

Webinaire Datafoncier du 15 juin 2021

Réponses aux questions du chat



Sommaire

Données foncières	page 3
Lovac	page 6
Construction de la base de données nationales des bâtiments	page 10
Vers une meilleure connaissance des zones d'activités en Saine-et-Marne	page 12
Vers des zones d'activités plus économiques en foncières	page 13

Données foncières

Après dépôt d'une demande pour les fichiers fonciers en cours d'année, faut-il faire une demande séparée pour obtenir le RFP ou une réactualisation de la demande FF fonctionne ?

La procédure de rafraîchissement ne concerne que les données actualisées demandées lors de la demande initiale (datant de moins d'un an), sur le même périmètre géographique et pour les mêmes finalités.

Si vous souhaitez obtenir d'autres données, il vous faut refaire une demande initiale d'obtention.

Y a-t-il une procédure particulière pour restaurer les données LOVAC ou DV3F ?

Vous trouverez sur le site Datafoncier des aides à l'installation et à la restauration des données :

<https://datafoncier.cerema.fr/ressources/fichiers-fonciers>

Une [vidéo de démonstration](#) est également disponible.

Des tutoriels vidéos seront prochainement mis en ligne.

Quelles sont les conditions particulières d'utilisation et de diffusion des données pour des usages orientés à la recherche ?

Les données ne sont accessibles que pour les chercheurs, dans le cadre strict d'un projet de recherche.

Si les résultats de l'étude peuvent être diffusés (sous réserve de respecter l'anonymat des données), les données en elle-même ne peuvent pas être diffusées, ni redistribuées.

A-t-on des informations sur un éventuel retour des données qui ont disparu? (Identifiants mutation et locaux, infos vendeurs et acquéreurs)

Pas de nouvelles informations de la part de la DGFIP sur la restauration des données disparues depuis l'ouverture de DVF

Par contre, la donnée DVF devrait être accessible sur l'Alsace-Moselle à partir de 2022.

Quels sont les millésimes présents dans DV3F V5?

Les données DVF de 2010 à mi 2020 et les données Fichiers fonciers de 2009 à 2020.

Dans la V6 qui est en cours de production, seront présentes les données DVF de 2010 à fin 2020 et les données Fichiers fonciers de 2009 à 2020.

Sous quel délai la V6 de DV3F sera disponible ?

DV3F V6 devrait être disponible courant juillet.

Quelle est la fréquence de mise à jour de DV3F ?

Mise à jour semestrielle au même rythme que la livraison des données DVF.

La mise à jour peut parfois être un peu plus longue, dans l'attente de la sortie d'un nouveau millésime des Fichiers fonciers.

Quelle est la différence entre le Référentiel Foncier Présumé Public et les données des personnes morales mises à disposition par la DGFIP en open data depuis mars dernier ?

Le RFP contient les propriétés publiques avec une classification des acteurs publics qui permet de mieux les identifier. Ces classifications sont issues d'un retraitement et ne sont pas forcément présentes dans le fichier MAJIC personnes morales open data.

Les informations sur la caractérisation des biens sont aussi plus importantes dans le RFP.

Dans les données MAJIC personnes morales open data, il y a finalement peu d'informations sur les locaux et les parcelles.

De plus, si le RFP ne concerne que les terrains bâtis ou non bâtis de l'État et des établissements publics, les données MAJIC open data concernent le public et le privé.

es outils SIG. Ce sera un service complémentaire avec même processus d'obtention que GetDF

Y aura-t-il des stockages historiques de données à 10 / 20 ans ?

A ce jour, tous les millésimes Fichiers fonciers (depuis 2009) sont conservés pour répondre à certains usages tels que la mesure de la consommation d'espace ou la production de DVF. La date de conservation des données doit être rediscutée prochainement.

Y aura t-il une procédure particulière pour l'obtention des API données foncières ?

A ce jour, les API ne sont pas encore disponibles. Nous vous tiendrons informés via notre site Datafoncier dès cette procédure sera opérationnelle. Il ne devrait pas y avoir de procédure supplémentaire à celle actuellement en place pour bénéficier de ce service.

Y'aura t-il une contribution financière pour l'adhésion ?

