



Méthodologie d'analyse des données OPERAT

Version 2024.1

Sommaire

1	Méthodologie de fiabilisation des données	3
1.1	Présentation des données.....	3
1.2	Termes et abréviations utilisés.....	3
1.3	Scores de fiabilité	3
1.4	Analyse de fiabilité.....	5
2	Utilisation des scores de fiabilité	6
2.1	Répartition des consommations et des surfaces déclarées.....	6
2.1.1	Répartition par année de consommation	6
2.1.2	Répartition géographique par département.....	7
2.1.3	Répartition par (méta)catégorie d'activité.....	7
2.2	Evolution à périmètre constant.....	7
2.3	Consommation d'énergie finale surfacique	8
	Annexe 1 : Détail des champs de données.....	9
	Annexe 2 : Détail du calcul des scores de fiabilité	11
	Annexe 3 : Catégories d'activité.....	16



La présente note détaille la méthodologie utilisée pour la première analyse des données OPERAT publiée par l'ADEME : [Performance énergétique du parc tertiaire : quel bilan de l'utilisation de la plateforme OPERAT en 2022-2023 ?](#)

Contact : Roméo Juge, Chef de projet, OID (romeo.juge@o-immobilierdurable.fr)

1 Méthodologie de fiabilisation des données

1.1 Présentation des données

L'analyse présentée ici porte sur les **données de la base OPERAT au 31/01/2024**. Chaque ligne correspond à une EFA et une année de consommation (année de référence, 2020, 2021, 2022) données. Les champs de données (les colonnes) sont détaillés dans l'[Annexe 1 : Détail des champs de données](#).

Au 31 janvier 2024, 739 700 déclarations avaient été effectuées sur OPERAT, représentant 235 483 EFA distinctes.

1.2 Termes et abréviations utilisés

- Surface moyenne annuelle non vacante : Surface moyenne annuelle – Surface moyenne annuel local vacant ;
- Période d'activité principale : [date de début d'activité ; date de fin d'activité] de la sous-catégorie d'activité majoritaire ;
- C_{EF} : consommation surfacique annuelle en énergie finale (EF) PCI, toutes énergies confondues ($kWh_{EF}/m^2/an$), obtenue en divisant la consommation totale en EF PCI par la surface non vacante ;
- (S)CAP-**M** : (sous)-catégorie d'activité **majoritaire** au sens de la surface moyenne annuelle ;
- (S)CAP-**S** : (sous)-catégorie d'activité **secondaire** au sens de la surface moyenne annuelle.

1.3 Scores de fiabilité

Afin de quantifier la fiabilité des données, dans le but de filtrer les déclarations considérées comme incomplètes ou erronées permettant ainsi de construire des échantillons représentatifs, des scores de fiabilité ont été élaborés. Ceux-ci sont adaptés des scores développés par l'Observatoire de l'Immobilier Durable dans le [Baromètre de la performance énergétique et environnementale des bâtiments \(BPE\)](#).

Ces scores sont calculés pour chaque déclaration de la base, c'est-à-dire pour chaque EFA et chaque année de consommation. Deux scores sont calculés, mesurant chacun la fiabilité de caractéristiques distinctes, bien qu'interdépendantes :

- Le score **PS**, compris entre 0 et 9, mesure la fiabilité des données relatives au **périmètre** de déclaration et à la **surface** ;
- Le score **E**, compris entre 0 et 4, mesure la fiabilité des données de **consommation énergétique**.

Ces scores se présentent sous forme d'échelles. Le calcul d'un score consiste à faire passer plusieurs tests successifs à une ligne. Chaque score commence à 0 et, pour atteindre un score donné, une ligne doit passer tous les tests antérieurs. Par exemple, pour qu'une ligne obtienne le score PS4, elle doit passer les tests PS1, PS2, PS3 et PS4.

Une description succincte des tests pour chaque score est donnée ci-dessous. Le détail du calcul des scores est quant à lui donné dans l'[Annexe 2 : Détail du calcul des scores de fiabilité](#).

Enfin, plusieurs champs de données, générés lors du calcul des scores, sont ajoutés à la base initiale. Ils sont détaillés dans l'[Annexe 1 : Détail des champs de données](#).

Le score **E** permet de conserver les déclarations pour lesquelles :

1. La consommation totale en énergie finale (kWh PCI) est strictement positive et $PS \geq 5$.
2. La consommation surfacique en énergie finale, C_{EF} (kWh/m²/an), qui peut désormais être calculée, est comprise à l'intérieur des bornes définies par le 1^{er} filtre (filtre physique).
3. C_{EF} est comprise à l'intérieur des bornes définies par le 2^{ème} filtre (filtre statistique), plus restrictif et qui dépend de la distribution des C_{EF} déclarées sur la même catégorie d'activité.
4. C_{EF} est stabilisée, c'est-à-dire les déclarations pour lesquelles le périmètre est stabilisé ($PS \geq 8$), et C_{EF} n'a pas trop varié par rapport à la déclaration de l'année précédente. Pour l'année de référence et pour 2020, toutes les lignes passent automatiquement le test.

Les différents tests des scores de fiabilité peuvent être modulés en fonction du niveau d'exigence requis. A titre d'exemple, les bornes (en surface ou en énergie) peuvent être modifiées, la variation maximale d'une surface ou d'une consommation d'une année sur l'autre peut également être ajustée.



Certains tests (PS8, E2, E3, E4) sont rétroactifs, au sens où l'atteinte d'un palier par une ligne dépend des consommations de l'EFA sur les années précédentes ou des consommations des autres EFA appartenant à la même catégorie. Il est donc important, à chaque nouvelle version de la base, de recalculer ces scores de fiabilité sur la totalité de la base de données.

