



RÉPUBLIQUE
FRANÇAISE

Liberté
Égalité
Fraternité



Géosciences pour une Terre durable

brgm

Document à accès immédiat

Projet « De l'Eau pour Demain » - Analyse de la consommation en eau potable à partir des fichiers de facturation - Guide méthodologique et application



Rapport final

BRGM/RP-73025-FR

Version 1 du 29 août 2023

Étude réalisée dans le cadre des projets de recherche et développement

Hévin G., Rinaudo JD., Montginoul M. (INRAE)

| Vérificateur : | Approbateur : |
|---|---|
| Nom : BOISSON Alexandre | Nom : DRZEWIECK Pauline |
| Fonction : Ingénieur hydrogéologue | Fonction : Directrice |
| Date : 29/08/2023 | Date : 30/10/2023 |
| Signature :  | Signature :  |

Le système de management de la qualité et de l'environnement du BRGM
est certifié selon les normes ISO 9001 et ISO 14001.

Contact : qualite@brgm.fr



Avertissement

Ce rapport est adressé en communication exclusive au demandeur, au nombre d'exemplaires prévu.

Le demandeur assure lui-même la diffusion des exemplaires de ce tirage initial.

La communicabilité et la réutilisation de ce rapport sont régies selon la réglementation en vigueur et/ou les termes de la convention.

Le BRGM ne saurait être tenu comme responsable de la divulgation du contenu de ce rapport à un tiers qui ne soit pas de son fait et des éventuelles conséquences pouvant en résulter.

Votre avis nous intéresse

Dans le cadre de notre démarche qualité et de l'amélioration continue de nos pratiques, nous souhaitons mesurer l'efficacité de réalisation de nos travaux.

Aussi, nous vous remercions de bien vouloir nous donner votre avis sur le présent rapport en complétant le formulaire accessible par cette adresse <https://forms.office.com/r/yMgFcU6Ctg> ou par ce code :



Mots clés : Consommation en eau potable ; Fichiers de facturation ; Tendence.

En bibliographie, ce rapport sera cité de la façon suivante :

Hévin G., Rinaudo JD., Montginoul M. (INRAE) 2023. Projet « De l'Eau pour Demain » - Analyse de la consommation en eau potable à partir des fichiers de facturation -

Guide méthodologique et application. Rapport final V1. BRGM/RP-73025-FR, 50 p., 6 fig., 1 ann.

© BRGM, 2023, ce document ne peut être reproduit en totalité ou en partie sans l'autorisation expresse du BRGM.
IM003-MT008-P2-09/03/2023

Synthèse

Le projet “Eau Pour Demain”, mené par les syndicats d'eau potable de Bretagne (SDEAP22, SMG35, Eau du Morbihan), le conseil départemental du Finistère et le BRGM, avec le soutien financier de l'Agence de l'Eau Loire Bretagne et le conseil régional de Bretagne vise à développer une meilleure compréhension des usages de l'eau et du fonctionnement des ressources en eau régionales. L'objectif est de développer de nouveaux outils de gestion pour une meilleure adaptation au changement climatique en Bretagne.

Ce rapport rédigé dans le cadre des travaux de la Tâche 1 du projet axé sur la caractérisation de la consommation en eau potable sur le réseau public, présente une méthode permettant d'analyser le niveau et les tendances d'évolution de la consommation des différentes catégories d'usagers à partir des données contenues dans les fichiers de facturation.

Cette méthode, développée par INRAE en partenariat avec le BRGM dans le cadre de projets de recherche antérieurs, a été adaptée et améliorée dans le cadre du projet Eau Pour Demain, en collaboration avec les syndicats bretons gestionnaires du service d'eau potable associés au projet. Ce projet est soutenu financièrement par l'Agence de l'Eau Loire Bretagne et le Conseil Régional de Bretagne.

Ce rapport se focalise sur une présentation de la méthode afin de permettre à des collectivités qui souhaiteraient réaliser une étude similaire de rédiger un cahier des charges. L'analyse proposée dans cette méthode permet (i) de mieux appréhender la part de la consommation totale imputable à différents types d'abonnés ; (ii) de calculer des ratios unitaires de consommation pour une grande diversité d'usagers ; (iii) d'analyser l'hétérogénéité spatiale de ces ratios de consommation et (iv) de caractériser les tendances d'évolution de la consommation des différents types d'abonnés.

Le rapport présente les résultats d'une application de la méthode à deux secteurs de distribution dans le Centre Morbihan, gérés par la SAUR pour le compte du syndicat départemental Eau du Morbihan. Le rapport présente différents types de graphiques qui peuvent être demandés dans le cadre d'une consultation adressée à un prestataire spécialisé ou dans le cadre d'un contrat de délégation. La méthode a également été appliquée à deux autres collectivités, le Syndicat du Lié (département des Côtes d'Armor). Les résultats de cette analyse ont alimenté une réflexion portée dans un autre rapport du projet coordonné par le Syndicat des Eaux des Côtes d'Armor (Rivallan, J., Jehanno, F., Cauet, Y., Ainaoui, M., Barraï, N., Berrehouc, G., 2022. Connaissance de la consommation d'eau potable distribuée par les réseaux publics. De l'Eau Pour Demain).

Sommaire

| | |
|--|----|
| 1. Introduction : pourquoi analyser les fichiers de facturation ? | 7 |
| 1.1. Une mauvaise connaissance de la consommation | 7 |
| 1.2. Les fichiers de facturation, une source d'information sous-exploitée | 8 |
| 1.3. Mieux connaître la consommation pour mieux gérer | 8 |
| 1.3.1. <i>Prévision de la demande</i> | 8 |
| 1.3.2. <i>Cibler les mesures d'économie d'eau</i> | 8 |
| 1.3.3. <i>Simuler l'impact de changement de politique tarifaire</i> | 9 |
| 1.3.4. <i>Construire un observatoire de la consommation</i> | 9 |
| 1.4. Objectif et contenu de ce document | 9 |
| 2. Comment analyser les fichiers de facturation ? | 13 |
| 2.1. Les conditions d'utilisation des données | 13 |
| 2.1.1. <i>Précautions relatives à l'utilisation de données personnelles (RGPD)</i> | 13 |
| 2.1.2. <i>Extraction des données</i> | 13 |
| 2.2. Les étapes du traitement des données | 14 |
| 2.2.1. <i>Extraction et mise en forme du fichier de facturation</i> | 14 |
| 2.3. Nettoyage des données | 14 |
| 2.4. Calcul d'une consommation sur 365 jours | 15 |
| 2.5. Typologie des abonnés | 16 |
| 2.6. Statistiques descriptives | 19 |
| 2.7. Classification des abonnés selon leur niveau de consommation | 20 |
| 2.8. Test de tendance d'évolution de la consommation en eau | 20 |
| 2.9. Graphiques et cartes | 21 |
| 2.10. Exportations de l'ensemble des résultats sous deux formats | 21 |
| 2.11. Les outils pouvant être utilisés | 22 |
| 2.11.1. <i>Outils issus de la recherche</i> | 22 |
| 2.11.2. <i>Autres outils</i> | 23 |
| 3. Visualisation des résultats | 25 |
| 3.1. Résultat de la classification | 25 |
| 3.1.1. <i>Analyse globale des effectifs et des volumes facturés</i> | 25 |
| 3.1.2. <i>Analyse spatiale</i> | 26 |
| 3.2. Analyse des ratios de consommation unitaire par type | 27 |
| 3.2.1. <i>Analyse globale</i> | 28 |
| 3.2.2. <i>Analyse spatiale</i> | 28 |
| 3.3. Analyse des distributions statistiques | 29 |
| 3.3.1. <i>Boîte à moustache</i> | 29 |
| 3.3.2. <i>Courbes de densité</i> | 31 |
| 3.4. Analyse de la dynamique temporelle | 32 |
| 3.4.1. <i>Évolution des effectifs et des volumes facturés</i> | 32 |
| 3.4.2. <i>Analyse des tendances globales</i> | 34 |
| 3.4.3. <i>Analyse des dynamiques spatio-temporelles</i> | 36 |
| 3.4.4. <i>Analyse des tendances individuelles</i> | 37 |
| 3.5. Analyses complémentaires | 39 |
| 4. Conclusion | 41 |
| 4.1. Une source d'information à valoriser | 41 |
| 4.2. Une analyse à réaliser en routine | 41 |
| 5. Références bibliographiques | 43 |

Liste des figures

| | |
|---|----|
| Figure 1 : Principe de calcul d'une consommation sur 365 jours. | 15 |
| Figure 2 : Typologie des abonnés - Favre et Montginoul (2020)..... | 17 |
| Figure 3 : Présentation du principe d'une boîte à moustache..... | 20 |
| Figure 4 : Présentation du principe du test de tendance de Mann Kendall. | 21 |
| Figure 5 : Représentation schématique de la circulation des données au cours de l'analyse... .. | 22 |
| Figure 6 : L'analyse des fichiers de facturation intégrée dans une stratégie de gestion du fichier clientèle et facturation..... | 42 |

Liste des annexes

| | | |
|----------|----------------------------|----|
| Annexe 1 | Exemple de convention..... | 45 |
|----------|----------------------------|----|

1. Introduction : pourquoi analyser les fichiers de facturation ?

Le projet "Eau Pour Demain", mené par les syndicats d'eau potable de Bretagne (SDEAP22, SMG35, Eau du Morbihan), le conseil départemental du Finistère et le BRGM, avec le soutien financier de l'Agence de l'Eau Loire Bretagne et le conseil régional de Bretagne vise à développer une meilleure compréhension des usages de l'eau et du fonctionnement des ressources en eau régionales. L'objectif est de développer de nouveaux outils de gestion pour une meilleure adaptation au changement climatique en Bretagne.

La consommation en eau potable et les déterminants de son évolution restent des éléments relativement méconnus. Ainsi, en Bretagne, si, comme dans beaucoup d'autres régions en France, une baisse a été observée jusqu'en 2016, la tendance s'est inversée à partir de la sécheresse de 2017, la consommation par abonné augmentant alors, sans que l'on en identifie précisément les raisons et les abonnés concernés.

Pour les experts des collectivités concernées, cette évolution à la hausse pourrait s'expliquer par des changements démographiques (une augmentation de la population permanente), de mode de vie (un télétravail qui permet d'utiliser sa résidence secondaire de manière plus intense, accentué depuis l'épisode COVID de 2020) ou d'usage de l'eau (notamment des consommations dites « de confort »). Elle pourrait aussi résulter de l'évolution de certaines activités industrielles (notamment agroalimentaires) ou de la fréquentation touristique, ou encore refléter l'abandon ou les reports de ressources propres (forages privés) par de gros consommateurs agricoles ou agroalimentaires. Ces hypothèses restent difficilement vérifiables par manque de données permettant d'analyser finement la consommation des différentes catégories d'usagers.

1.1. UNE MAUVAISE CONNAISSANCE DE LA CONSOMMATION

De manière générale, les services d'eau potable connaissent relativement mal la typologie de leurs abonnés et leurs consommations. Ils ne disposent en général que du volume total facturé et du nombre total d'abonnés, ce qui ne permet que de calculer une consommation moyenne par abonné. Dans le meilleur des cas, il est possible de faire une distinction entre abonnés assimilés domestiques et gros consommateurs. Il n'existe généralement pas de classification plus fine des abonnés (ou peu fiable si existante) permettant de suivre l'évolution des consommations par secteur économique et, au sein de chaque secteur, par type d'abonnés (par exemple camping, hôtel, restaurants au sein du secteur touristique).

Ces valeurs globales de consommation unitaire par abonné cachent une grande diversité. En effet, on regroupe souvent sous le terme « usagers domestiques et assimilés » des clients aussi différents que des ménages avec compteurs individualisés, des immeubles comprenant plusieurs logements, leur local poubelles parfois équipé d'un compteur, des commerces, des restaurants, des petites entreprises diverses, des écoles et autres bâtiments administratifs, etc. La catégorie « gros consommateurs » peut inclure, des industries, un hôpital, un collège ou une piscine municipale, etc.

Or, au sein de ces deux grandes catégories, la consommation de chacun de ces types d'abonnés peut évoluer de manière très différente sans que le gestionnaire du service ne puisse l'observer, ne disposant que de tendances moyennes et globales. Il est fort possible qu'une baisse tendancielle marquée d'un type d'utilisateur soit compensée par une hausse tout aussi forte d'un autre type de consommateur, la résultante étant une apparente stabilité, ce qui peut induire des erreurs de diagnostic et par la suite de prévision.

1.2. LES FICHIERS DE FACTURATION, UNE SOURCE D'INFORMATION SOUS-EXPLOITÉE

Le fichier de facturation constitue une formidable source de données, à l'heure actuelle sous-exploitée.

Ce fichier contient en effet des informations qui, via un traitement particulier, permettent de commencer à mieux caractériser le type d'usage associé à chaque abonné. Ces informations sont contenues dans des champs destinés à assurer la facturation, comme le nom et l'adresse de l'abonné, le diamètre de son compteur, etc.

Le nom de l'abonné est à lui seul une source d'information essentielle, puisqu'il contient souvent de manière explicite des termes décrivant l'activité économique : boulangerie, camping, hôtel, restaurant, etc. Une simple lecture du fichier permettrait ainsi de coder les abonnés, affectant chaque client à une catégorie prédéfinie. Cette opération serait néanmoins très fastidieuse, notamment pour les collectivités gérant plusieurs dizaines de milliers d'abonnés. D'où l'idée de développer un algorithme qui réalise ce classement des abonnés de manière automatique.

Cet algorithme permettrait ainsi d'aider à classer les abonnés selon une typologie décrivant les activités et les usages, donc de fiabiliser à terme la base de données, non seulement à des fins d'observatoire de la consommation, mais également de suivi de la facturation et d'accompagnement des opérations de recouvrements.

