen de rotation entre 2010 et 2011 selon le parc et l'âge des occupants



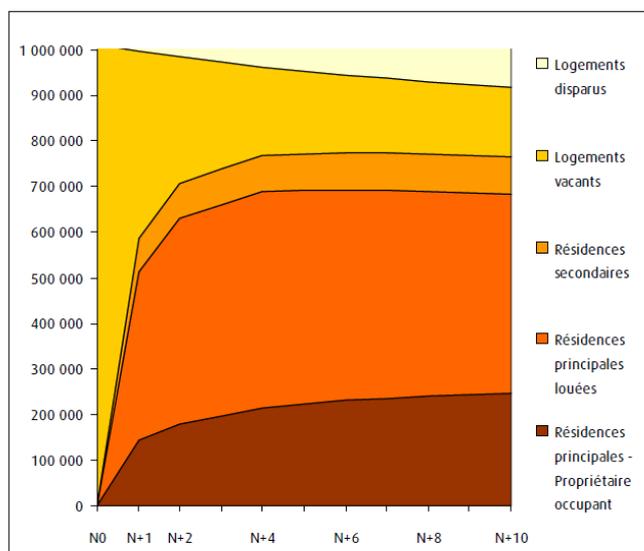
### Distribution des revenus imposables par unité de consommation des propriétaires occupants en 2011



- **Publication : Le point sur .... n° 170 – juillet 2013 / CGDD – observations et statistiques « Une fois sur deux, la vacance d'un logement est suivie d'une réoccupation rapide et durable »**

### Chronogramme des logements entrés en vacance durant l'année 2000

En nombre de logements

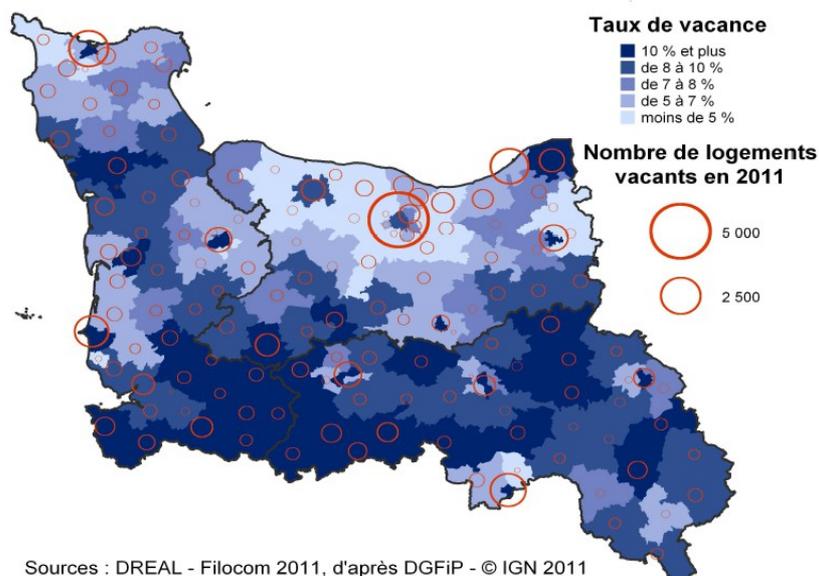


Lecture : sur 1 million de logements entrés en vacance en 2000, on compte 4 ans plus tard 215 000 résidences principales occupées par leur propriétaire, 474 000 résidences principales louées, 80 000 résidences secondaires, 236 000 logements vacants, le reste des logements ayant disparu.