1.4 Analyse de fiabilité

Des scores de fiabilité ont été calculés pour chaque ligne de la base de données transmises au 31/01/2024. Le tableau ci-dessous donne le pourcentage de déclarations restantes après application des différents tests. Par exemple, 82,5% des lignes obtiennent un score $PS \geq 7$ et $E \geq 3$, et 71,1% des lignes obtiennent le score maximal $PS = 9$ et $E = 4$.

Lignes restantes (%)						
PS \ E	0	1	2	3	4	
0	100.0					
1	97.2					
2	97.0					
3	96.4					
4	95.7					
5	94.3	92.2	89.2	88.0		
6	94.2	92.0	89.1	87.9		
7	88.0	86.2	83.5	82.5		
8	83.4	81.9	79.5	78.6		76.0
9	77.9	76.5	74.3	73.5		71.1

En fonction des analyses effectuées, différents niveaux de fiabilité (i.e. scores PS/E) pourront être sélectionnés.

2 Utilisation des scores de fiabilité

Cette section détaille la méthodologie utilisée pour les 3 types d'analyses présentées dans le rapport :

- Répartition des consommations et surfaces déclarées
- Evolutions à périmètre constant
- Indicateurs de consommation surfaciques par (méta)catégorie d'activité

2.1 Répartition des consommations et des surfaces déclarées

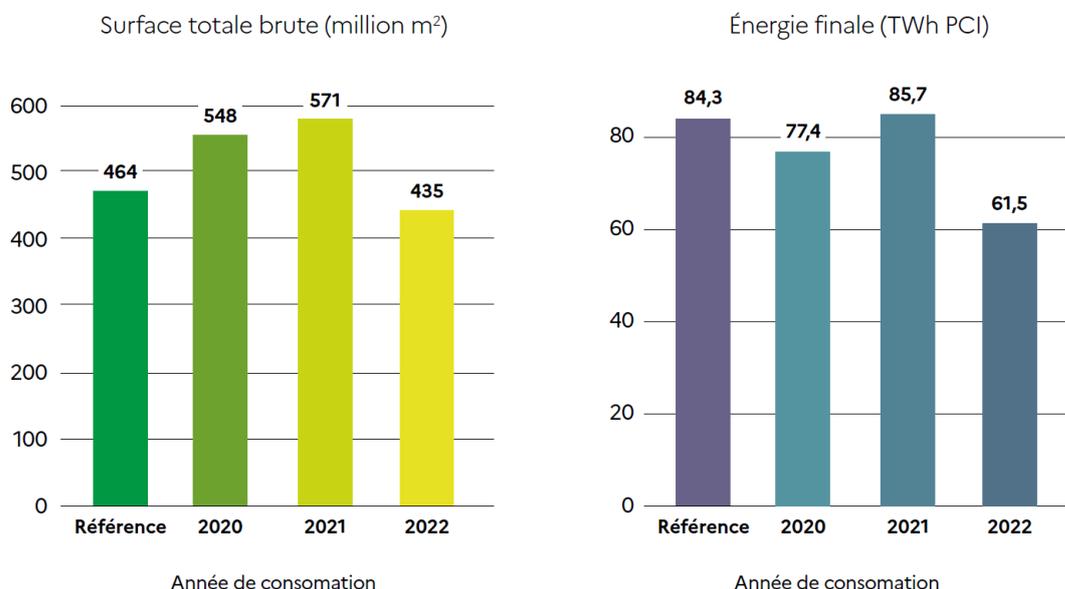
On détaille ici les analyses portant sur la répartition de la surface (millions de m²) et de la consommation finale totale (TWh PCI) déclarées par année de consommation (pp. 9, 11, 13 de l'étude), par département (p. 14), par (méta)catégorie d'activité (pp. 15-17).

On sélectionne pour cela les **82,5%** de déclarations avec **PS ≥ 7** et **E ≥ 3**. Ce faisant, on conserve les déclarations dont les surfaces et/ou la consommation d'énergie a évolué de manière significative entre deux années successives ainsi que les déclarations ne commençant pas au 1^{er} janvier de l'année. Cela permet d'obtenir un échantillon fiable aussi large que possible sur lequel effectuer des analyses macroscopiques. L'évolution du nombre de déclarations, de la surface totale brute et de la consommation totale en énergie finale (EF) en fonction de l'année de consommation sont données ci-dessous.

2.1.1 Répartition par année de consommation

La surface totale brute et la consommation finale totale restantes après application des filtres, pour les différentes années de consommation, est reproduite ci-dessous

Évolution des surfaces et consommations totales déclarées sur les différentes années



Source : [Performance énergétique du parc tertiaire : quel bilan de l'utilisation de la plateforme OPERAT en 2022-2023 ?](#) [p.9]

La répartition par année de consommation de la consommation par type d'énergie (pp. 11, 13) s'obtient simplement en sommant la consommation totale (en TWh PCI) pour chacune de ces sources.

2.1.2 Répartition géographique par département

La répartition géographique de cette même surface et de cette même consommation (p.14 de l'étude) est calculée sur les données de consommation 2021, en sommant la surface et la consommation déclarées sur chaque département.

2.1.3 Répartition par (méta)catégorie d'activité

Afin d'obtenir la répartition de la surface et de la consommation totales déclarées par (méta)catégorie d'activité, des retraitements supplémentaires sont nécessaires.