1.3. MIEUX CONNAÎTRE LA CONSOMMATION POUR MIEUX GÉRER

L'exploitation des informations contenues dans le fichier de facturation peut répondre à de nombreux besoins.

1.3.1. Prévoir la demande future

Le premier est le besoin de prévision d'évolution de la consommation. Lorsqu'une collectivité réalise un schéma directeur d'alimentation, elle doit faire des hypothèses sur l'évolution future de la demande afin de dimensionner les ouvrages à réaliser et les ressources à mobiliser. Cette prévision est aussi nécessaire pour la planification financière à moyen et long terme (estimation de l'assiette de facturation et détermination des tarifs permettant un équilibre financier).

Or, l'état de l'art en matière de prévision de la demande préconise de calculer les besoins futurs par type d'usagers, en précisant pour chacun une hypothèse d'évolution des effectifs et d'évolution du ratio de consommation unitaire (Rinaudo, 2013). Si l'on ne dispose pas de ces informations par catégorie d'usagers, on ne peut que faire une prévision globale basée sur un ratio unitaire par habitant, plus aléatoire et moins précis.

1.3.2. Cibler les mesures d'économie d'eau

Le deuxième besoin est relatif à la mise en œuvre de programmes d'économies d'eau. Connaître les ratios de consommation par catégorie d'usager est essentiel pour :

- Identifier les catégories de consommateurs dont la consommation moyenne continue d'augmenter ou stagne (cibles prioritaires) et celles qui baissent spontanément (pas de besoin d'actions ciblées) ;
- Estimer le volume total pouvant être économisé par le biais de mesures volontaristes en matière d'économie d'eau (ex. distributions de kits hydro économes chez les abonnés

domestiques) ; cette estimation suppose en effet de connaître le nombre d'abonnés de la population ciblée et leur consommation moyenne et totale, l'économie étant ensuite estimée sur la base de résultats d'expériences similaires conduites par d'autres collectivités (% de gain possible).

L'analyse des fichiers de facturation peut aussi aider à appréhender la diversité des consommations au sein des catégories. Par exemple, les hôtels peuvent avoir un niveau de consommation très hétérogène (rapportée au nombre de lits). Caractériser cette hétérogénéité et comprendre quelles sont les caractéristiques des plus gros consommateurs (ex. nombre d'étoiles) peut aider à mieux cibler les leviers d'action pour les inciter à faire des économies.

1.3.3. Simuler l'impact de changement de politique tarifaire

Lors de la mise en place d'une nouvelle politique de tarification, le gestionnaire souhaite généralement anticiper les effets sociaux et économiques.

Le fait de connaître la consommation moyenne de différentes catégories d'abonnés (par ex. domestiques habitant en appartement ou en maison, petits commerces, hôtels, secteur tertiaire privé, etc.) permet de simuler l'impact que pourrait avoir un changement de tarification sur la facture de chaque catégorie.

Ceci peut aider à anticiper les conséquences sociales ou économiques de scénarios tarifaires envisagés sur chaque catégorie d'abonnés, mais aussi d'estimer l'impact global des scénarios tarifaires sur les recettes futures et donc de mieux planifier les investissements à long terme.

1.3.4. Construire un observatoire de la consommation

Les données agrégées actuellement exploitées (consommation moyenne par habitant) sont entachées de nombreuses incertitudes, notamment liées à la variabilité des dates de relève des compteurs, mais aussi des fuites et dégrèvements. Ceci rend difficile l'analyse des tendances de consommation.

Disposer d'une analyse plus précise des évolutions de consommation basée sur une exploitation détaillée du fichier de facturation apporte une plus grande prévision et fiabilité de suivi. L'analyse répétée des fichiers de facturation permet ainsi de construire, année après année, un observatoire fiable des consommations. Cet observatoire permettra de suivre la dynamique spatiale, sectorielle et temporelle de la consommation.

1.4. OBJECTIF ET CONTENU DE CE DOCUMENT

Ce rapport présente une méthode permettant d'analyser la valeur de distribution statistique et les tendances d'évolution de la consommation des différentes catégories d'utilisateurs, à partir des données contenues dans les fichiers de facturation. Cette méthode, développée par INRAE en partenariat avec le BRGM dans le cadre de projets de recherche antérieurs, a été améliorée et adaptée aux cas des collectivités étudiées dans le projet « *de l'Eau pour Demain* ». Elle est sommairement présentée dans ce rapport afin de permettre à des collectivités qui souhaiteraient réaliser une étude similaire de rédiger un cahier des charges. La méthode a été appliquée à trois collectivités bretonnes à titre de test. Des illustrations des résultats obtenus pour l'une de ces collectivités, située dans le département du Morbihan, sont présentées pour alimenter la réflexion de collectivités qui souhaiteraient préciser leurs attentes vis-à-vis d'un prestataire dans un cahier des charges.

Le présent document vise avant tout à permettre à des élus et techniciens des services en charge de la compétence eau potable de comprendre :

- Les étapes d'analyse des fichiers de facturation ;
- Le type de résultats qui peuvent en résulter ;
- Les éléments à inclure dans un cahier des charges visant à reproduire ce type d'analyse.

1.5. TERRAINS D'APPLICATION

La méthode décrite dans le présent document a été appliquée à quatre secteurs d'étude qui ont été identifiés par les partenaires du projet « *De l'Eau pour Demain* » (Tableau 1 et Figure 1). Au total, ces quatre terrains d'étude représentent 55 communes et plus de 60 000 abonnés. Les données analysées portent sur une dizaine d'années pour trois de ces secteurs et uniquement sur 3 ans pour le quatrième.

| Nom du secteur d'étude | Communes | Abonnés (2019) | Période étudiée |
|--|---|----------------|-----------------|
| Communauté de Communes Pays Bigouden Sud | Ile Tudy, Loctudy, Tréméoc, Pont-l'Abbé, Combrit, Tréguennec, Saint-Jean-Trolimon, Plomeur, Penmarc'h, Treffiagac, Guilvinec | 20,692 | 2011-2021 |
| Syndicat du Pays du Lié | Hémonstoir, Loudéac, La Prénessaye, Plémet, Saint-Maudan, Saint-Barnabé, La Chèze, Plumieux, Coëtlogon, Saint-Etienne-du-gué-de-l'isle, Le Cambout | 7,997 | 2011-2020 |
| Centre Morbihan secteur Veolia | Damgan, Ambon, Billiers, Azral, Muzillac, Noyal-Muzillac, Peillac, Saint-Vincent-sur-Oust, Saint-Jacut-les-pins, Saint-Perreux, Saint-Jean-la Poterie, Saint-Gorgon, Allaire, Béganne, Rieux, Théhillac, Nivillac, Saint-Dollay | 19,989 | 2019-2021 |
| Centre Morbihan, secteur SAUR | Evellys, Moréac, Saint-Allouestre, Buléon, Guéhenno, Billio, Plumelec, Saint-Jean-Brévelay, Bignan, Moustoir-AC, Plumelin, La Chapelle Neuve, Locminé, | 12,405 | 2013-2021 |

Tableau 1 : Secteurs étudiés dans le projet « *De l'Eau pour Demain* ».

:

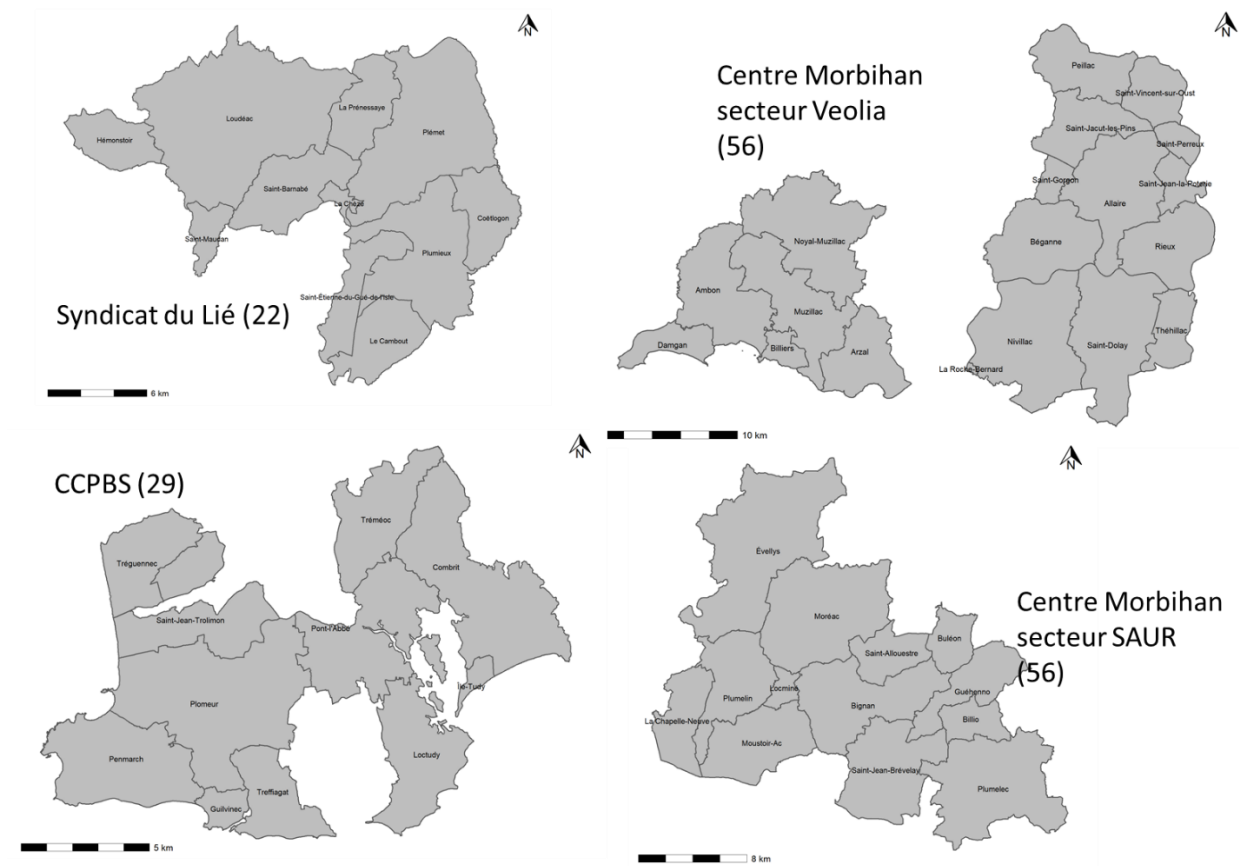


Figure 1 : Carte des secteurs étudiés dans le cadre du projet « De l'Eau pour Demain ».

2. Comment analyser les fichiers de facturation ?

2.1. LES CONDITIONS D'UTILISATION DES DONNÉES

2.1.1. Précautions relatives à l'utilisation de données personnelles (RGPD)

Le fichier de facturation contient des données individuelles. Son utilisation est de ce fait soumise au Règlement Général sur la Protection des Données ou RGPD (voir [Comprendre le RGPD | CNIL](#)). Si son analyse est confiée à un tiers, une convention devra donc être rédigée entre le donneur d'ordre (la collectivité compétente) et le prestataire.

Rappelons que :

- La collectivité compétente est propriétaire des fichiers même si ce fichier est renseigné et géré informatiquement par une entreprise exploitante ;
- Elle peut librement utiliser ce fichier, ou en confier l'exploitation à un prestataire pour son compte, si l'analyse projetée est liée à la réalisation du service pour lequel le fichier a été créé. Elle n'a dans ce cas pas besoin de réaliser une déclaration à la CNIL ;
- En revanche, si la méthode prévoit de croiser le fichier de facturation avec une autre source de données individuelles, des précautions de traitement devront être observées, dans le respect du RGPD.

La convention établie entre la collectivité compétente et le prestataire qui va exploiter le fichier vise avant tout à garantir la sécurité des données personnelles. Elle spécifie notamment :

- L'interdiction d'utiliser les données en dehors du cadre précis de l'étude ;
- L'obligation d'héberger les données sur des systèmes informatiques sécurisés ;
- L'obligation d'informer les autorités compétentes et les personnes concernées en cas de fuite / violation des données à caractère personnel ;
- L'obligation de détruire tous les fichiers contenant des données personnelles à l'issue de l'étude.

Un exemple de convention établie entre les collectivités et le BRGM est annexé à la fin de ce document.

2.1.2. Extraction des données

Il existe plusieurs logiciels de gestion des données de facturation. Les grandes sociétés exploitantes ont développé leur propre outil de gestion de bases de données (SAUR, VEOLIA, SUEZ). Les régies et d'autres entreprises exploitantes peuvent également avoir développé leur propre outil. Enfin, des sociétés de service informatique peuvent vendre des outils spécifiques aux régies.

Du fait de la multiplicité de ces outils informatiques et de l'absence de standardisation et de normes, les données ne sont jamais organisées de la même manière. La première étape de la préparation du cahier des charges d'une telle étude consiste donc à analyser l'organisation des bases de données et de s'assurer des modalités possibles de leur extraction en vue de leur analyse.

Lorsque le logiciel de facturation est opéré par un prestataire de service indépendant, il peut facturer le service d'extraction des données ou refuser de l'exécuter si cela n'a pas été prévu au contrat (situation rencontrée par les auteurs dans le cas de très petites collectivités).

2.2. LES ÉTAPES DU TRAITEMENT DES DONNÉES

2.2.1. Extraction et mise en forme du fichier de facturation

Les logiciels de gestion des données de facturation sont généralement conçus sous forme de base de données relationnelles qui contiennent plusieurs tables attributaires dont, le plus souvent :

- Une table décrivant les branchements (diamètre, adresse, type et numéro de compteur...);
- Une table décrivant les abonnés (noms, adresse, numéro d'abonné);
- Une table décrivant les payeurs (peut être différent de l'abonné);
- Une table décrivant les consommations (numéro du compteur, date de relevé, index relevé);
- Autres tables selon les logiciels et les prestataires de service / fermiers.