Les données foncières sont diffusées gratuitement aux bénéficiaires. Aucune adhésion ni contribution n'est demandée.

Est-ce que les API données foncières permettront de récupérer via postgresql des objets géométriques avec les données attributaires associées sous Qgis?

Oui, il s'agira de flux qui permettront de récupérer du format geojson lisible via les outils SIG. Ce sera un service complémentaire avec même processus d'obtention que GetDF

Est-ce que les formations sont éligibles et inscrites au CFP (anciennement DIF droit individuel de formation) ?

Les formations Fichiers fonciers et DV3F ne sont pas éligibles au CFP.

Pour être éligibles au CFP, les formations doivent aboutir à :

- un diplôme, un certificat de qualification professionnelle (CQP) ou bloc de compétences, inscrits au RNCP (Répertoire national des certifications professionnelles),
- une certification ou habilitation enregistrée dans le répertoire spécifique des certifications et habilitations.

La formation doit de plus être datadocké (ce qui ouvre le financement par l'OPCO de la structure demandeuse) : le Cerema n'est pas datadocké. Le Cerema s'est plutôt engagé dans une démarche de certification Qualiopi, label qui remplacera Datadock au 1er janvier 2022.

Est-il envisageable de faire des formations à distance? Pour les DOM ?

Les formations Fichiers fonciers et DV3F se font pour l'instant en présentiel. Mais nous ne sommes pas fermés à proposer d'autres formats.

Une expérience en distanciel a été faite lors du confinement. Le distanciel est apparu plutôt adapté à la partie thématique.

Les échanges effectués sur la BAL fonctionnelle dafoncier@cerema.fr sont-ils accessibles (comme pour le labo.Qgis par exemple) ?

Non car il n'y a pas que des échanges techniques qui intéresseraient potentiellement tout le monde. Une FAQ va être mise en place sur la site Datafoncier, une partie des questions sera basée sur les questions fréquentes de la BAL.

Lovac

Quel est le délai moyen pour obtenir LOVAC après l'envoi de la demande ?

Suite à la mise en place d'une nouvelle procédure d'accès début mai, les délais de mise à disposition sont relativement courts : compter environ une semaine suite au dépôt de la demande.

Si le délai d'une semaine est passé et que nous n'avons pas de nouvelles pour Lovac, est-ce en raison d'un refus ?

Non, vous recevrez un message spécifique en cas de refus ou de problème spécifique lié à votre demande d'accès.

Les collectivités locales doivent-elles justifier l'utilité pour accéder aux données Lovac ?

Les collectivités signent un acte d'engagement qui comporte effectivement des mentions relatives à l'obligation d'utiliser ces données exclusivement dans le cadre d'une politique locale de lutte contre la vacance de logement (études et/ou démarches incitatives auprès des propriétaires de logements vacants).

En tant que Communauté de communes, est-il possible de refournir un traitement à partir de Lovac à une commune membre ?

Non, les communes de l'EPCI doivent de leur côté effectuer leur demande sur le site du CEREMA (<https://datafoncier.cerema.fr/lovac>)

Quels organismes sont ayants droits des données Lovac ?

Les organismes ayants-droit des données LOVAC sont les services de l'Etat, l'Agence nationale de l'habitat, les collectivités locales et leurs groupements dotés d'une fiscalité propre,

La liste des ayants-droits des données LOVAC est plus restreinte que pour les Fichiers Fonciers. Il s'agit des ayants-droits du fichier 1767BISCOM sur les logements vacants, qui constitue l'une des deux sources du fichier LOVAC.

Les autres organismes (agences d'urbanisme, établissements publics fonciers d'État et locaux, ADIL, bureaux d'études, SAFER, syndicats mixtes, ...) peuvent accéder aux données LOVAC dans le cadre d'une prestation réalisée pour le compte d'un ayant-droit.

Si les EPF sont éligibles, les données LOVAC peuvent-elles être transmises aux communautés de communes ou d'agglomération par les EPF pour contacter les propriétaires par exemple ?

Un EPF accéder aux données LOVAC dans le cadre d'une prestation réalisée pour le compte d'un ayant-droit, service de l'État ou collectivité territoriale et leurs groupements dotés d'une fiscalité propre.