Pour les **EFA composées de plusieurs catégories d'activité**, on renseigne la surface de la catégorie majoritaire et celle de la catégorie secondaire dans leur catégorie respective. Toute catégorie additionnelle composant l'EFA (pour les EFA composée de plus de 2 catégories) sera comptée dans la catégorie majoritaire.¹ A noter qu'une très large majorité des déclarations (93%) ont une (86%) ou deux (7%) catégories d'activité renseignées.

Exemple : pour une EFA de 1 000 m² composée de 800 m² de bureaux (catégorie majoritaire), 150 m² de commerces (catégorie secondaire) et 50 m² de salles serveurs, on comptera 850 m² de bureaux et 150 m² de commerces.

Concernant la consommation d'énergie, il n'est pas possible de séparer les consommations des différentes catégories d'activité composant l'EFA. On attribuera donc les consommations au prorata de la surface occupée par la catégorie d'activité au sein de l'EFA, en comptant les consommations des locaux vacants comme nulle.

Exemple : pour une EFA de 1 000 m², avec une consommation totale annuelle de 100 000 kWh PCI, composée de 700 m² de bureaux (catégorie majoritaire), 150 m² de commerces (catégorie secondaire), 100 m² de locaux vides et 50 m² de salles serveurs :

- Consommation bureaux = $(700 + 50) / (1000 - 100) * 100\ 000 = 83\ 333$ kWh
- Consommation commerces = $150 / (1000 - 100) * 100\ 000 = 16\ 667$ kWh

On regroupe ensuite les EFA en **métacatégories**. Ces-dernières sont définies dans l'[Annexe 3 : Catégories d'activité](#).

2.2 Evolution à périmètre constant

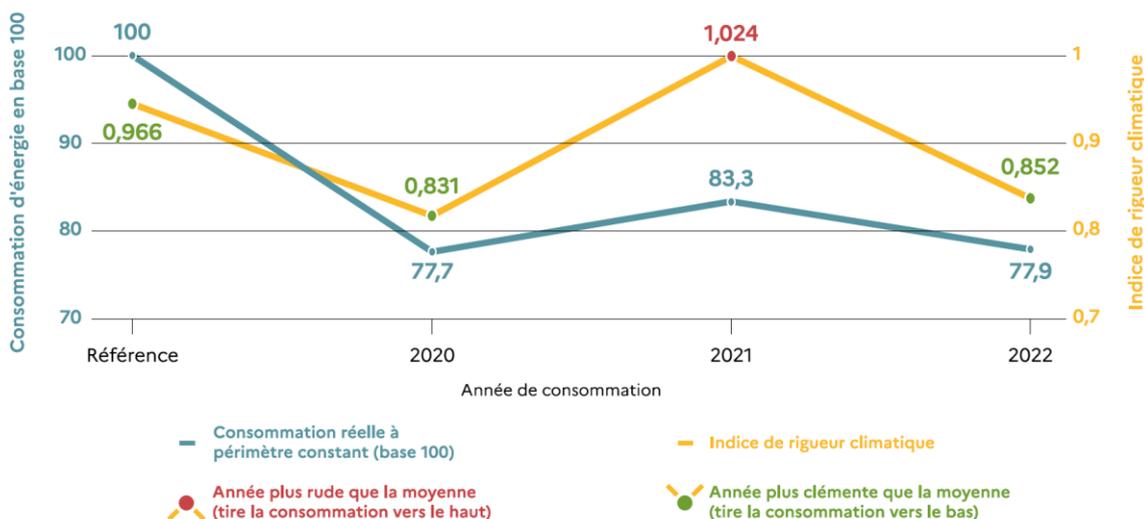
Pour donner une **évolution annuelle** globale de la consommation d'énergie de la base, on sélectionne les lignes ayant obtenu le score maximum (**PS = 9** et **E = 4**), afin de ne conserver que les déclarations dont la surface et la consommation d'énergie sont stabilisées² et commençant au 1^{er} janvier, afin que toutes les consommations d'une année donnée correspondent à la même période d'activité principale, ce qui représente **71,1%** des lignes de la base.

On utilise ensuite la méthode de calcul utilisée par le SDES dans la [bilan énergétique de la France](#) : pour chaque année de consommation N, année de référence incluse, on calcule la variation relative par rapport à l'année N-1 à périmètre constant (i.e. sur les EFA déclarées à la fois en années N-1 et N), en fixant la consommation de l'année de référence à 100, ce qui permet de tracer l'évolution relative de la consommation d'énergie du parc, malgré l'évolution du périmètre d'une année sur l'autre.

¹ Seules les deux catégories occupant la plus grande surface (catégories majoritaire et secondaire) étaient disponibles dans les données analysées.

² Voir tests PS8 et E4.

Évolution des consommations à périmètre constant (en base 100 sur la période de référence 2010-2019) et indice de rigueur climatique



Source : [Performance énergétique du parc tertiaire : quel bilan de l'utilisation de la plateforme OPERAT en 2022-2023 ?](#) [p.10]

Pour obtenir l'évolution à périmètre constant de la part relative de chaque énergie, on effectue un calcul similaire sur chaque source d'énergie.

2.3 Consommation d'énergie finale surfacique

Le dernier type d'analyse effectué concerne le calcul des ratios de consommation surfaciques (en kWh/m²/an) moyens par (méta)catégorie d'activité (pp. 18-20).