La première étape consiste donc à réaliser une extraction de cette base de données pour créer un unique fichier qui sera ensuite exploité. Ce fichier unique comporte une ligne par abonné et toutes les informations utiles sont ajoutées sous forme de colonnes.

Points de vigilance :

- Cette extraction doit être réalisée par le gestionnaire de la base de données de facturation, qui fournit alors un ou plusieurs fichiers « export » en format Excel ou texte (.txt, .csv). Il est essentiel de bien lui préciser la nature des champs requis. L'analyse des tendances d'évolution de la consommation requiert le plus grand nombre d'années possible (10 ans idéalement).
- Le gestionnaire du logiciel de facturation peut ne pas pouvoir extraire les données sous forme d'un seul fichier. Il fournira alors plusieurs fichiers (correspondant aux tables de sa base) qui devront être intégrés par l'acteur chargé de l'analyse des données. Cette étape peut être relativement complexe d'un point de vue informatique. Elle doit être réalisée par un binôme disposant de compétences informatiques et une bonne expertise relative aux données de facturation (pour effectuer le bon choix de variables).
- Dans le cas où le nombre d'abonnés est élevé et la période d'étude importante, le fichier de données extraites de la base de facturation peut avoir une taille dépassant la capacité des tableurs de type Excel ou équivalent. Pour manipuler ces données, il faudra alors avoir recours à des outils plus puissants permettant de programmer le traitement des données, en utilisant des logiciels comme R (libre) ou Matlab ou des outils de programmation (Python). Cette difficulté peut à elle seule justifier les services d'un prestataire.

2.3. NETTOYAGE DES DONNÉES

La première étape consiste à effectuer un nettoyage des données, consistant à rechercher des valeurs aberrantes dans chaque champ. Cette analyse, qui doit être effectuée manuellement, consiste notamment :

- À rechercher les valeurs minimales et maximales de chaque champ ; on détectera ainsi des aberrations comme par exemple un compteur de diamètre 200 mm (au lieu de 20 mm).

- À rechercher les valeurs négatives, issues de calculs ou reflétant des dégrèvements liés à des fuites (consommation).
- À vérifier les effectifs d'abonnés présents chaque année ; il se peut en effet que le périmètre des collectivités compétentes ait changé au cours du temps ; on peut ainsi avoir extrait des données pour un nombre de communes variables, et donc un effectif d'abonnés différent. Dans ce cas, deux stratégies peuvent être choisies :
 - Si l'objectif principal est de caractériser les tendances d'évolution, il faudra alors faire le choix d'un périmètre constant et ne garder que les abonnés des communes (Unités de distribution ou secteurs), pour lesquels on dispose de données pendant toute la période étudiée.
 - Si l'objectif est au contraire de faire une analyse détaillée de la variabilité spatiale sur une période récente, on conservera toutes les communes, donc l'échantillon maximal, mais seront exclues les données historiques pour les années pour lesquelles toute la population actuelle d'abonnés n'est pas présente.

2.4. CALCUL D'UNE CONSOMMATION SUR 365 JOURS

Le fichier de données brutes contient, pour chaque abonné, des dates de relevés et un index correspondant à chaque date. Ces deux informations permettent de recalculer, pour chaque abonné et chaque année, une consommation théorique sur 365 jours. Il est aussi possible que le fichier contienne un volume facturé associé à une date de relevé de compteurs (dans ce cas l'index du compteur n'est pas connu).

La Figure 2 illustre le calcul réalisé dans le cas où le fichier fourni contient une consommation facturée à une date donnée (et non pas les index) :

- Dans cet exemple, on sait que l'abonné a reçu une facture pour 152 m³, relevée au 20/11/2019 ; puis une consommation facturée de 195 m³ au 13/11/2020, et enfin une troisième de 132 m³ au 26/11/2021.
- La première donnée est inexploitable, car on ne connaît pas la date du relevé antérieur ; la consommation 2020 peut être évaluée très précisément ; celle de 2021 est calculée en extrapolant la consommation mesurée sur 11 mois de 2021 au mois de décembre.

Les observations correspondant à une consommation sur l'année inférieure à 6 mois (moins de 180 jours) sont supprimées, car cela supposerait d'extrapoler ce qui a été observé sur 7 mois aux 5 mois manquants (source d'erreur majeure en raison de la saisonnalité de la consommation). Les consommations négatives sont également supprimées.

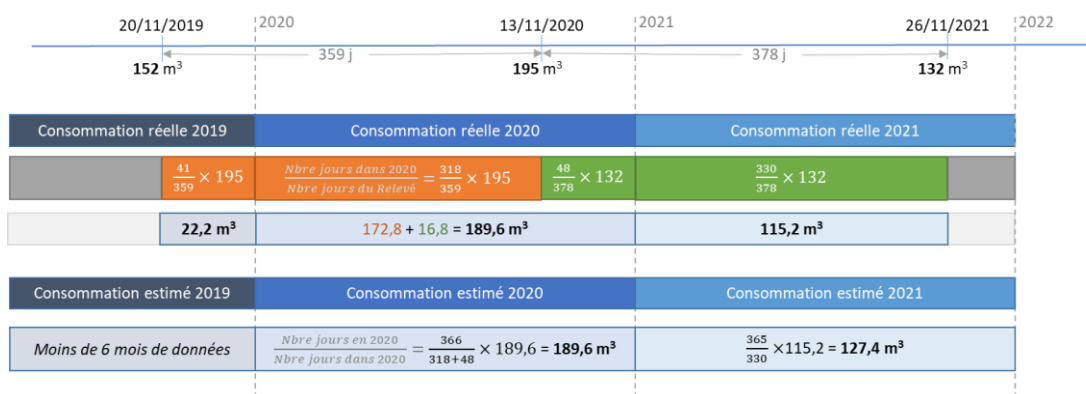


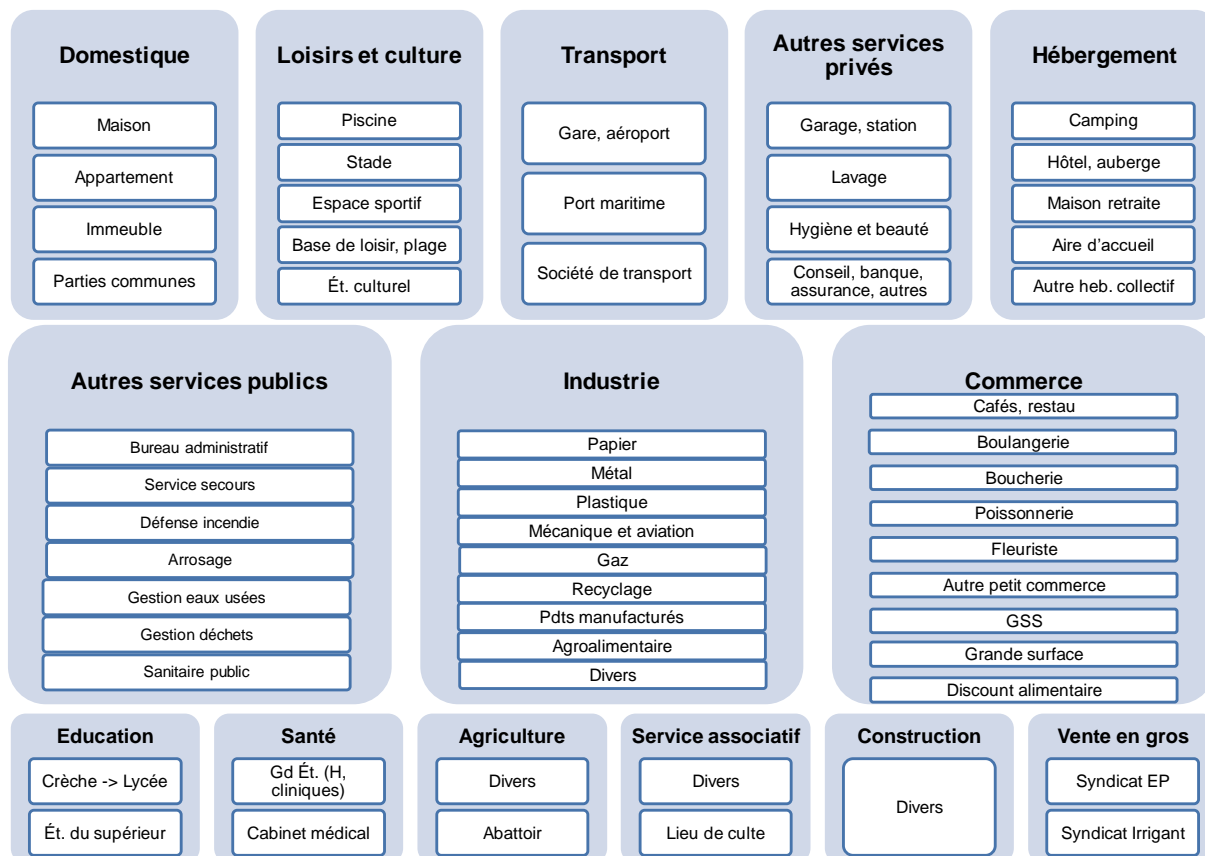
Figure 2 : Principe de calcul d'une consommation sur 365 jours.

Lors de ce calcul, il faut rechercher les cas dans lesquels le numéro de compteur a changé (remplacement). L'index du compteur passe alors d'une valeur positive (la dernière relevée avant le changement du compteur) à zéro (installation du nouveau compteur), ce qui va donner une valeur négative de consommation avec un calcul automatisé.

2.5. TYPOLOGIE DES ABONNÉS

L'étape suivante consiste à classer les abonnés selon une typologie de consommateurs préétablie. Il n'existe pas de typologie universelle répondant à tous les besoins. C'est au commanditaire de l'étude de définir la typologie des consommateurs qui lui paraît pertinente en fonction de l'utilisation qu'il compte faire des données. La typologie optimale sera différente selon que l'on veut effectuer une simulation de l'impact d'un scénario tarifaire ou cibler des actions d'économies d'eau :

- Cette typologie peut être très sommaire (domestique, économique, administrative) ou très détaillée.
- La définition des types peut être basée sur la nature de l'activité, mais aussi sur des critères de consommation (par exemple : un petit consommateur agricole si sa consommation est inférieure à 1000 m³/an ou gros consommateur agricole si supérieur).
- La typologie utilisée dans le code développé par INRAE (voir section suivante), en lien avec la Fédération Nationale des Régies Concédantes ou en Régie (FNCCR), est présentée en Figure 3.



Légende : EP : Eau potable ; Et : Établissement ; Gd : Grand ; GSS : grandes surfaces spécialisées ; H : Hôpital ; Heb : hébergement.

Figure 3 : Typologie des abonnés - Favre et Montginoul (2020)

Une fois la typologie établie, un programme informatique devra être développé pour affecter chaque abonné dans une classe de cette typologie. Deux grandes méthodes sont possibles :

- La première repose sur une **recherche lexicale** dans les différents champs de la base de données, notamment noms et adresses. Cela consiste à construire un registre de mots qui sont associés à chaque type puis à rechercher dans le fichier tous les abonnés dont les champs cibles contiennent ce mot. Quelques exemples sont présentés dans le tableau ci-après. Cette recherche lexicale peut être complétée par des conditions portant sur d'autres champs comme la taille minimale ou maximale des branchements, ou une consommation maximale.
- La deuxième approche consiste à **rechercher des correspondances** entre le fichier de facturation et des bases de données externes. La base de données SIRENE de l'INSEE qui décrit les établissements économiques peut ainsi être utilisée. Un algorithme devra alors être développé pour rechercher une correspondance entre le nom de l'abonné et la raison sociale de l'entreprise.

Quelle que soit la méthode choisie, la classification restera imparfaite et incomplète : une partie des abonnés ne pourra pas être classifiée et certaines erreurs de classifications sont inévitables. Il faut considérer cette classification automatique comme un « débroussaillage » (ou screening) qui peut être amélioré par la suite, via une analyse manuelle des abonnés qui n'ont pas été classés avec la procédure automatique. Son intérêt est de permettre rapidement et facilement la classification de plusieurs dizaines de milliers d'abonnés, ce qui serait coûteux à réaliser manuellement.