Peut-on aussi avoir accès au fichier 1767-ResSec ?

Le fichier des résidences secondaires est communicable aux communes sur leur demande auprès de leur Service de Fiscalité Directe Locale (SFDL).

Existe-t-il un dictionnaire de données pour la partie 1767 BISCOM des données LOVAC ? A quelle échéance la publication du dictionnaire est-elle prévue ?

Un dictionnaire en ligne spécifique est en cours de construction et va être mis en ligne très rapidement.

Le fichier LOVAC est plus propre que le croisement simple ?

LOVAC 2020 contient l'intégralité des données du 1767 bis com 2020 enrichies avec les fichiers fonciers 2020 (ou 2019 si local non présent dans FF2020). Chaque donnée source est respectée et il n'y a aucun redressement. Par contre en cas de différences entre les deux sources (propriétaire, vacance) cela peut nous alerter sur leur fiabilité et inciter à la prudence quant à leur utilisation.

Y aura t il des formations LOVAC ?

Pas encore, mais c'est envisagé dans les prochains mois.

Quel est le nom des 4 tables Lovac ?

- lovac_exh : logements et locaux (dépendances) vacants, toutes durées confondues : ce fichier permet d'accéder à l'intégralité du fichier 1767BISCOM sur les logements vacants croisé avec les Fichiers Fonciers.
- lovac_fil : logements vacants du parc privé vacants depuis au moins deux ans : c'est ce jeu de données qu'il est recommandé d'utiliser en priorité. Il permet de repérer et caractériser la vacance structurelle du parc privé de logements, cible du plan national de lutte contre les logements vacants
- extff_2020_stat : données agrégées à la commune sur l'ensemble du parc de logements
- extff_2020 : ce fichier permet d'accéder à l'ensemble du parc de logements d'un territoire. Cela permet notamment de comparer le profil des logements vacants à celui de l'ensemble du parc de logements.

Le contenu de la livraison LOVAC Postgres.sql est différent de la livraison SHP (1table en moins) ?**Pourquoi y a-t-il 3 tables postgis contre 4 pour les autres formats de données ?**

La table lovac_fi n'est pas présente dans le dump Postgresql, car la puissance de Postgresql permet de la reconstituer en moins de 10 seconde au niveau national avec la requête suivante (pour le millésime 2020) :

SELECT *

FROM lovac.log_vac2020_ff19_20_ff2

WHERE to_number(debutvacance,'9999')<=2017 AND groupe NOT IN ('1','2','3','4','5','6','9') AND aff='H'
AND nature in ('APPART','MAISON') AND ff_ccthp='V'

Faites-vous, comme l'INSEE, une différence avec les résidences secondaires ?

Les définitions de la vacance et des résidences secondaires sont légèrement différentes de celles utilisées dans les recensements de population.

D'après le recensement :

- Un logement est vacant s'il est inoccupé et :
 - proposé à la vente, à la location,
 - déjà attribué à un acheteur ou un locataire et en attente d'occupation,
 - en attente de règlement de succession,
 - conservé par un employeur pour un usage futur au profit d'un de ses employés,
 - sans affectation précise par le propriétaire (logement vétuste, etc.).
- Une résidence secondaire est un logement utilisé pour des séjours de courte durée (week-ends, loisirs, ou vacances). Les logements meublés mis en location pour des séjours touristiques sont également classés en résidences secondaires.

D'après les sources fiscales :

- les logements vacants sont vides de meubles et libres de toute occupation pendant au moins une année au 1er janvier de l'année d'imposition. Une occupation de moins de 90 jours consécutifs au cours de l'année ne remet pas en cause la vacance fiscale.
- les résidences secondaires sont des locaux meublés, affectés à l'habitation, non occupés en permanence, et soumis à la taxe d'habitation.

La vacance est calculée sur l'acquittement de la taxe d'habitation. Quelle perspective de calcul de la vacance avec la suppression de la taxe d'habitation ?

La taxe d'habitation sur les résidences principales va disparaître mais les taxes sur les logements logements vacants (TLV et THLV) et résidences secondaires (THRS) seront maintenues.

Avez-vous des retours d'utilisation ? Des retours d'études ?