L'objectif est ici de construire, pour chaque (méta)catégorie d'activité et chaque année de consommation, un échantillon d'EFA représentatif de sa (méta)catégorie. Pour cela, le périmètre est resserré et des retraitements supplémentaires sont effectués (le % de **déclarations restantes** après chaque étape est indiqué entre parenthèses) :

- Sont d'abord sélectionnés les déclarations avec **PS ≥ 9** et **E ≥ 3**. Ce faisant, ne sont conservées que les déclarations dont le périmètre est stabilisé et dont la période d'activité débute au 1^{er} janvier, afin de disposer de données comparables entre deux années consécutives. Sont néanmoins conservées les déclarations dont la consommation a évolué significativement entre deux années consécutives, étant donné le contexte sanitaire **(73%)** ;
- Enfin, sont sélectionnées les déclarations dont la catégorie d'activité majoritaire (CAP-M) représente plus de **90%** de la surface totale, afin de ne conserver que les EFA représentatives de leur catégorie d'activité **(52%)**.

Enfin, pour chaque (méta)catégorie, on calcule la moyenne des ratios surfaciques.

Annexe 1 : Détail des champs de données

Le tableau ci-dessous détaille les champs de données (i.e. les colonnes) de la base OPERAT analysée. Les champs générés à l'issue du calcul des scores de fiabilité sont donnés en bas du tableau.

Intitulé	Type	Description
id	int	Identifiant de la déclaration
Annee Consommation	str	Année de consommation
Date de début de la déclaration	date	Date de début de déclaration des consommations énergétiques de la CAP-M
ID interne EFA	int	Identifiant interne de l'EFA (anonymisé)
IUBs EFA	str	Identifiant unique bâtiminaire de l'EFA
Statut	str	Statut de la déclaration (« Validée » ou « Prête ou à compléter »)
Commune	str	-
Departement	str	-
Region	str	-
Station meteo	str	Définies dans le décret tertiaire.
Zone climatique	str	
Classe altitude	str	
Categorie d'activite majoritaire	str	CAP majoritaire (CAP-M) au sens de la surface moyenne annuelle
Surface moyenne annuelle catégorie d'activité majoritaire (m²)	float	Surface moyenne annuelle de la CAP-M
Categorie d'activite secondaire	str	CAP secondaire au sens de la surface moyenne annuelle
Surface moyenne annuelle catégorie d'activité secondaire (m²)	float	Surface moyenne annuelle de la CAP-S
Surface totale brute (m²)	float	Surface brute de l'EFA renseignée par l'assujetti (somme des surfaces de toutes les SCAP déclarées)
Surface moyenne annuelle totale chauffée	float	-
Surface moyenne annuelle totale refroidie	float	-
Surface moyenne annuelle totale froid process	float	-
Categorie d'activite principale	str	Champ facultatif renseigné par l'assujetti, servant à des fins purement statistiques. Non utilisé dans la présente analyse.
Sous Categorie d'activite majoritaire	str	SCAP majoritaire au sens de la surface moyenne annuelle
Surface moyenne annuelle sous-catégorie d'activité majoritaire (m²)	float	Surface moyenne annuelle correspondante
Sous-catégorie d'activité secondaire	str	SCAP secondaire au sens de la surface moyenne annuelle
Surface moyenne annuelle sous-catégorie d'activité secondaire (m²)	float	Surface moyenne annuelle correspondante
Dont Data Center	str	'OUI' si au moins une des SCAP déclarée appartient à la CAP data center
Nb sous categories distinctes	int	
Nb categories distinctes	int	Nombre de CAP déclarées par l'assujetti
Surface moyenne annuelle	float	
Consommation totale (kWh PCI)	float	Somme des consommations des différentes énergies exprimée en kWh d'énergie finale PCI
Ratio (kWh PCI / m2)	float	Consommation totale en EF PCI divisée par la surface moyenne annuelle
Surface moyenne annuel local vacant (m²)	float	Lorsqu'un local vacant est déclaré l'EFA par l'assujetti
Surface moyenne annuelle catégorie "Stationnement" (m²)	float	-
Electricite (kWh)	float	-
Gaz naturel - reseaux (kWh)	float	-
Gaz naturel liquefie (kg) - kWh	float	-
Gaz propane (kg) - kWh	float	-
Gaz propane (m3) - kWh	float	-
Gaz butane (kg) - kWh	float	-

Gaz butane (m3) - kWh	float	-
Fioul domestique (l) - kWh	float	-
Charbon - agglomerés et briquettes (kg)- kWh	float	-
Houille (kg) - kWh	float	-
Bois - Plaquettes d'industrie (kg) - kWh	float	-
Bois - Plaquettes forestières (kg) - kWh	float	-
Bois - Granules (pellets) ou briquettes (kg) - kWh	float	-
Bois - Buches (stère)-kWh	float	-
Reseau de chaleur (kWh)	float	-
Reseau de froid (kWh)	float	-
Consommation usages thermiques connues	str	« OUI »/ « NON »
Consommation usages chauffage (kWh PCI)	float	-
Consommation usages refroidissement (kWh PCI)	float	-
Consommation usages froid process (kWh PCI)	float	-
Consommation totale IRVE	float	IRVE = Infrastructure de Recharge de Véhicule Électrique
NB Bornes recharge IRVE	int	-
Code CAP-M	str	Liste de codes donnés dans la V5.0 de l'outil d'import csv des consommations
Code CAP-S	str	
P&S Score	int	Score PS obtenu
S_inf (m²)	float	Borne surfacique inférieure du filtre PS7
S_sup (m²)	float	Borne surfacique supérieure du filtre PS7
E Score	int	Score E obtenu
Ratio (kWh PCI / m2) – Corrigé	float	C _{EF} calculée à l'issue du test E1 sur la surface non vacante
E_inf abs (kWh/m².an)	float	Borne énergie inférieure du test E2
E_sup abs (kWh/m².an)	float	Borne énergie supérieure du test E2
E_inf stat (kWh/m².an)	float	Borne énergie inférieure du test E3
E_sup stat (kWh/m².an)	float	Borne énergie supérieure du test E3

Nouveaux champs de données générés à l'issue du calcul des scores de fiabilité

Annexe 2 : Détail du calcul des scores de fiabilité

Score PS

1

La surface totale brute est > 0

2

Une catégorie d'activité majoritaire (CAP-M) est renseignée et celle-ci est présente dans la liste des choix possibles

Les déclarations dont la CAP-M, la CAP-S, la SCAP-M ou la SCAP-S n'est pas dans la liste des catégories actives définies dans l'[outil d'import csv](#) (v5).