| Domestique | Habitat individuel | maison | Civilité = Madame, Monsieur + diamètre de compteur inférieur à 15 mm ou non renseigné Nom = logement de fonction) |
|-------------------|----------------------------------|-------------------------|--|
| | | appartement | Parmi les individuels, si mention d'habitat en immeuble dans l'adresse (étage, appartement, bâtiment, résidence, etc.) |
| | Habitat collectif | immeuble | Titre et Nom = copropriétés, Syndicat, SCP, etc. + diamètre de compteur supérieur à 15 mm Nom = HLM, ILM, Entreprises Sociales pour l'Habitat, Nantaise d'Habitations, Atlantique d'Habitations + diamètre de compteur supérieur à 15 mm |
| | Parties communes | parties communes | Nom = HLM ou copropriété) + type de compteur mentionnant un compteur « divisionnaire » |
| Agriculture | Divers | agriculture divers | Statut et nom = GFA, SCE, GAEC, ASA, COOP, EARL ou GFF, école d'agriculture, terrain agricole, herbage |
| | Abattoirs | abattoirs | Nom = SOLIPAG ABATTOIRS |
| services publics | Services administratifs | bureaux administratifs | Titre = EPIC, SPLA, ADMIN et Nom = Chambre d, préfecture, inspection académique, la Poste, finance publique, trésorerie générale, direction départementale, département, conseil général, pôle emploi, sécurité sociale, caisse primaire, caisse allocation, ville de, ministère, agglomération, communauté de commune, commune, métropole, municipal, mairie, consulat, ambassade, palais de justice, tribunal) |
| | Services de police et de secours | police et de secours | Nom = police, gendarmerie, caserne, capitainerie, pompier, militaire, commissariat, douanes |
| | Bornes incendie | défense incendie | Nom = Réserve eau incendie, centre départ incendie, services incendie |
| | Arrosage des espaces verts | arrosage | Nom = Arrosage, espace vert, giratoire, jardins des églantiers, jardins familiaux... |
| | Assainissement | assainissement | Nom = eaux usées, station épuration, assainissement) |
| | Gestion des déchets | déchet | Nom = Sivom, déchet, décharge |
| | Sanitaires publics | sanitaires publics | Nom = WC publics, bain-douche, sanitaires publics |
| | Autres équipements urbains | autre équipement urbain | Nom = marchés, halles, cimetière |
| Commerce | Café, restaurant | restaurant | Nom = bar, café, brasserie, restaurant, pizzeria, crêperie, snack, discothèque, Mc Do, KFC, Courtepaille, Hippopotamus, Casino cafeteria |
| | Boulangerie | boulangerie | Nom = Boulangerie, pâtisserie, sandwicherie |
| | Boucherie | boucherie | Nom = boucherie, charcuterie, traiteur, crèmerie, fromagerie |
| | Poissonnerie | poissonnerie | Nom = poissonnerie |
| | Fleuriste | fleuriste | Nom = fleuriste |
| | Autres petits commerces | autres commerces | Nom = tabac, journaux, presse, vidéo club, librairie, opticien, pharmacie, décoration, informatique, bureau de change, primeur, alimentation, épicerie, Naturalia, Bio coop, bijouterie, boutique, auto-moto, auto-école |

| | | | |
|-------------|---|-------------------------------|--|
| | Grands magasins et grandes surfaces spécialisées | Grandes surfaces spécialisées | Nom = liste de grandes enseignes et centre commercial |
| | Grandes surfaces de vente des produits alimentaires | super et hypermarché | Nom = liste grandes enseignes |
| | Discount alimentaire | discount alimentaire | Nom = liste grandes enseignes) |
| Éducation | Établissements petite enfance et scolaire | crèche et école | Nom = crèche, multi-accueil, garderie, centre aéré, école, groupe scolaire, primaire, lycée, collège, CFA, OGEC de, BTP CFA, DDEC, handicapés) |
| | Établissements d'enseignement supérieur | enseignement supérieur | Nom = faculté, université, IUFM, IUT |
| Hébergement | Tourisme hôtelier | hôtel | Nom = hôtel, auberge |
| | Tourisme camping | camping | Nom = camping |
| | Maison de retraite | maison retraite | Nom = retraite, EPAHD, senior, âgées + liste départementale si disponible |
| | Aire d'accueil gens du voyage | Aire accueil | Nom = gens du voyage, terrain d'accueil |
| | Autre hébergement collectif | hébergement collectif | Nom (pénitentiaire, prison, maison d'arrêt, Crous, amis de l'icam, Cité U, Adelis, foyer, hébergement, orphelinat, refuge, CHRS, petites sœurs des pauvres, charbon blanc/bleu, ADAPEI, une famille un toit, APAJH, ANJORRANT, accueil d'urgence, SENEVE, 102 GAMBETTA, logement urgence, centre St Yves, centre Ametis) |

Tableau 2 : Exemple de lexiques associant des mots clefs avec des catégories d'usagers pour la typologie INRAE – FNCCR

2.6. STATISTIQUES DESCRIPTIVES

Les statistiques descriptives peuvent ensuite être calculées idéalement pour chaque catégorie d'usagers et dans chaque commune. Outre la moyenne, il est utile de calculer la médiane et si possible de construire une représentation graphique de type "boîtes à moustaches" pour bien visualiser la distribution des valeurs. Ce type de graphique résume graphiquement les principales données statistiques : le minimum, les quartiles et le maximum. Il permet en outre de visualiser les valeurs extrêmes, représentées par des points noirs sur la droite de la Figure 4 ci-après. Si ces valeurs sont nombreuses, on pourra aller les consulter pour vérifier qu'il s'agit ou pas d'erreur de classification. Ces graphiques sont en effet très utiles pour explorer le résultat de la classification automatique.

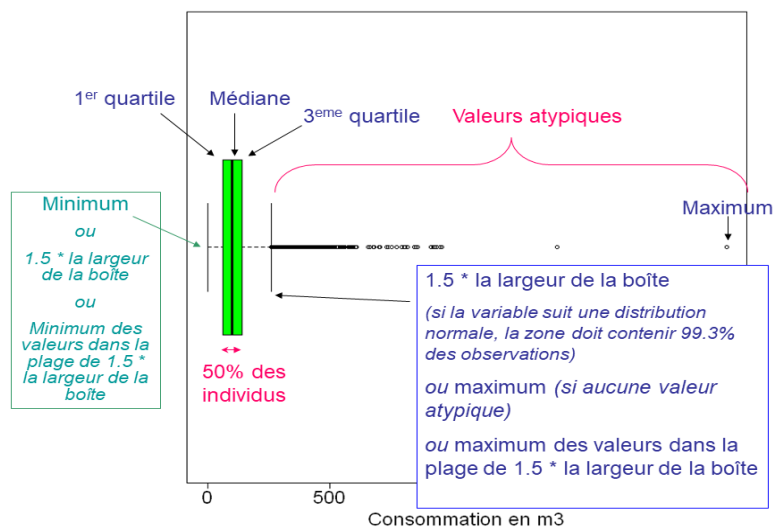


Figure 4 : Présentation du principe d'une boîte à moustache

2.7. CLASSIFICATION DES ABONNÉS SELON LEUR NIVEAU DE CONSOMMATION

Au sein de chaque type d'abonné, il peut être intéressant de répartir les abonnés en 3, 4 ou 5 classes de consommation (de petit à grand).

Cela permet, par exemple, de tracer des courbes d'évolution des abonnés domestiques ayant une petite (< 60 m³), moyenne (60-140 m³) et grosse consommation (> 140 m³). L'objectif est d'identifier si l'une de ces classes de consommateurs évolue différemment des autres (par exemple, les gros consommateurs peuvent être caractérisés par une baisse plus forte que les autres, car ce sont eux qui ont mis en place des mesures d'économies d'eau).

La définition des bornes de petite / moyenne / grosse consommation doit alors être réalisée à dire d'expert pour chacun des types.

2.8. TEST DE TENDANCE D'ÉVOLUTION DE LA CONSOMMATION EN EAU

Il est également possible de chercher à caractériser la **tendance individuelle** de la consommation de chacun des milliers d'abonnés inclus dans les fichiers.

Pour cela, on applique à chaque abonné (donc au total à plusieurs milliers d'abonnés) un test statistique non paramétrique comme le test de Mann Kendall. Ce test permet de détecter une tendance décroissante, croissante ou stationnaire. La tendance peut ne pas être significative, auquel cas le test n'est pas conclusif. Le test donne, pour chaque individu, un résultat associé à une probabilité : l'individu X présente une tendance décroissante avec une probabilité de 0,9 ou de 0,3.

L'utilisation de ce test nécessite idéalement de disposer d'une série chronologique d'au moins 8 à 10 années. La tendance ne peut donc pas être estimée pour des abonnés ayant été présents pendant une période plus courte.

Une fois le calcul réalisé pour chaque abonné, il peut être analysé la proportion de tendances décroissantes ou croissantes. On peut avoir 20 % des abonnés en décroissance forte, 75 % stable et 5 % en croissance forte par exemple. Alternativement, on pourrait avoir 80 % de décroissance faible et 20 % sans tendance.

Test de Mann Kendall - le principe

- Analyse de la tendance pour chaque consommateur
- Détection d'une non stationnarité (croissante, décroissante)
- Incertitude = seuil de confiance 5%
- Non paramétrique (pas d'hypothèses sur la distribution, pas d'ajustement de coefficients)

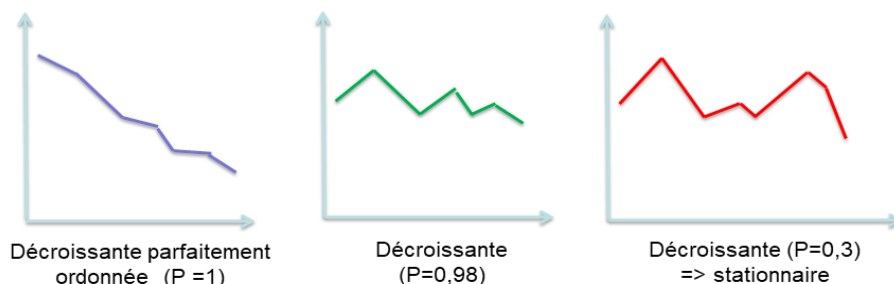


Figure 5 : Présentation du principe du test de tendance de Mann Kendall.

2.9. GRAPHIQUES ET CARTES

L'ensemble des résultats obtenus doivent pouvoir être spatialisés sur une carte des communes du périmètre du service. Cela permet de rapidement visualiser les différences et de chercher à comprendre les facteurs géographiques (proximité du littoral ou du centre urbain) ou socio-économiques (communes à revenus moyens élevés) qui les expliquent.

2.10. EXPORTATIONS DE L'ENSEMBLE DES RÉSULTATS SOUS DEUX FORMATS

Il est utile de demander au prestataire de terminer son analyse par une exportation des données initiales enrichies de toutes les nouvelles variables (type, tendance) sous un format Excel. Ceci permettra à la collectivité de réaliser des analyses plus poussées par secteur géographique ou par type d'usagers via un outil facilement accessible et manipulable. Cela permettra éventuellement aussi de réaliser une vérification manuelle des abonnés qui n'ont pas pu être classés automatiquement et de réintégrer le résultat de l'analyse (et notamment de la classification par type d'usage) dans le logiciel de facturation (Figure 6).

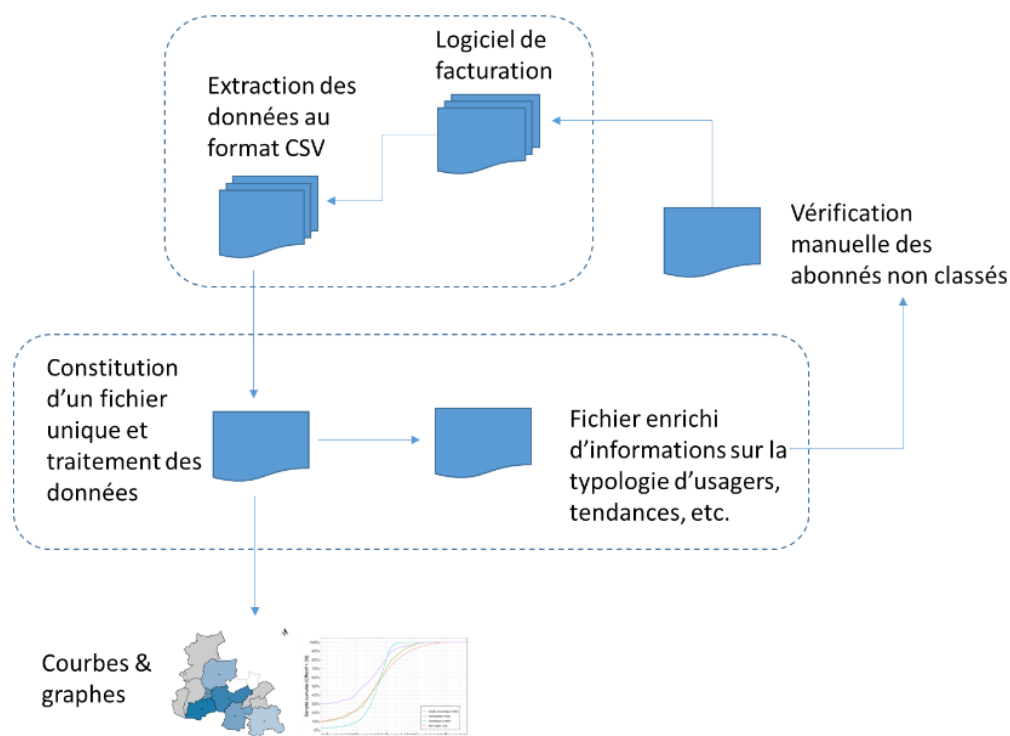


Figure 6 : Représentation schématique de la circulation des données au cours de l'analyse.

2.11. LES OUTILS POUVANT ÊTRE UTILISÉS

Il n'existe pas à ce jour d'outil commercial (logiciel) permettant de réaliser ce type d'analyse de manière standardisée.

2.11.1. Outils issus de la recherche

Les centres de recherche (INRAE en particulier, en collaboration avec le BRGM) ont développé un outil précurseur depuis une dizaine d'années. Cet outil a été développé en utilisant le logiciel libre R. Les codes ne sont pas interfacés, ce qui rend leur utilisation complexe pour un non-spécialiste de R. Ils peuvent néanmoins être mobilisés en sollicitant directement INRAE via un conventionnement (Contact : marielle.montginoul@inrae.fr).

L'outil INRAE s'appuie sur des algorithmes conçus dans le cadre de différents projets de recherche : la convention de recherche avec la Communauté Urbaine de Nantes pour le "suivi et l'analyse des consommations d'eau des usagers de l'eau" (2005-2012) ; le projet ANR Villes Durables « Eau et 3E » (<https://eau3e.hypotheses.org/le-projet>) (2009-2013) ; enfin la convention de recherche avec Nantes Métropole, Veolia, Montpellier Méditerranée Métropole, Syndicat Mixte Garrigues Campagne, ville de Metz (2014-2019) sur la mise en place d'un observatoire des consommations en eau. Ces différentes actions de recherche ont donné lieu à un certain nombre de publications, essentiellement sous forme de rapports et de communications à destination d'un public opérationnel (Favre et Montginoul, 2018, 2020 ; Montginoul et al., 2013a, b ; Montginoul et al., 2010, 2013c ; Montginoul et al., 2017).

Le code informatique développé permet de réaliser les différentes étapes écrites dans la section précédente et de les traduire ensuite sous forme d'un rapport automatisé produit sous un format Word (donc qui permet ensuite une mise en forme spécifique par l'utilisateur final). Ce rapport édité

automatiquement s'accompagne également d'un fichier Excel, présentant les différentes statistiques, utilisables également ensuite par l'utilisateur.