Source : Filocom

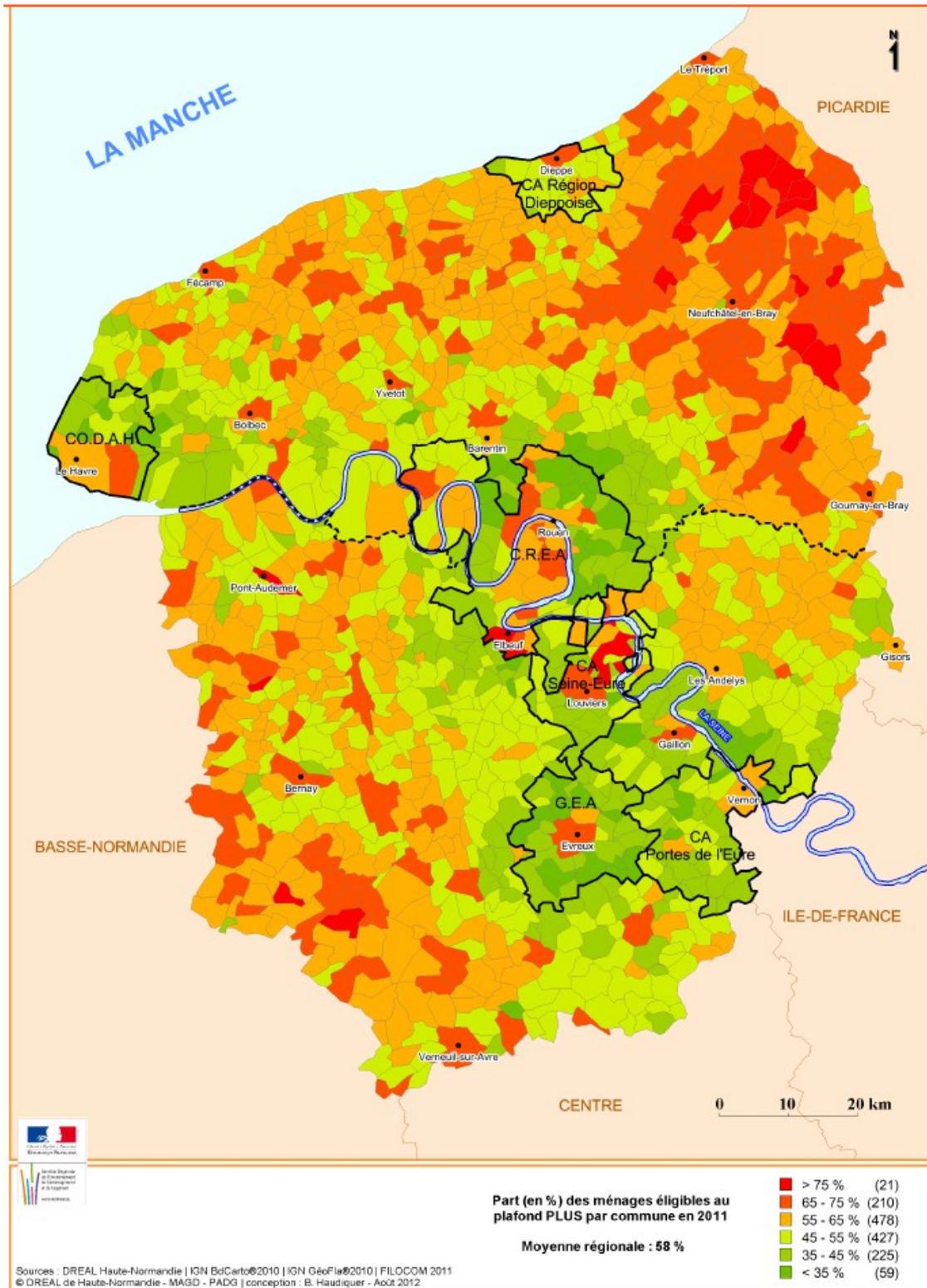
- **Publication : Premiers résultats – Basse-Normandie « La structure du parc de logement en 2011 » août 2012**

### La vacance



Sources : DREAL - Filocom 2011, d'après DGFIP - © IGN 2011

Part des ménages éligibles au plafond HLM au 1er janvier 2011



## Pour en savoir plus.....

.... liens vers les sites intranet et documents bibliographiques

### ◦ **Ministère de l'Écologie, du Développement Durable et de l'Énergie**

**CETE Nord-Picardie** Ce site propose notamment :

- **les téléchargements :**
  - **du Guide Filocom** – juillet 2001. Ce document fournit aux utilisateurs de Géokit un ensemble de définitions permettant une utilisation pertinente de FILOCOM et constitue une référence pour conduire des travaux détaillés à partir de la base de données individuelles dont dispose le MEEDDM/CGDD/SOeS. Ce guide présente l'intérêt de décrire chacun des champs de la base, ses limites et les utilisations qui peuvent en être faites.
  - **du dictionnaire des variables et données de cadrage 95-11.** Ce document (2012) est destiné aux utilisateurs avertis des fichiers Filocom. Il vient en complément du guide d'utilisation des fichiers Filocom édité en juillet 2001 par la DAEL : « Le fichier Filocom – Une base de données sur les logements et leur occupation ». Il est le recueil des évolutions du fichier depuis la première exploitation, il constitue donc l'historique des exploitations successives des fichiers Filocom.
  - **D'un rapport comparatif entre le recensement de la population et Filocom.** Rapport de décembre 2003 comparant des exploitations relatives à FILOCOM 1999 (informations au 1/1/1999) au recensement général de la population (RP) en date du mois de mars 1999. Ce travail décrit, mesure et explique, lorsque cela est possible, les écarts entre ces deux sources sur les mêmes variables.

**Commissariat général au développement durable** – page consacrée à Filocom

### **Géokit3**

- **Source Filocom / méthodologie et définition :** présentation de la source et documentation sur les variables
- **Organisation de l'univers Filocom non secrétisé**

### ◦ **Exemples d'exploitation**

#### **Publications du CGDD / Observation et statistiques**

<http://www.statistiques.developpement-durable.gouv.fr/>

- **Références / juin 2013** – L'état du logement en 2011
- **Le point sur... n° 170 / juillet 2013** – Une fois sur deux, la vacance d'un logement est suivie d'une réoccupation rapide et durable

#### **DREAL Basse-Normandie**

<http://www.basse-normandie.developpement-durable.gouv.fr/>

- Publication « premiers résultats » août 2012
  - La structure du parc de logements en 2011
  - Les conditions d'occupation des logements en 2011

#### **DREAL Haute-Normandie**

<http://www.haute-normandie.developpement-durable.gouv.fr/>

- Comité régional de l'habitat de Haute-Normandie rapport 2011

© 2014 Cerema

La reproduction totale ou partielle du document doit être soumise à l'accord préalable du Cerema.