3

Surface moyenne annuelle \leq Surface totale brute

Ce test de cohérence permet de vérifier que les consommations n'ont pas été déclarées sur une période de plus d'une année, en conservant une marge d'erreur de 0,5%, correspondant à une déclaration de 1 an + 2 jours. En d'autres termes, la déclaration passe le test si :

$$\text{Surface moyenne annuelle} \leq 1,005 \times \text{Surface totale brute}$$

Pour rappel, la surface moyenne annuelle est calculée à partir de la surface de chaque sous-catégorie (SCAP) composant l'EFA et de la durée de la période de déclaration des consommation pour cette SCAP :

$$\widehat{S}_{tot} = \sum_{SCAP} S_{SCAP} \frac{N_{jour}}{365}$$

Avec \widehat{S}_{tot} la surface moyenne annuelle, S_{SCAP} la surface totale brute déclarée pour une SCAP donnée et N_{jour} la durée en jours de la période de déclaration pour cette SCAP.

4

L'EFA n'est pas vacante

Ce test permet de ne sélectionner que les EFA non-vacante, en conservant une marge d'erreur de 1%, i.e. en n'excluant que les EFA vacantes à 99% ou plus.

En d'autres termes, la déclaration passe le test si :

$$\text{Surface moyenne annuelle local vacant} < 0,99 \times \text{Surface moyenne annuelle}$$

5

La période d'activité principale couvre une année complète

On sélectionne ici les déclarations dont la surface moyenne annuelle est égale à la surface totale brute (i.e. la période d'activité de toutes les sous-catégories composant l'EFA est égale à 1 an), en conservant une marge d'erreur de 1%, correspondant à environ 5 jours. On conserve donc les déclarations couvrant au minimum $365 - 5 = 360$ jours.

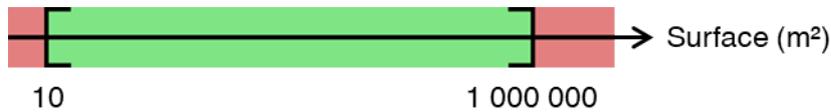
En d'autres termes, la déclaration passe le test si :

$$\text{Surface moyenne annuelle} \geq 0,99 \times \text{Surface totale brute}$$

6

La surface est comprise entre les bornes définies (1^{er} filtre physique)

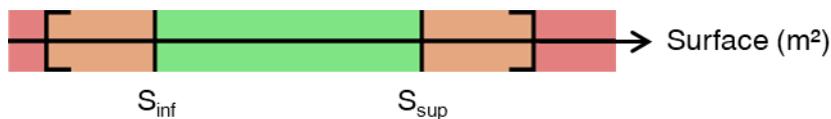
Ce test permet d'exclure les déclarations dont la surface (moyenne annuelle ou totale brute, désormais égales) est considérée comme aberrante (erreurs de saisies).



7

La surface est comprise entre les bornes plus restrictives, qui dépendent de la catégorie d'activité (2^{ème} filtre physique)

Ce test permet d'exclure les déclarations dont la surface est considérée comme anormalement petite ou grande au regard de la (des) CAP à laquelle (auxquelles) elle appartient.



Des bornes inférieure (S_{inf}) et supérieure (S_{sup}) sont définies pour chaque CAP. Elles sont données dans l'Annexe 3 : Catégories d'activité. Celles-ci peuvent être de 3 types :

- Valeur numérique (m^2)
- « Defaut » : minimum (si inférieure) ou maximum (si supérieure) des bornes parmi toutes les valeurs numériques définies
- « Aucune » : pas de borne définie (les bornes du test précédent s'appliquent).

Remarque : une CAP peut avoir une borne inférieure mais pas de borne supérieure et inversement.

Puisqu'une EFA peut être composée de plusieurs CAP, les bornes d'une EFA s'obtiennent en pondérant les bornes des CAP qui la composent par leur poids relatif dans la surface totale :

$$S_{inf(sup)} = \frac{S_{CAP-M}}{S_{EFA}} S_{inf(sup)}^{CAP-M} + \frac{S_{CAP-S}}{S_{EFA}} S_{inf(sup)}^{CAP-S} + (S_{EFA} - S_{CAP-M} - S_{CAP-S}) \frac{S_{inf(sup)}^{défaut}}{S_{EFA}}$$

Diagramme de l'équation ci-dessus :

- Le premier terme $\frac{S_{CAP-M}}{S_{EFA}} S_{inf(sup)}^{CAP-M}$ est lié à : Surface occupée par la CAP-M au sein de l'EFA (numérateur) et Surface totale de l'EFA (surface non vacante) (dénominateur).
- Le deuxième terme $\frac{S_{CAP-S}}{S_{EFA}} S_{inf(sup)}^{CAP-S}$ est lié à : Surface occupée par la CAP-S au sein de l'EFA (numérateur) et Borne inférieure (supérieure) de la CAP-M composant l'EFA (déclaré).
- Le troisième terme $(S_{EFA} - S_{CAP-M} - S_{CAP-S}) \frac{S_{inf(sup)}^{défaut}}{S_{EFA}}$ est lié à : Surface occupée par les éventuelles autres CAP composant l'EFA (numérateur) et Borne inférieure (supérieure) de la CAP-S composant l'EFA (déclaré).
- Le terme $S_{inf(sup)}^{défaut}$ est lié à : Inf : min des S_{inf} / Sup : max des S_{sup} (déclaré).