2.11.2. Autres outils

Les sociétés fermières (SUEZ, VEOLIA, SAUR) développent leurs propres programmes de recherche et des outils internes visant à répondre à cette demande croissante des collectivités territoriales. L'outil SAUR a été utilisé dans le cadre du projet « De l'Eau pour Demain » (Boisson et al, 2024), à titre exploratoire.

3. Visualisation des résultats

Cette section présente une large gamme de types de graphiques, courbes et cartes qui peuvent être produits pour visualiser les résultats d'une analyse des fichiers de facturation.

Il s'agit d'exemples pouvant inspirer la rédaction d'un cahier des charges. Les figures s'appuient sur l'analyse des données d'un secteur du centre Morbihan, géré par la SAUR, pour le compte de la collectivité Eau du Morbihan.

3.1. RÉSULTAT DE LA CLASSIFICATION

3.1.1. Analyse globale des effectifs et des volumes facturés

La répartition des abonnés peut être visualisée en pourcentage de l'effectif d'abonnés et en pourcentage du volume facturé total. Cette représentation peut être réalisée pour deux niveaux de classification, comme dans les graphiques ci-dessous : le premier décompose la consommation en quatre grands usages (domestique, activités économiques, administrations et non classés) (Figure 7). Le second décompose plus finement les activités économiques (Figure 8).

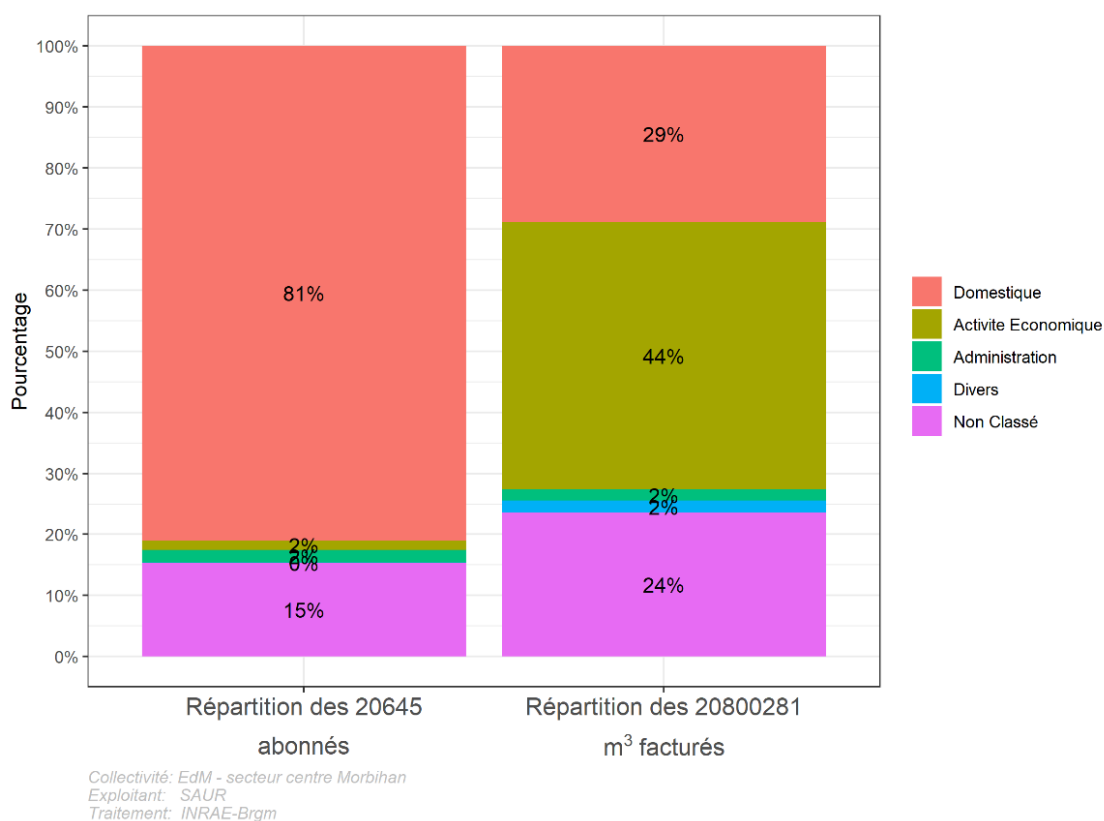


Figure 7 : Répartition des consommateurs par catégorie d'abonnés entre 2011 et 2021 (source : INRAE-Brgm)

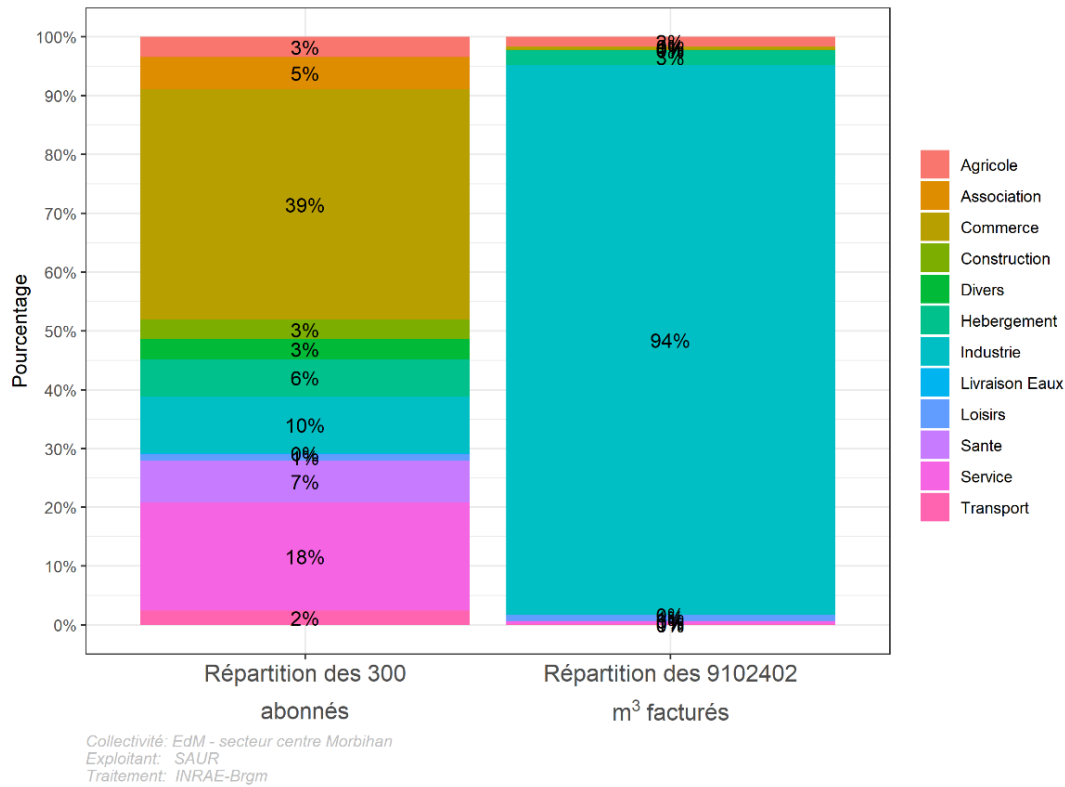


Figure 8 : Répartition des abonnés économiques par activité entre 2011 et 2021 (source : INRAE-Brgm)

3.1.2. Analyse spatiale

Un ensemble de cartes peut être réalisé pour représenter l'effectif de chaque catégorie d'abonnés et le pourcentage des clients qu'ils représentent. Quelques exemples sont présentés ci-dessous pour la période 2011-2021 (Figure 9). Le même type de cartes peut être produit pour chaque année.

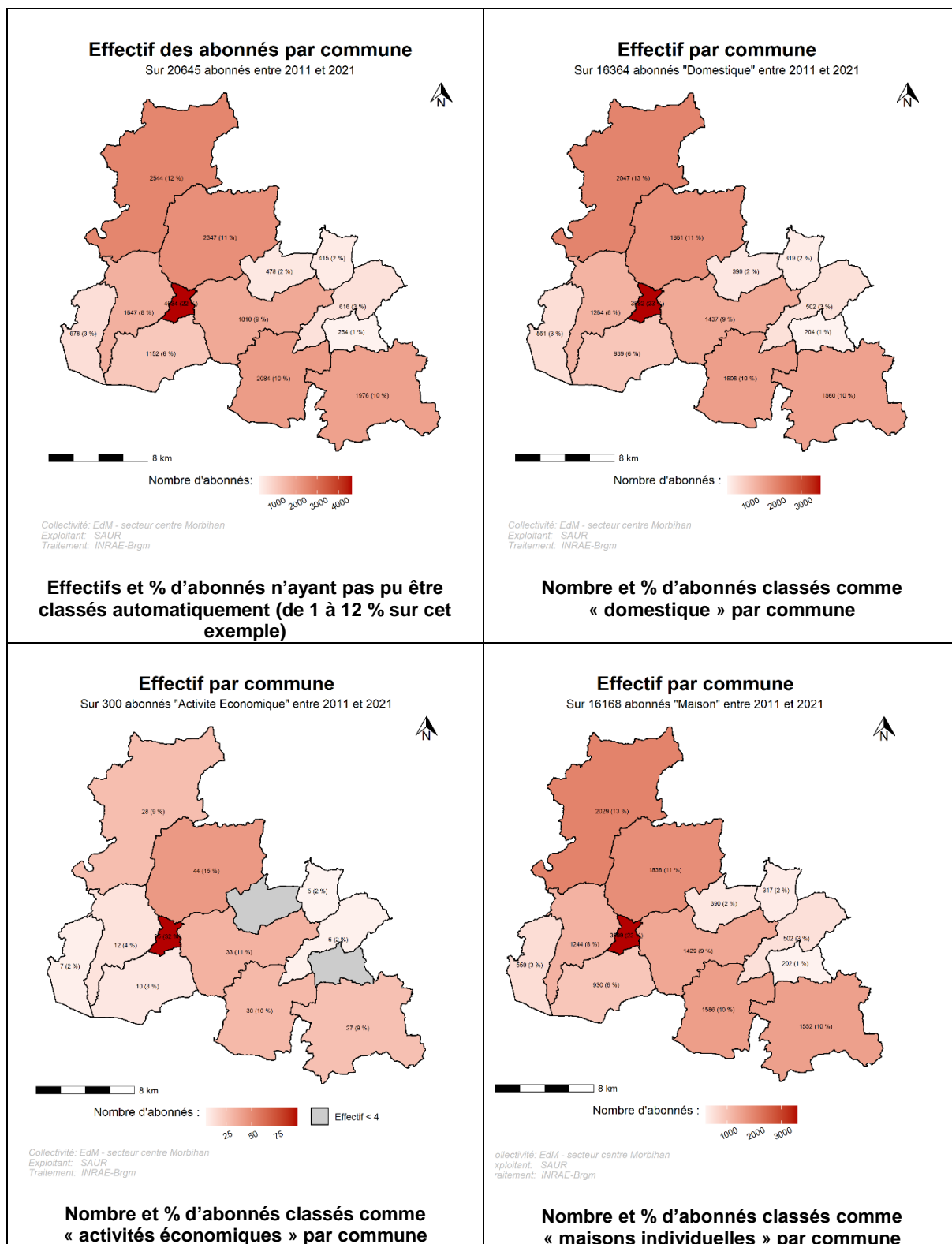


Figure 9 : Carte du nombre d'abonnés par grande catégorie (Total haut à gauche, domestiques haut droite, activités économiques bas gauche et maisons individuelles bas droite)

3.2. ANALYSE DES RATIOS DE CONSOMMATION UNITAIRE PAR TYPE

3.2.1. Analyse globale

Les ratios unitaires de consommation peuvent être présentés sous une forme graphique (Figure 10) montrant la moyenne et la médiane. La médiane permet de s'affranchir des valeurs extrêmes. C'est le cas dans le graphique ci-dessous, où une consommation industrielle très importante tire la moyenne vers le haut (55 000 m³) alors que la moitié des abonnés industriels consomment moins de 192 m³ (médiane).