Nous disposons d'un nombre croissant de retours d'utilisations qui seront prochainement partagés sur la page dédiée aux données LOVAC sur le site du CEREMA. Ces études sont également partagées via la plateforme d'échanges sur les logements vacants. Les demandes d'accès à cette plate-forme sont à adresser à : logement-vacant.dhup.dgaln@developpement-durable.gouv.fr

A quoi correspond la base 1767 BISCOM (taxe d'habitation) ?

Ce fichier fiscal recense l'ensemble des logements vacants et précise notamment la durée de vacance, ce qui permet de distinguer la vacance de longue durée dite structurelle, de la vacance de courte durée, dite frictionnelle.

Le fichier LOVAC croise les informations du fichier 1767BISCOM avec celles des Fichiers Fonciers (caractéristiques des logements et du propriétaire).

Quelle est la différence entre les logements ccthp2019=V(ff20) et ccthp2018=V(ff2019) et ccthp2017=V(ff2020) et les logements de lovac2020 vacants depuis au moins 2 ans?

L'expérience a montré que la variable ccthp des fichiers fonciers correspondait à l'occupation du millésime N-1, contrairement à d'autres sources comme Filocom par exemple.

La requête suivante :

```
SELECT debutvacance,ff_ccthp,count(*) as nb
FROM a0000_logements_vacants.log_vac2020_ff19_20_ff2
WHERE debutvacance in ('2020','2019') and nature in ('APPART','MAISON')
GROUP BY debutvacance,ff_ccthp
ORDER BY debutvacance desc,ff_ccthp
```

debutvacance	ff_ccthp	nb
2020	G	21
2020	L	1243
2020	N	4141
2020	P	248
2020	T	10166
2020	V	21
2020		272568
2019	B	4
2019	G	576
2019	L	13298
2019	N	5139
2019	P	5726
2019	T	18941
2019	V	1615092
2019	X	1
2019		21874

Donne un résultat qui laisse penser au même décalage entre FF2020 et le 1767biscom 2020.

Le fichier LOVAC logement privé vacant depuis plus de deux ans est basé sur la variable debutvacance du 1767biscom et en plus sur ff_ccthp='V', car selon les retours terrain cela permet de retirer de nombreux logements non vacants ou vacants mais non mobilisables.

Quelle est la différence entre ccthp2019 des FF2020 et lovac2020 (Vacant N-1) ?

La variable CCTHP (code occupation à la taxe d'habitation) est issue des Fichiers Fonciers. Elle décrit l'occupation du bien au 1^{er} janvier N-1.

Comment est constitué la donnée lorsqu'il n'y a pas de taxation des logements vacants ?

Le fichier LOVAC recense tous les logements vacants, qu'ils soient taxés ou non sur la vacance.

Pourrait-on disposer des variables agrégées des niveaux régionaux et nationaux afin de pouvoir comparer avec les niveaux départementaux et infra ?

Les données agrégées à des échelons régionaux et nationaux devraient être prochainement publiées.

Existe-t-il l'équivalent Lovac pour la vacance des locaux d'activité ?

A notre connaissance, il n'y a pas d'équivalent sur les locaux d'activité. Des observatoires locaux existent mais pas encore de démarche nationale.

Une piste potentielle pourrait être d'exploiter les données de la cotisation foncière des entreprises.

Est-il possible d'appréhender la vacance par copropriété avec les seules données Lovac ?

Le registre national des copropriétés dispose d'identifiants, mais effectivement les identifiants uniques de bâtiment sont encore un sujet de recherche/analyse. Le CSTB y travaille. Mais pas seul

Construction de la base de données nationale des bâtiments

Qui seront les ayants droits ?

Par principe, un maximum de donnée sur les bâtiments sera disponible en OpenData, mais pour certain attributs (ex : ceux issus des Fichiers Fonciers), nous devons respecter les règles de diffusions imposées par les producteurs de données.

A quelle échéance les données seront sur data.gouv ?

Le périmètre des données est en cours de discussion.

Le code source est-il ouvert ?

Un travail avec Etalab est en cours pour identifier les codes avec le meilleur potentiel de réutilisation.

Quand l'appli GoRenove sera-t-elle disponible à la consultation ?