8

Le périmètre est stable d'une année à l'autre (la surface a varié de moins d'un facteur 2 par rapport à l'année précédente et la CAP n'a pas changé)

Ce test permet d'exclure les déclarations dont la surface a beaucoup varié par rapport à l'année précédente sans que la catégorie ne change, ou celles pour lesquelles la catégorie a évolué. En d'autres termes, il permet de conserver les EFA dont la catégorie et la surface sont « stabilisées », c'est-à-dire n'ont pas trop varié pendant au moins 2 années consécutives.

La déclaration passe le test si :

- L'EFA a fait l'objet d'une déclaration en année N-1
- Cette déclaration a obtenu un score PS = 7
- Le nombre de CAP distinctes, la CAP-M et la CAP-S n'ont pas changé
- La surface moyenne annuelle en année N, S_N , est telle que :

$$\frac{1}{X} \times S_{N-1} \leq S_N \leq X \times S_{N-1} \text{ où } X = 2$$

Remarque : pour la première année de déclaration (2020), comme pour l'année de référence, toutes les déclarations passent automatiquement le test (à condition d'avoir un score PS = 7).

9

La période d'activité principale couvre de janvier à décembre

Ce test permet de s'affranchir d'éventuels biais d'interprétation dus à des variations de période d'activité entre déclaration d'une même année de consommation. Pour rappel, toutes les déclarations qui ont passé le test PS5i couvrent désormais une année complète.

La déclaration passe le test si la date de début d'activité est comprise entre le 1^{er} et le 5 janvier de l'année de consommation.

Score E

1

Une consommation totale en EF positive non-nulle est renseignée
ET
PS \geq 5

En prenant PS \geq 5, ce test permet d'éliminer d'emblée les déclarations avec des surfaces incohérentes, les EFA vacantes et les EFA dont la période de déclaration ne couvre pas une année complète. On ne conserve ensuite que les déclarations pour lesquelles la consommation totale en énergie finale (kWh PCI) est strictement positive (> 0).

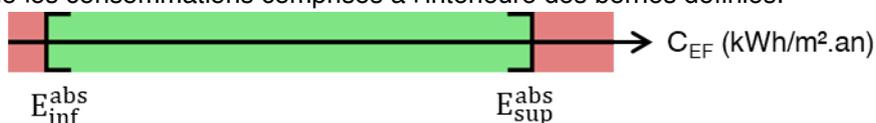
On peut alors calculer la consommation totale surfacique :

$$C_{EF} [kWh \text{ PCI}/m^2] = \frac{\text{Consommation énergétique totale en EF}}{\text{Surface annuelle moyenne non vacante}}$$

2

C_{EF} comprise entre les bornes absolues (filtre physique)

Ce premier filtre permet d'exclure les consommations considérées comme aberrantes. Ne sont conservées que les consommations comprises à l'intérieure des bornes définies.



- **Borne inférieure** E_{inf}^{abs} . Consommation d'un ordinateur standard (50W) pendant 228 jours (jours ouvrés sur un an moins 5 semaines), à raison de 7h/jour (79.8 kWh/an), plus la consommation annuelle moyenne d'une ampoule LED (9 kWh/an), pour une surface de 10 m². $E_{inf}^{abs} = (79.8 + 9)/10 \approx 10$ kWh/m²/an
- **Borne supérieure** E_{sup}^{abs} . Deux cas de figure :
 - Si l'EFA contient un data center : consommation estimée d'un très grand data center avec une densité énergétique maximale et un taux d'utilisation de 100% (100 000 kWh/m²/an)
 - Sinon : consommation maximale observée d'un data center parmi les EFA composées uniquement de « Salles serveurs et centres d'exploitation informatique », avec PS \geq 7 et E = 2. Sur la base transmise le 31/01/2024, cette borne vaut 17 374 kWh/m²/an.

Calcul de la consommation maximale d'un data center :

- Densité énergétique maximum (>96% des cas) d'un data center ([Uptime Institute, 2021](#)) = 50 kW/rack
 - Emprise au sol d'un rack standard d'après [Rack Solution](#) = 0.74 m² → densité énergétique = 68 kW/m²
 - Taux d'utilisation = 100% et sans interruption (8760h/an) → conso IT = 595 680 kWh/m²/an
 - PUE = 1.6 (PUE moyen, [Uptime Institute, 2021](#))
 - Surface équipements IT/Surface totale = 10% (DEET, Arrêté Vabs III)
- Consommation surfacique = 595 680 x 1.6 x 0.1 \approx 100 000 kWh/m²/an

3

C_{EF} comprise entre les bornes relatives (filtre statistique)

Une fois le filtre précédent appliqué, on applique un deuxième filtre, statistique cette fois, qui consiste à retirer les valeurs extrêmes, en utilisant la règle empirique dite des « 3 sigma ».

Etant donné que différentes catégories d'activité (CAP) peuvent avoir des profils énergétiques très variables, et que les consommations évoluent d'année en année, il faut définir des bornes différentes pour chaque CAP et pour chaque année de consommation.