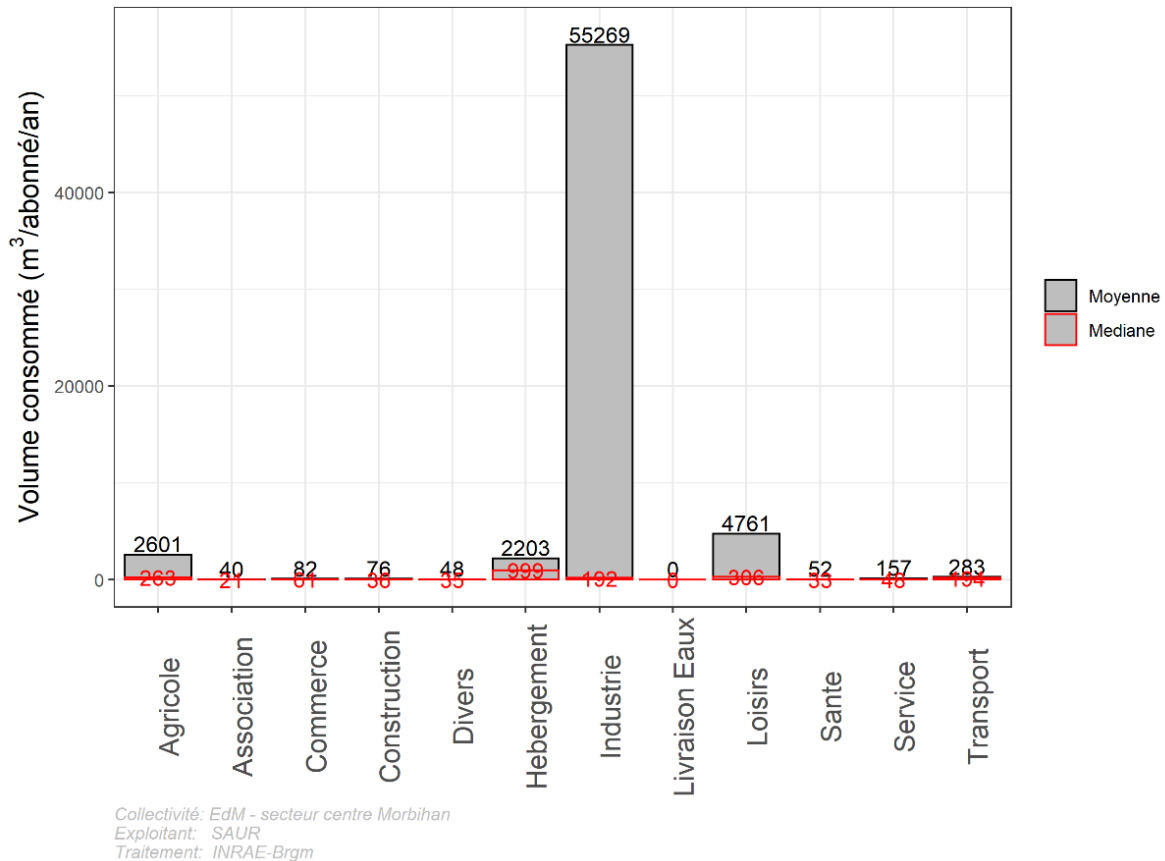


Figure 10 : Consommation moyenne et médiane par catégorie d'activité économique (période 2011-2021) – source INRAE-Brgm

3.2.2. Analyse spatiale

Une représentation cartographique des consommations moyenne et médiane par catégorie d'utilisateur est un bon support d'exploration de la variabilité spatiale des consommations. La Figure 11 illustre l'intérêt de ce type de représentation spatiale des résultats.

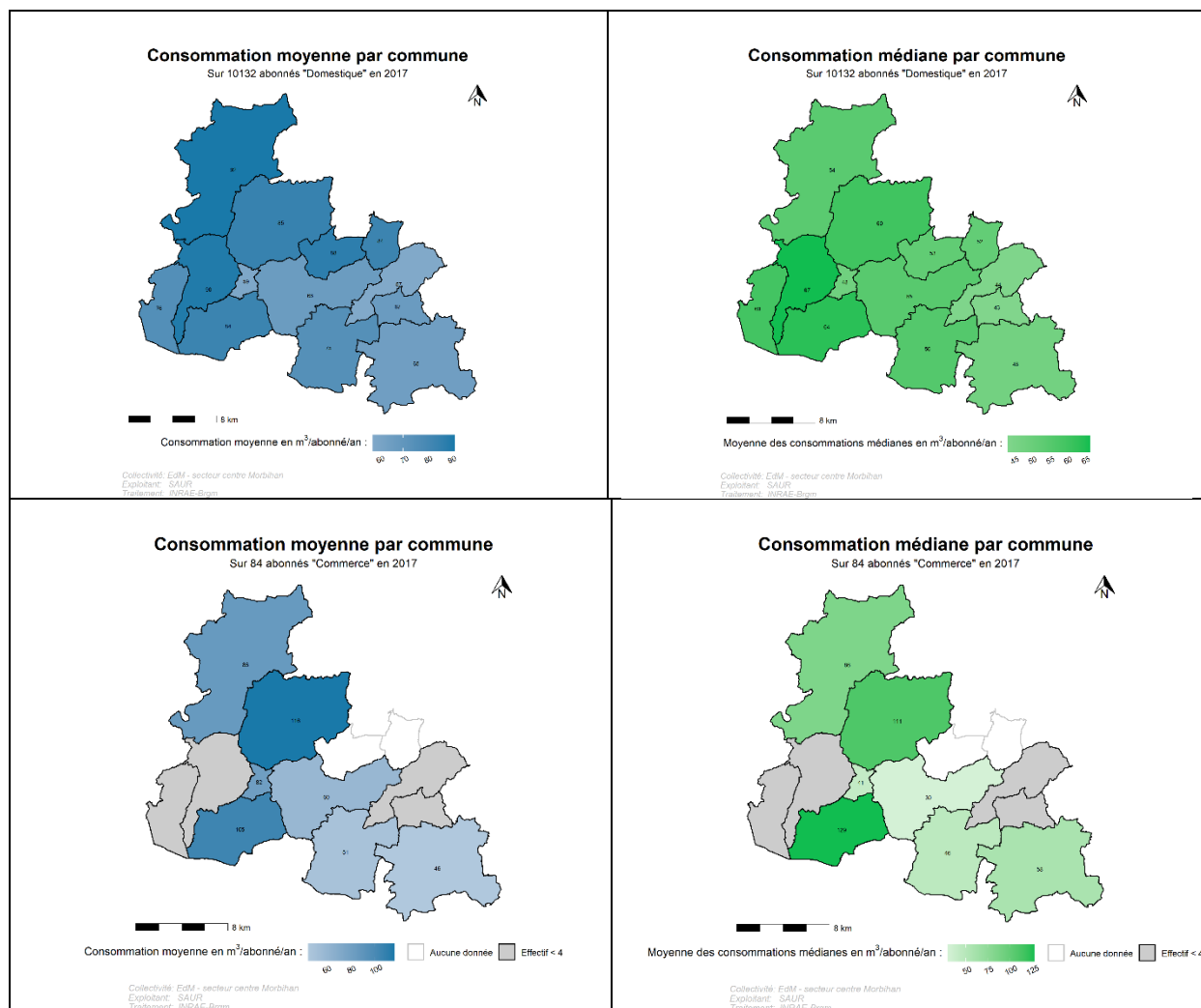


Figure 11 : Représentation spatiale de la consommation moyenne / médiane pour différents types d'abonnés. En haut abonnés domestiques, en bas, commerces. Source : INRAE-Brgm

3.3. ANALYSE DES DISTRIBUTIONS STATISTIQUES

Une analyse de la distribution statistique des valeurs de consommation au sein de chaque catégorie d'usagers est utile pour appréhender la diversité des consommateurs, mais aussi pour repérer d'éventuelles erreurs de classement avec l'algorithme automatique.

3.3.1. Boîte à moustache

La première représentation possible de la distribution statistique est le graphique dit en « boîte à moustache ». Ce graphique montre comment se répartissent les valeurs de consommation autour de la médiane (barre verticale noire) et de la moyenne (point rouge). Le rectangle gris contient 50 % des observations. La plage représentée par les traits noirs horizontaux contient 90 % des valeurs. Les points qui se situent au-delà de cette barre horizontale ne représentent donc que 5 % des valeurs les plus basses et 5 % des valeurs les plus hautes.

Le premier graphique ci-dessous (Figure 12), décrivant la distribution de la consommation des abonnés domestiques, met en évidence plusieurs points :

- Une consommation moyenne et médiane des maisons supérieure aux appartements ;
- Pour les appartements, 90 % des valeurs sont inférieures) 130 m³/an et il y a peu de valeurs très élevées ;
- Pour les maisons, 90 % des valeurs sont inférieures à 175 m³/an, mais il y a de nombreuses valeurs très élevées ; ces valeurs mériteraient d'être analysées manuellement pour en comprendre la cause (par exemple présence d'un grand jardin maraîcher, ou d'une activité économique conduite au sein de la maison) ; il est également possible que ces valeurs révèlent une erreur de classification automatique, qui doit alors être corrigée manuellement.

Ces boîtes à moustaches peuvent être réalisées pour toutes les catégories d'usage définies dans la typologie et pour chaque année (celles présentées ici sont pour l'année 2017).

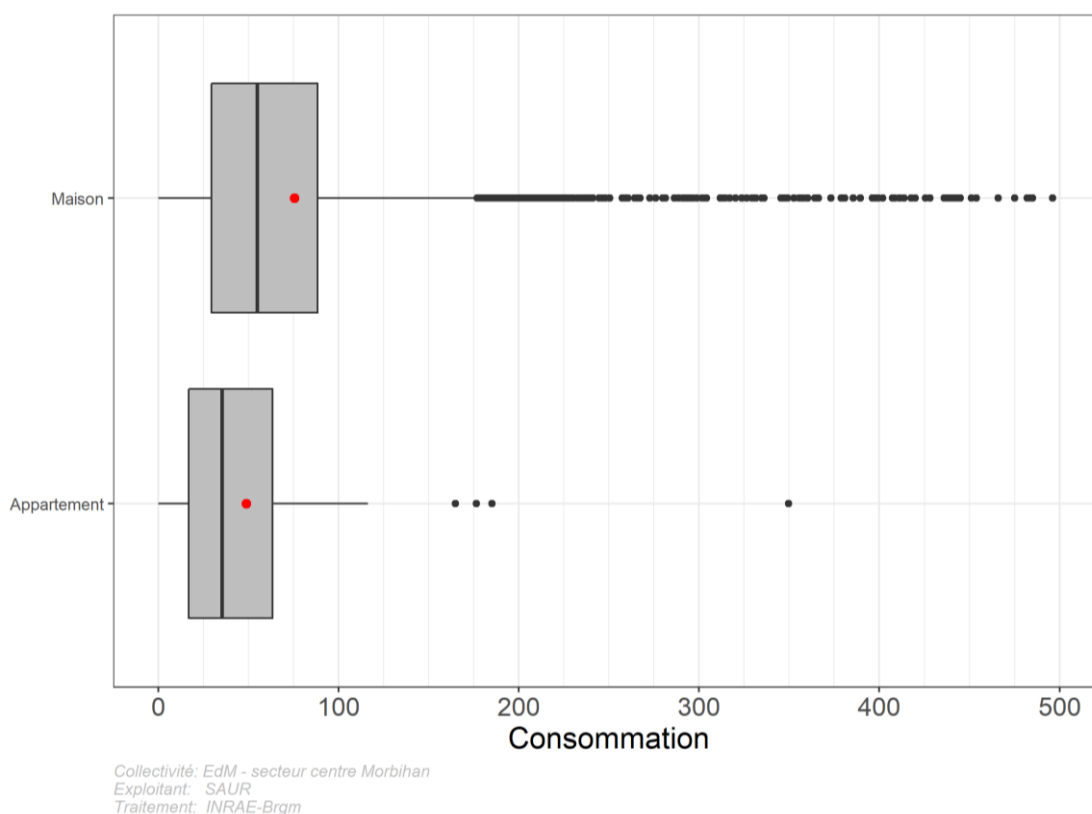


Figure 12 : Représentation de la distribution statistique des consommations pour les abonnés domestiques en maison individuelle (haut) et en appartement (bas) en 2017 – graphique dit « en boîte à moustaches » - Source : INRAE-Brgm

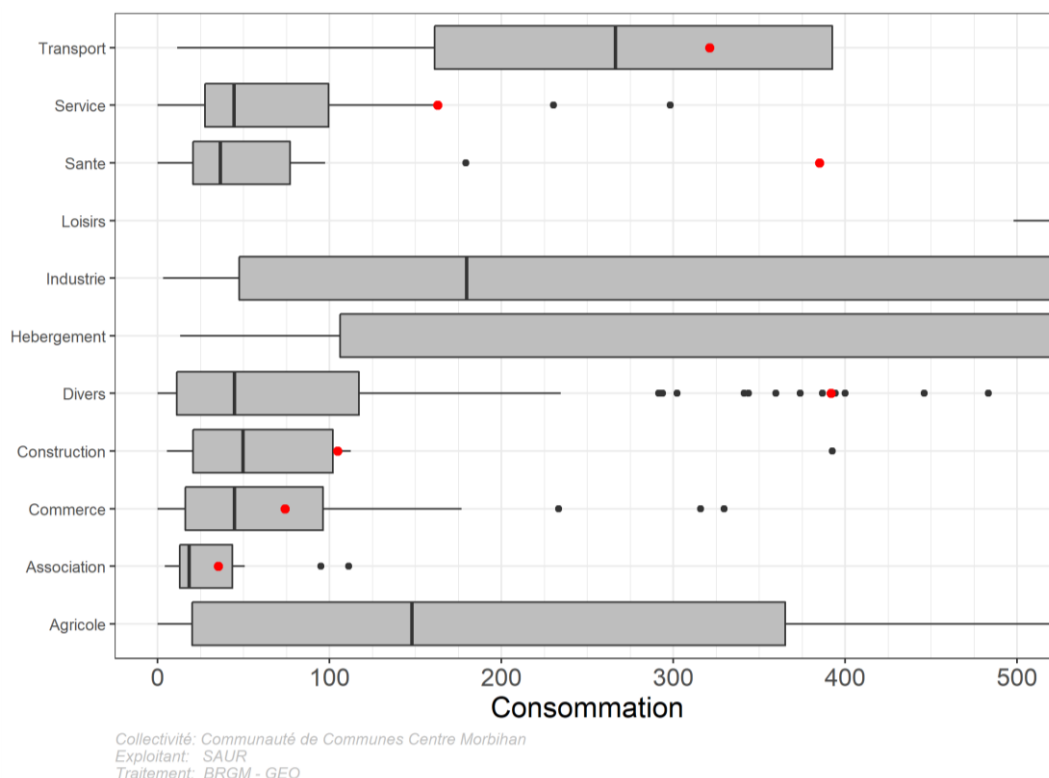


Figure 13 : Représentation de la distribution statistique des consommations pour les catégories d'activités économiques en 2017 – graphique dit « en boîte à moustaches » - Source : INRAE-Brgm

3.3.2. Courbes de densité

La seconde représentation possible est la courbe de distribution cumulée. Elle permet de lire simplement quel est le pourcentage des abonnés qui consomment moins qu'une valeur donnée. Par exemple, sur le graphique suivant (Figure 14), on peut voir que :

- Au point A, on a 80 % des abonnés domestiques qui consomment moins de 90 m³/an ;
- Au point B, 95 % des abonnés domestiques consomment moins de 175 m³/an.