L'application GORENOVE rentre en service au 1er janvier 2022.

Quels sont les logiciels utilisés pour les traitements ? Quel matériel informatique ?

Pour les logiciels, c'est du PostgreSQL+PostGIS pour les opérations de croisement géospatial, ainsi que du code Python pour les éléments de workflow et d'optimisation combinatoire non-gérables par un système de gestion de bases de données

Et pour le matériel informatique : un serveur de calcul de 48 cœurs et 188 Go de RAM, sous OS Linux.

Pour les autres types d'opérations (prédictions, traitements des adjacences, masques solaires, calculs énergétiques), c'est Spark + Python.

L'Alsace sera-t-elle dispo ?

Oui, l'alsace est dans le périmètre. Mais nous ne traitons que les bâtiments/parcelles/TUP ayant des géométries vectorielles « non-fictives ». Si la source est complète, pas de problème. Si l'informations géométriques récupérées auprès de la BDTopo et des Fichiers Fonciers sont fictives ou inexistantes, non.

La mise en œuvre de la directive bâtiments tertiaires est bien l'un des enjeux de cette BDD ?

Pour le moment, la BDNB vise essentiellement le périmètre fixé dans le cadre du projet Go-Rénove (financement PROFEEL). Ce périmètre cible le « bâti résidentiel ». Dans notre workflow actuel, nous n'avons pas exclu les bâtiments « non-résidentiels », mais une grande marge d'amélioration existe pour ce segment de parc. Dans un futur proche, nous espérons pouvoir développer ce segment dans le BDND. Toutes propositions/collaborations sont les bienvenues.

Une barre d'immeuble qui aurait plusieurs adresses et plusieurs entrées est-elle considérée comme un bâtiment ou comme plusieurs bâtiments comme dans Majic ?

La réponse apportée ci-après cible la « BDNB » et non « Majic ».

Dans la BDNB un bâtiment peut avoir plusieurs adresses. Et une adresse peut desservir plusieurs bâtiments.

Nous avons développé des algorithmes d'attributions de données au prorata du nombre de locaux et/ou surfaces déclarés dans la Fichiers Fonciers (FF). Pour rappel, dans les FF, les entités « bâtiment » (idbat) sont constituées d'un ou plusieurs « local ». Chaque local dispose d'une « adresse postale ».

Sera-t-il possible d'identifier, dans une parcelle où il y a plusieurs bâtiments, le bâtiment d'habitation ?

C'est précisément sur ce type de métrique évaluant la qualité des croisements géométriques que nous travaillons aujourd'hui. L'apport de la base de données Arcep/HD-THD nous aide à correctement identifier les bâtiments d'habitation sur les parcelles multi-bâtiments.

Vous prenez en compte la notion de parcelles référentes qui sont souvent sans bâti mais qui récupèrent les attributs des locaux ?

Nous utilisons les parcelles cadastrales, mais surtout les TUP qui englobe toutes les parcelles adjacentes appartenant à un même propriétaire. Cela permet de simplifier les configurations de propriétés multi-parcelles.

L'emprise au sol des FF correspond ici à dcnt13 ou dcntarti?

Du point de vue géométrie côté FF, nous n'utilisons que le champ « geompar » de la table des TUP. Pour la BDNB, les seules emprises au sol qui nous intéressent sont celles des bâtiments. Nous les récupérons de la BDTopo et d'OSM.

Quels sont les coefficients de détermination de ces régressions linéaires ?

Il ne s'agit pas vraiment de régression. Ici, il s'agit de comprendre les relations entre les données de surfaces/ hauteurs entre la BD Topo et les FF sur les TUP simples, de manière à alimenter les algorithmes d'optimisations. Il s'agit de construire les règles métiers. Des coefficients sont calculés par groupes d'usages de bâtiments, mais nous ne les injectons pas directement dans le solveur combinatoire. La question posée est trop technique pour qu'une réponse complète puisse y être apporté par ce type de communiqué.

Une approche différenciée locale/territoriale permettrait elle d'augmenter le pourcentage d'appareillage ?