La méthode de calcul des **bornes pour une CAP et une année de consommation (N)** données est la suivante :

- (1) Sélection du périmètre de calcul :
 - Sélection des EFA avec une seule CAP renseignée (la CAP dont on veut calculer les bornes)
 - Sélection des EFA avec PS \geq 8 et E = 2, afin de garder les données les plus fiables, en conservant toutefois les déclarations ne commençant pas au 1^{er} janvier
 - Sélection des années de consommation N-1, N, N+1, afin de disposer d'un échantillon de taille suffisante centré autour de l'année N. Pour la première année, on sélectionnera les données N_min, N_min+1, N_min+2. De même, pour la dernière année (N_max), on sélectionnera les données N_max-2, N_max-1, N_max. Enfin, pour l'année de référence, on sélectionne uniquement les données sur cette-dernière.
- (2) La distribution de consommations surfaciques (kWh/m²/an) ainsi obtenue est dans la plupart des cas biaisée (skewed). On calcule la **distribution symétrique** des consommations surfaciques en

appliquant la fonction appropriée (celle donnant le plus petit [coefficient de Pearson en valeur absolue |G1|](#)) parmi les 4 fonctions suivantes : [aucune, sqrt, cbrt, log]. On calcule ensuite la **moyenne μ** et l'**écart-type σ** de la distribution symétrique obtenue.

- (3) On vérifie que la distribution a une taille suffisante (**min. 30 lignes**). Si ce n'est pas le cas, on ajoute chaque distribution de taille < 30 à la distribution « la plus proche », jusqu'à ce que toutes les distributions aient une taille de 30 au moins. Pour déterminer la distribution « la plus proche » d'une distribution donnée, procéder comme suit : appliquer le log à toutes les distributions (transformation la plus commune) et calculer leur moyenne. Ajouter la distrib < 30 lignes à celle dont la moyenne est la plus proche.
- (4) Enfin, on calcule les **bornes statistiques** pour cette distribution en prenant la transformation inverse de (**$\mu \pm 3\sigma$**). Par exemple, si la transformation utilisée est le log, la borne inférieure est $\exp(\mu - 3\sigma)$.

Une fois les bornes inférieure et supérieure calculées pour chaque CAP et chaque année de consommation, on peut calculer les **bornes de chaque déclaration** :

- Pour les EFA ne comportant qu'une CAP, on renseigne les bornes calculées précédemment
- Pour les EFA comportant plusieurs CAP, les bornes s'obtiennent en pondérant les bornes des CAP qui la composent par leur poids relatif dans la surface totale (voir test PS7).

4

La consommation est stable d'une année à l'autre (C_{EF} a varié de moins d'un facteur 2 par rapport à l'année précédente et)
ET
 $PS \geq 8$

Ce test permet de sélectionner les déclarations dont la surface et la catégorie (condition $PS \geq 8$), mais aussi la consommation sont « stabilisées », c'est-à-dire n'ont pas trop évolué par rapport à l'année précédente.

La déclaration passe le test si :

- La déclaration a obtenu un score $PS \geq 8$. Cela permet de s'assurer que l'EFA a fait l'objet d'une déclaration en année N-1, que le CAP n'a pas changé et que la surface est comprise entre les bornes définies par les tests précédents.
- La déclaration en année N-1 a obtenu un score $PS \geq 7$ et $E = 3$. Cela permet de s'assurer que la surface et la consommation de l'année N-1 sont cohérentes au regard des tests précédents.
- La consommation en année N, C_{EF} , est telle que :

$$\frac{1}{X} \times C_{EF,N-1} \leq C_{EF,N} \leq X \times C_{EF,N-1} \text{ (où } X = 2\text{)}$$

Remarque : pour la première année de déclaration (2020), comme pour l'année de référence, toutes les déclarations passent automatiquement le test (à condition d'avoir un score $PS \geq 8$).

Annexe 3 : Catégories d'activité

Le tableau ci-dessous dresse la liste des **60 catégories d'activité (CAP)**, telles que définies dans l'[outil d'import csv des consommations](#) (V5.0 compatible V6.0). Par souci de simplicité, un intitulé abrégé des CAP a été défini. Par ailleurs, **28 méta-catégories**, regroupant des catégories similaires, sont définies. Celles-ci sont utilisées à des fins statistiques, pour le calcul d'indicateurs de consommations énergétiques notamment. Enfin, la table donne les bornes surfaciques utilisées au test PS 7 pour chaque CAP.