Ce type d'information est très utile si l'on veut par exemple fixer des paliers pour une tarification progressive.

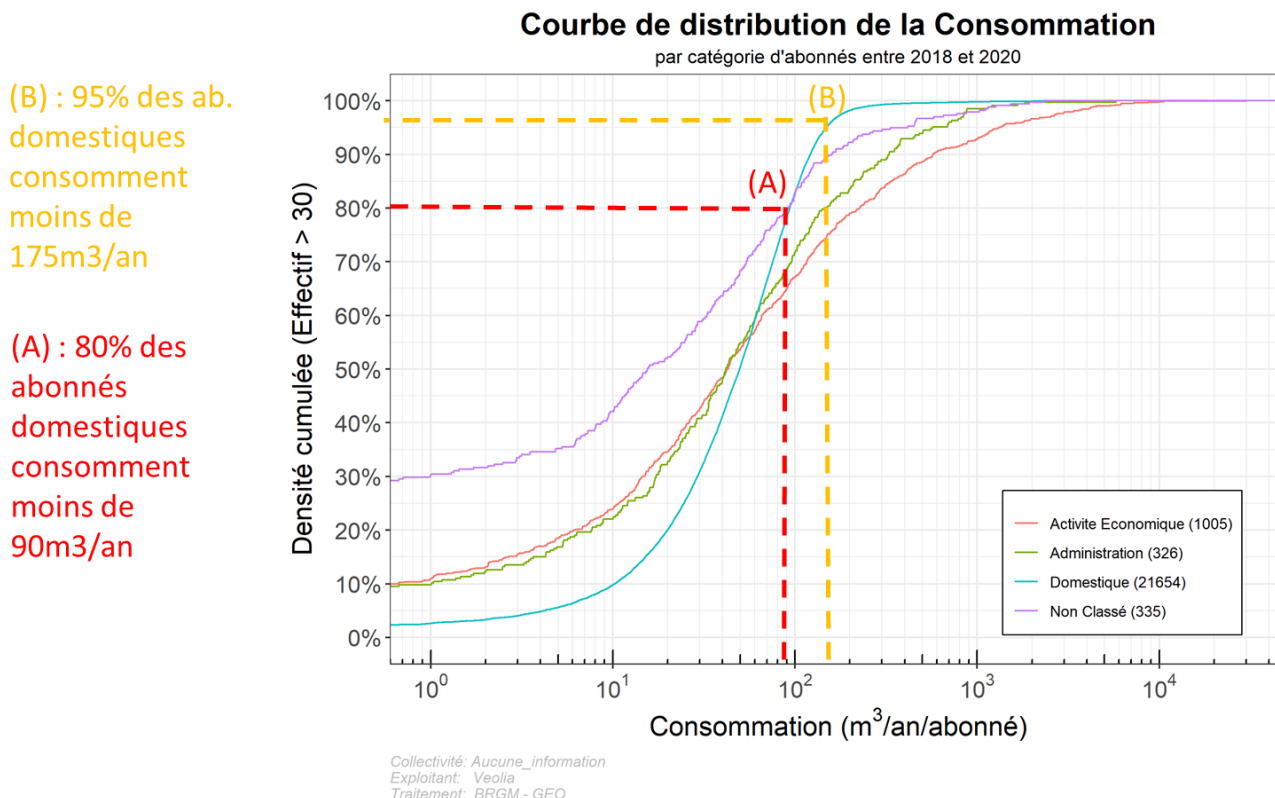


Figure 14 : Représentation de la distribution de la consommation par une courbe de densité cumulée. Chaque courbe de couleur représente un grand type d'abonnés. Source INRAe-Brgm.

3.4. ANALYSE DE LA DYNAMIQUE TEMPORELLE

3.4.1. Évolution des effectifs et des volumes facturés

La Figure 15 décrit l'évolution de la clientèle économique dans le secteur étudié. Elle montre une stabilité de la composition de la clientèle, avec 82 % d'abonnés domestiques sur toute la période. Les données de la dernière année (2021) sont incomplètes, car une grande partie des compteurs n'a pas été relevée à la date de récupération des données.

L'évolution de la clientèle peut aussi être décrite en volume facturé par catégorie d'abonnés (Figure 16). Ce graphique montre également que la classification automatique des abonnés est une source majeure d'incertitude : quelques gros consommateurs qui arrivent en cours de période ne sont pas affectés à un type, faisant augmenter le volume non affecté. Cela invite à analyser manuellement ces abonnés non classés.

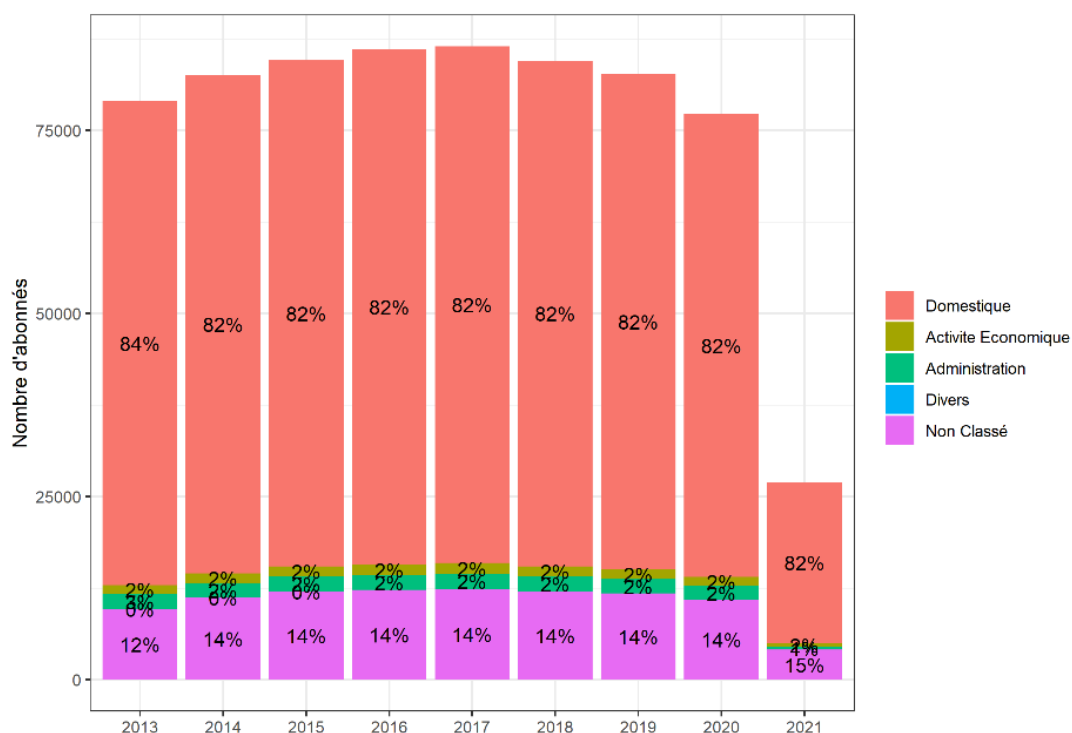


Figure 15 : Evolution de la clientèle en nombre d'abonnés (Source INRAe-Brgm)

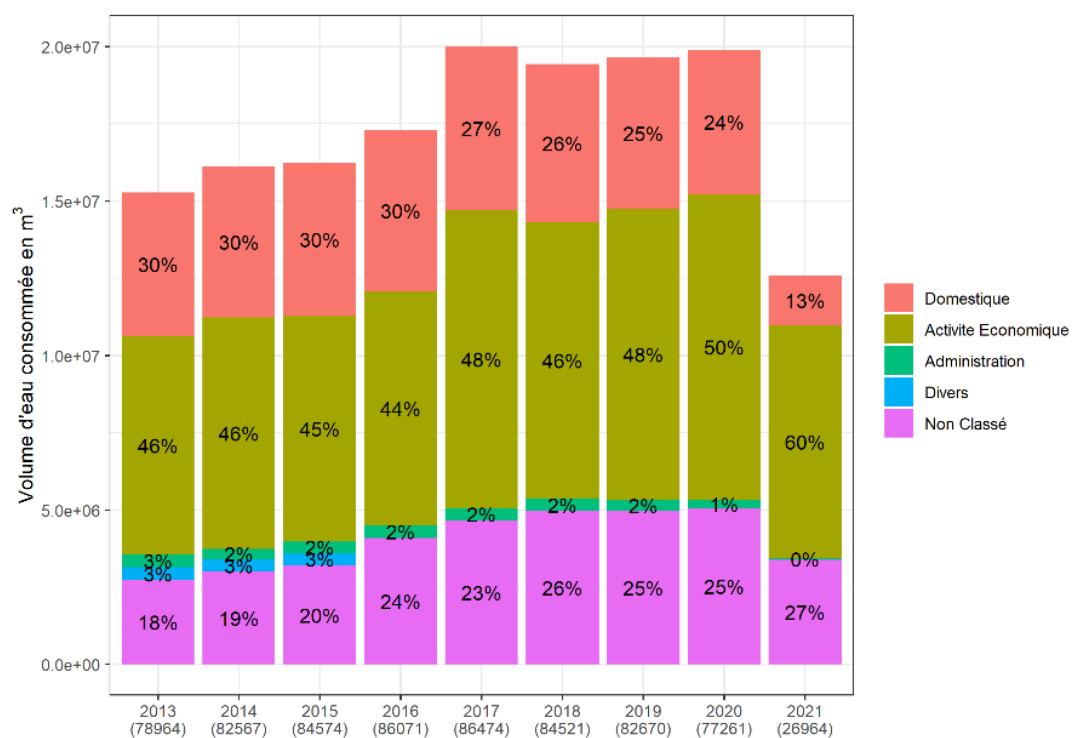


Figure 16 : Evolution des volumes facturés par catégorie d'abonnés (source INRAe-Brgm)

L'évolution des volumes facturés peut enfin être analysée pour chaque catégorie d'usage de la typologie. Sur la Figure 17, on voit l'évolution des volumes facturés aux usagers économiques, qui ne présente pas d'évolution notable.

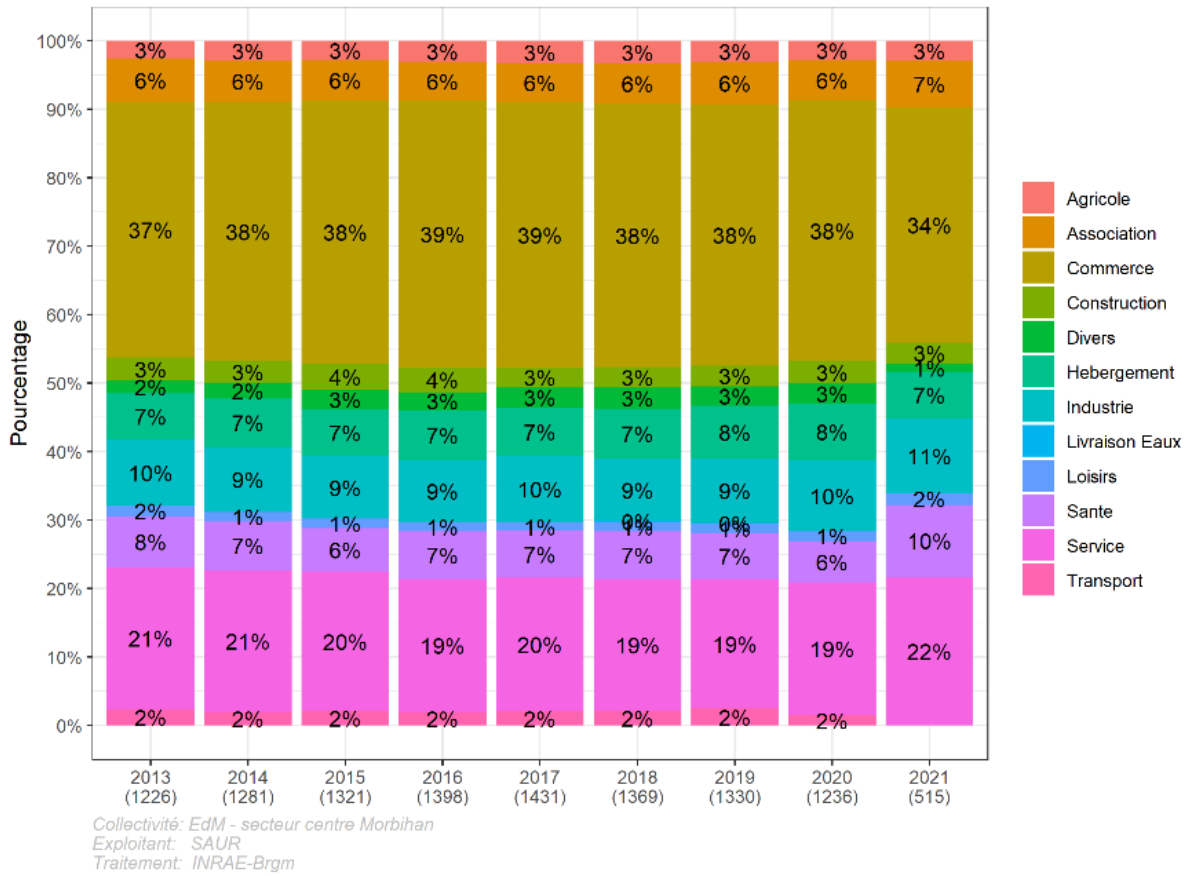


Figure 17 : Evolution de la répartition des abonnés économiques par type d'activité (source INRAE-Brgm)

3.4.2. Analyse des tendances globales

La courbe la plus utile est celle présentant l'évolution du ratio unitaire de consommation moyenne (ou médiane) par type d'usagers. Celles des abonnés domestiques résidant en appartement et en maison individuelle sont présentées ci-dessous à titre d'illustration (respectivement Figure 18 et Figure 19).