L'approche locale permettrait d'aller plus loin. C'est une question de rapport effort/bénéfice. Nous privilégions actuellement l'échelle nationale, mais ne sommes pas fermé à de futurs partenariats avec des collectivités.

Est-il possible d'avoir sur un département la correspondance des identifiants bd topo et Fichiers fonciers ?

Nous construisons effectivement cette liste de correspondance, peu importe l'échelle géographique (nationale, départementale, IRIS...). Le périmètre des données qui seront publiées en OpenData est en cours de discussion.

Y a-t-il un croisement avec le Référentiel Technique qui est une base de donnée des implantations de l'Etat ?

La base de données se concentre pour le moment sur les bâtiments de logements et tertiaires privés mais la base de données devrait être étendue aux bâtiments publics en 2022.

A propos de la BD nationale des bâtiments, avez-vous utilisé le répertoire des immeubles localisés de l'INSEE ?

Non, pas à ce jour.

Vers une meilleure connaissance des zones d'activités en Seine-et-Marne

Est-ce que la méthode est disponible ?

La DDT 77 communiquerait-elle la méthode aux autres DDT ?

Le support de présentation sera mis en ligne. Nous prendrons attache avec la DDT77 pour la question de la diffusion.

Concernant la vacance des locaux d'activités, les fichiers fonciers ne semblent pas fiables (problème de qualité de la base de la DGFIP) ?

Effectivement, les données FF sur la vacance (champ DNATLC) sont peu fiables. Mais il s'agit toujours, en l'absence de mieux, d'une donnée de contexte.

De quelle manière (les EPCI) se sont saisies de ces résultats ?

L'étude a été sollicitée par la DDT77. Les résultats ont permis d'enrichir la préparation des réunions avec les EPCI, mais nous n'avons pas de retour plus précis à ce sujet.

Quel champ utilisez-vous pour qualifier les types d'activités en présence ?

Nous avons utilisé conjointement les champs *cconlc* (disponibles depuis 2009) et *typeact* (dérivé du champ *ccocac* de la table pb21, disponible depuis 2017).

Dans le cas de bâtiments commerciaux, disposez vous de données sur les surfaces de vente ?

Nous n'avons pas effectué ce recensement.

Les fichiers LOCOMVAC recensant les locaux commerciaux et professionnels vacants au titre de la CFE de l'année précédente sont à investiguer ?

En effet. A condition d'y avoir accès. A ce stade, cela peut être intéressant, mais nous n'avons jamais eu l'opportunité de voir leur contenu et de déterminer leur fiabilité.

Vers des zones d'activités plus économiques en fonciers

Comment repérer actuellement les surfaces imperméables (bâti+revêtues)/ perméables quand on n'a pas l'OCSGE dispo?

On patiente jusqu'en 2022-2023 pour avoir une couverture nationale par l'OCSGE. A moins qu'il existe un Mode d'Occupation du Sol sur le territoire.

Comment avez-vous localisé les effectifs salariés (source ?) ?

Comment avez-vous estimé les emplois (pas toujours renseigné dans Sirene) ?

Fichier SIRENE opendata géocodé

http://data.cquest.org/geo_sirene/quotidien/

Pour chaque établissement, il est indiqué une tranche d'effectif salarié.

Nous sommes contraints de faire des approximations car la fourchette est parfois large (par ex [200-500], une moyenne est alors appliquée, puis vérifiée : en sommant par commune, si à peu près sur le chiffre de l'INSEE.

Dans l'Ain il existe un « réseau foncier » qui a fait un travail plus précis d'inventaire des ZAE, et qui a corrigé les positions et les données d'emploi des entreprises. On a donc une donnée d'emploi par ZAE plus fiable. J'ai pu profiter de ce travail disponible en open data :

<https://departement-ain.opendata.arcgis.com/search?tags=%C3%A9conomie>

Quelle est la source des données retombées fiscales ?

Comment avez-vous localisé la CVAE, nous n'avons que le SIREN pour le réaliser ?

Il nous a été mis à disposition par la collectivité (Clermont Métropole)

Les produits des différentes taxes à l'échelle de chaque collectivité sont en open data

<https://www.data.gouv.fr/fr/datasets/impots-locaux/>

Sur Clermont on avait le détail de ces produits pour chaque ZAE.