Méta-catégories	Catégories et sous-catégories	Catégories abrégées	S_inf (m²)	S_sup (m²)
Accueil petite enfance	Accueil petite enfance	Accueil petite enfance	Defaut	Defaut
Audiovisuel	Audiovisuel - Radio	Audiovisuel - Radio	Defaut	100000
Audiovisuel	Audiovisuel - Télévision et télédiffusion	Audiovisuel - TV	Defaut	57000
Blanchisserie	Blanchisserie dite "industrielle"	Blanchisserie	Defaut	Defaut
Bureaux	Bureaux – Services Publics - Banque	Bureaux	50	240000
Commerce alimentaire	Commerce - Grande Surface Alimentaire - Supérette (surface de vente < 400 m²)	GSA Supérette	Defaut	800
Commerce alimentaire	Commerce - Grande Surface Alimentaire - Petit supermarché (surface de vente comprise entre 400 m² et 1 000 m²)	GSA Petit Supermarché	400	2000
Commerce alimentaire	Commerce - Grande Surface Alimentaire - Grand supermarché (surface de vente comprise entre 1 000 m² et 2 500 m²)	GSA Grand Supermarché	1000	5000
Commerce alimentaire	Commerce - Grande Surface Alimentaire - Hypermarché (surface de vente supérieure à 2 500 m²)	GSA Hypermarché	2500	40000
Commerce alimentaire	Commerces et services de détail - Alimentaire	Détail - Alimentaire	Defaut	800
Commerce autre	Commerce - Grande Surface de Bricolage	GSB	1000	40000
Commerce autre	Commerce - Grande Surface Spécialisée - Equipement de la personne et loisirs	GSS - Personne & loisirs	1000	40000
Commerce autre	Commerce - Grande Surface Spécialisée - Equipement de la maison	GSS - Maison	1000	40000
Commerce autre	Commerce - Grande Surface Spécialisée - Equipement automobile et Moto	GSS - Auto & moto	1000	40000
Commerce autre	Commerces et services de détail - Equipement de la personne et loisirs	Détail - Personne & loisirs	Defaut	2000
Commerce autre	Commerces et services de détail - Equipement de la maison	Détail - Maison	Defaut	2000
Commerce de gros	Commerce de gros	Commerce de gros	Defaut	40000
Culture	Culture et spectacles - Bibliothèque et médiathèque	Culture - Bibliothèques	Defaut	200000
Culture	Culture et spectacles - Musée en bâtiment historique classé ou inscrit	Cultures - Musée patrimoine	Defaut	Aucune
Culture	Culture et spectacles - Musée en bâtiment sans protection patrimoniale	Cultures - Autres musées	Defaut	150000
Culture	Culture et spectacles - Salle d'exposition & Galerie d'art	Cultures - Exposition	Defaut	150000
Culture	Culture et spectacles - Salle de spectacles vivants	Culture - Spectacle vivant	Defaut	100000
Culture	Culture et spectacles - Cinéma	Culture - cinéma	Defaut	50000
Culture	Culture et spectacles - Espèces vivantes	Culture - Zoo	Defaut	Defaut
Enseignement	Enseignement Primaire	Enseignement Primaire	Defaut	130000
Enseignement	Enseignement Secondaire	Enseignement Secondaire	Defaut	130000
Enseignement	Enseignement Supérieur	Enseignement Supérieur	Defaut	Aucune

Enseignement	Enseignement - Formation continue pour adultes	Enseignement - Autres	Defaut	Defaut
Halles	Halles et marchés couverts	Halles	Defaut	Defaut
Hébergement touristique	Hébergement touristique de courte durée (auberge de jeunesse, centre-sportif, colonies de vacances, gîte d'étape et refuge de montagne)	Tourisme - Courte durée	Defaut	Defaut
Hébergement touristique	Résidence de tourisme et loisirs	Tourismes - Résidences	Defaut	Defaut
Hébergement touristique	Terrain de camping et parcs pour caravanes ou véhicules de loisirs	Camping	Defaut	Defaut
Hôtels	Hôtellerie	Hôtels	Defaut	Defaut
Imprimerie	Imprimerie et reprographie	Imprimerie	Defaut	Defaut
Justice	Justice - Tribunaux	Justice - Tribunaux	Defaut	120000
Justice	Justice - Etablissements pénitentiaires	Justice - Prisons	Defaut	170000
Justice	Justice - Protection Judiciaire de la Jeunesse	Justice - Jeunesse	Defaut	170000
Laboratoires	Laboratoires (hors périmètre médical : étalonnage, suivi écologique...)	Laboratoires	Defaut	Defaut
Local vacant	Local vacant	Local vacant	NA	NA
Logistique	Logistique	Logistique	Defaut	170000
Nuit & loisirs	Etablissement de nuit et de loisirs	Nuit & loisirs	Defaut	Defaut
Parc à thèmes	Parc d'attractions et parc à thèmes	Parc à thèmes	Defaut	Aucune
Parc d'expositions	Parc d'expositions	Parc d'expositions	Defaut	Aucune
Parties communes CC	Parties communes des centres commerciaux et des galeries commerciales	Parties communes CC	Defaut	Defaut
Restauration	Restauration - Débit de boissons	Restauration	Defaut	Defaut
Santé	Santé - Centre hospitalier public et privé	Santé - Centres hospitaliers	Defaut	Defaut
Santé	Santé - Etablissements médico-sociaux	Santé - EMS	Defaut	Defaut
Santé	Santé - Activités de santé libérales avec process	Santé - Activités libérales	Defaut	Defaut
Serveurs & IT	Salles serveurs et centres d'exploitation informatique	Serveurs & IT	Defaut	70000
Sports	Sports	Sports	Defaut	100000
Stationnement	Stationnement	Stationnement	Defaut	Aucune
Transport	Transport aérien de voyageurs - Site aéroportuaire	Transport aérien	Defaut	Aucune
Transport	Transport ferroviaire de voyageurs	Transport ferroviaire	Defaut	Aucune
Transport	Transport maritime ou fluvial de voyageurs	Transport navigable	Defaut	Aucune
Transport	Transport routier de voyageurs	Transport routier	Defaut	Aucune
Transport	Transport urbain de voyageurs (métro)	Transport urbain	Defaut	Aucune
Vente véhicules et engins	Vente et services véhicules légers	Vente véhicules légers	Defaut	40000
Vente véhicules et engins	Vente et services véhicules utilitaires et véhicules industriels	Vente utilitaires et véhicules industriels	Defaut	40000
Vente véhicules et engins	Vente et services motocycle	Vente motocycles	Defaut	40000
Vente véhicules et engins	Vente et services engins nautiques et de plaisance	Ventes engins nautiques	Defaut	40000