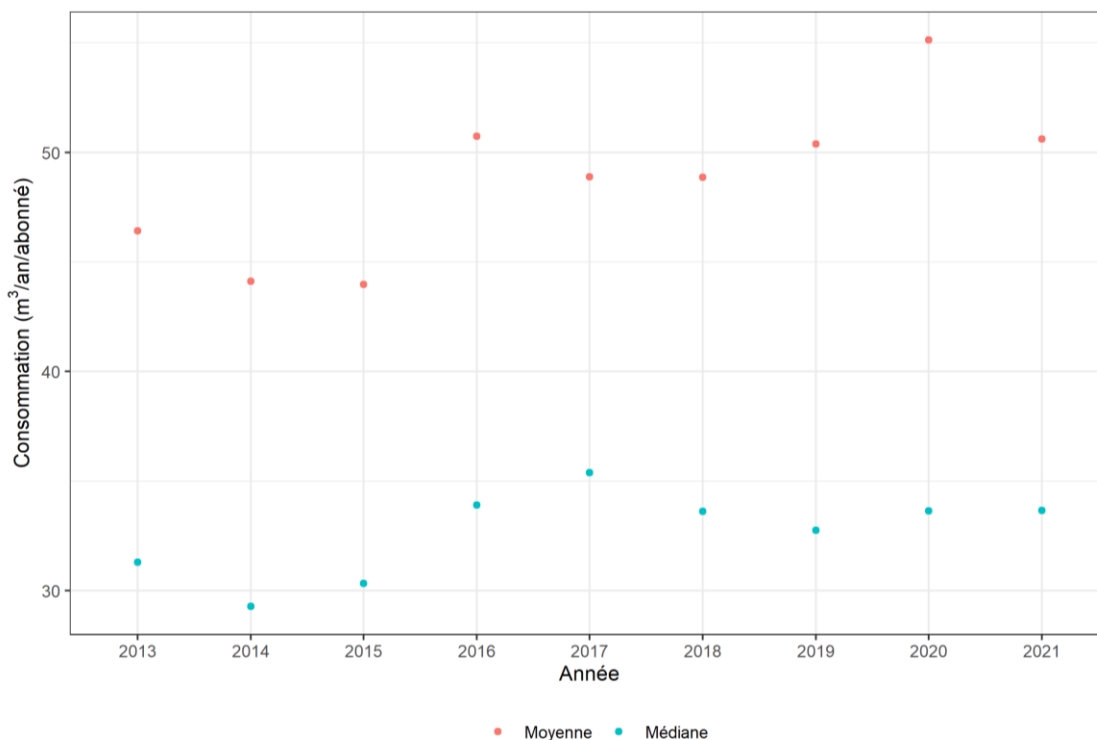


Figure 18 : Evolution de la consommation moyenne et médiane des abonnés domestiques résidant en appartement (source : INRAE-Brgm)

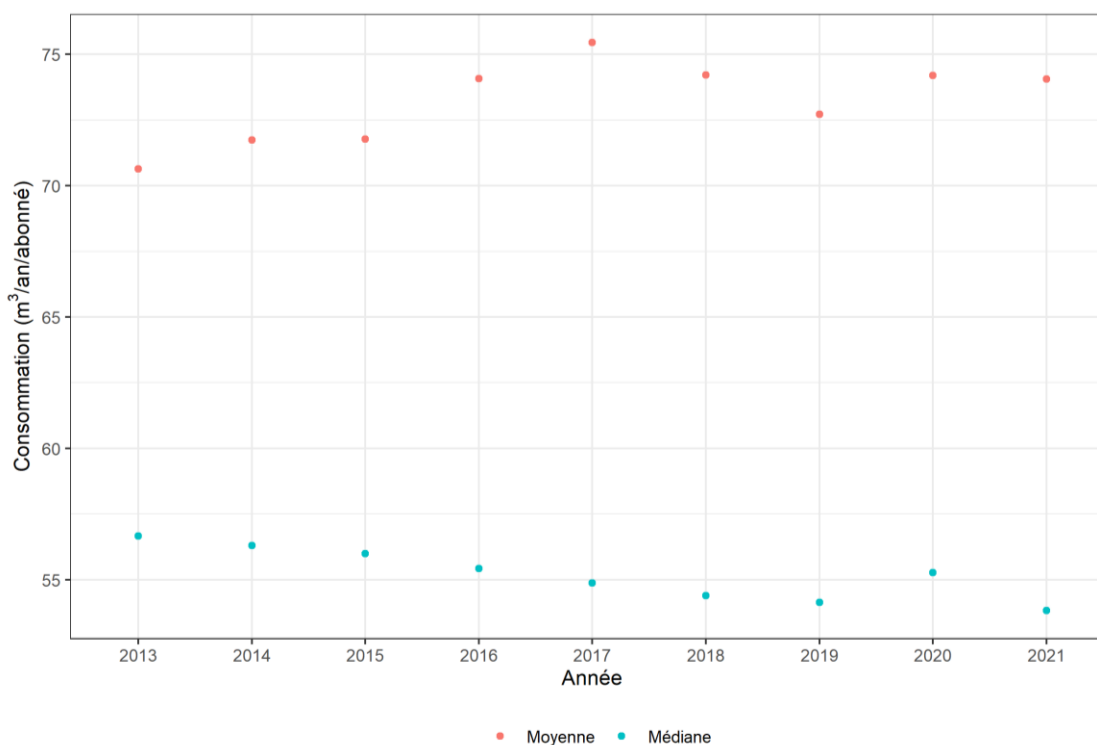


Figure 19 : Evolution de la consommation moyenne et médiane des abonnés domestiques résidant en maison individuelle (source : INRAE-Brgm)

La boîte à moustaches peut aussi être utilisée pour analyser plus finement l'évolution temporelle de la distribution des consommations. Sur le graphique ci-dessous (Figure 20), on constate par exemple que :

- les valeurs les plus élevées (> 100 000 m³) ont tendance à diminuer dans le temps (partie haute du graphique) ;
- la moyenne (points rouges) descend légèrement et la boîte grise contenant 50 % des valeurs se rétrécit, ainsi que la place des 90 % des valeurs (lignes grises) : il y a réduction progressive des écarts.

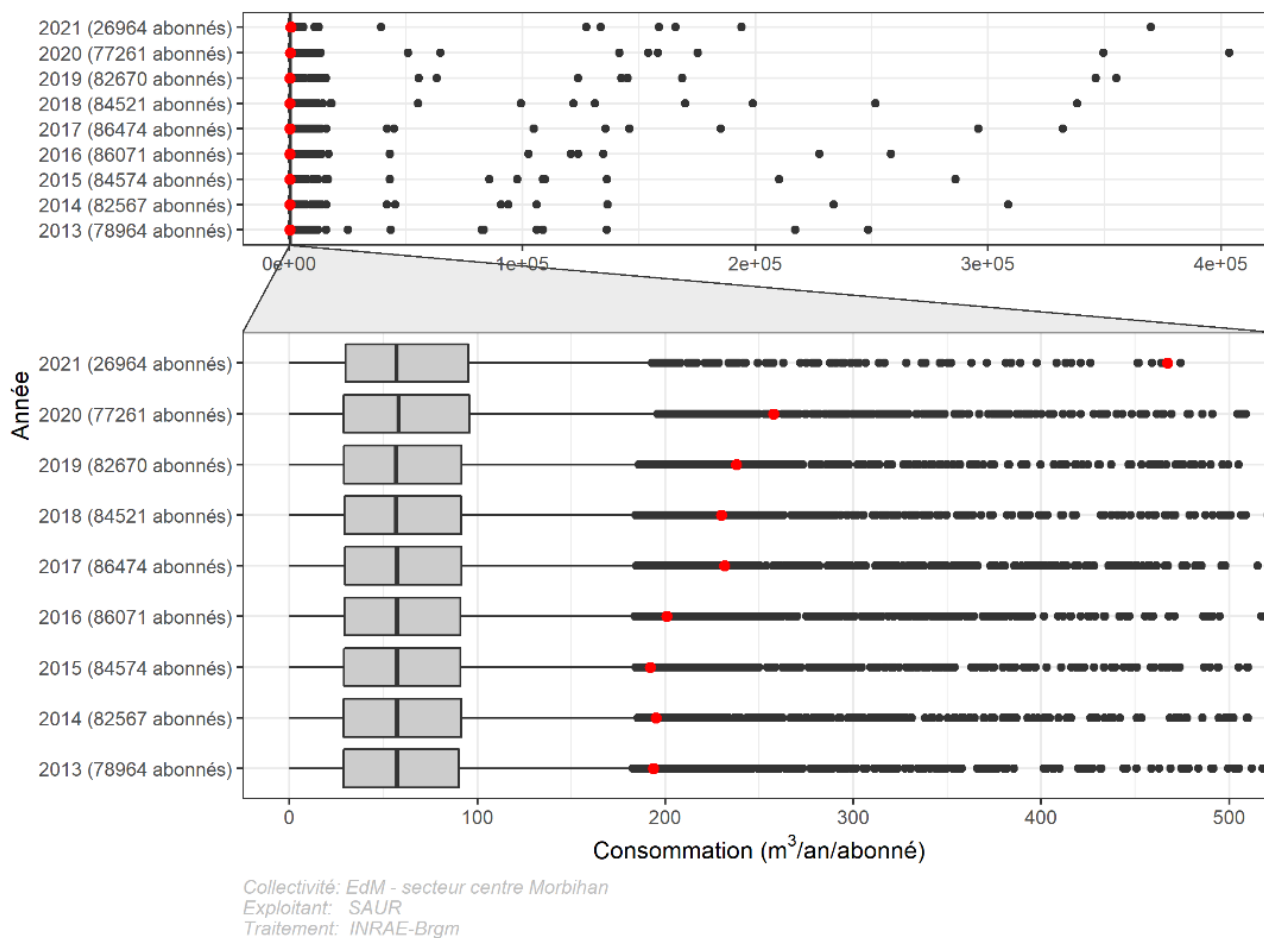


Figure 20 : Evolution de la distribution statistique des consommations représentée sous forme de boîtes à moustaches. Source : INRAe-Brgm

3.4.3. Analyse des dynamiques spatio-temporelles

La comparaison des cartes réalisées pour chaque année peut également aider à identifier des changements majeurs et à en rechercher la cause (Figure 21). Sur les deux cartes ci-dessous, on constate ainsi que la consommation des activités économiques est passée de 17 000 à 30 000 m³ dans une commune alors qu'elle a été divisée par deux dans une autre commune.

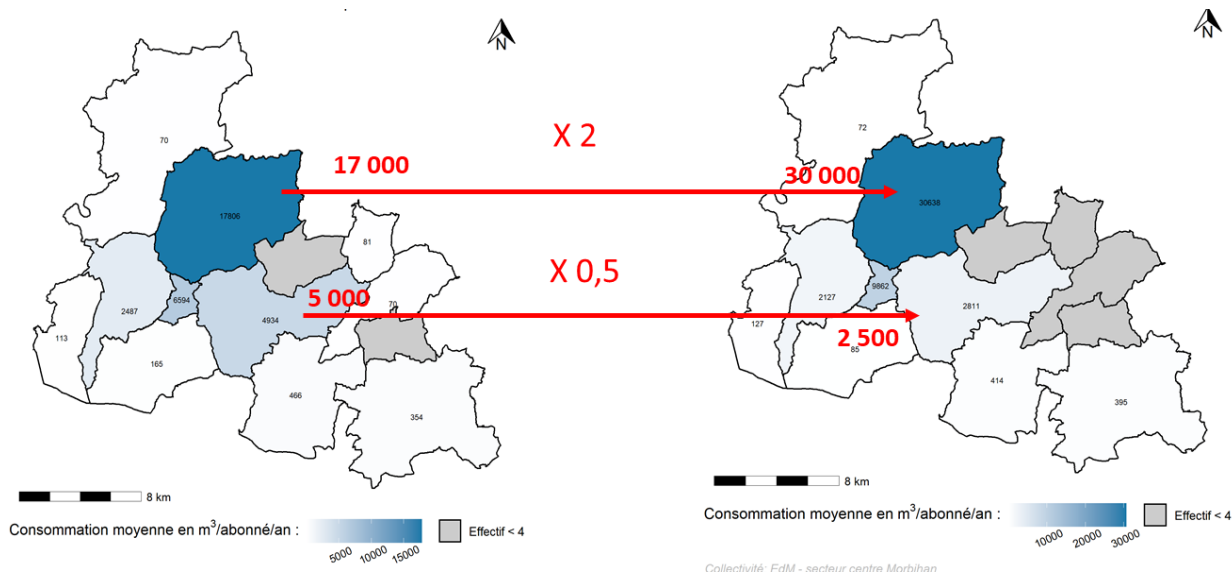


Figure 21 : Comparaison de la consommation moyenne des abonnés économiques entre 2015 et 2020. Source : INRAe-Brgm

3.4.4. Analyse des tendances individuelles

Les résultats du test de Man Kendal, qui détecte les consommateurs ayant une tendance croissante, décroissante ou stationnaire (voir section 2 du présent rapport), peuvent être représentés sous forme de carte (Figure 22). La tendance n'est pas calculée lorsqu'il n'y a pas assez de données (nombre d'années). Lorsque le test n'est pas conclusif, on dit qu'il n'y a pas de tendance (évolution semblant être aléatoire).

Les cartes de la Figure 22 montrent que l'on a globalement plus d'abonnés qui présentent une tendance croissante que décroissante, avec un écart pouvant être de 5 % dans certaines communes (par exemple 29 % croissant, 24 % décroissant).

La Figure 23 montre que la consommation moyenne des 3755 abonnés domestiques résidant en maison individuelle ayant une tendance croissante (triangles rouges orientés vers le haut) passe de 70 m³ en moyenne à 110 m³ entre 2013 et 2020. Ceux présentant au contraire une tendance décroissante (triangles verts, 4390) passent de 70 à 50 m³ sur la période étudiée.

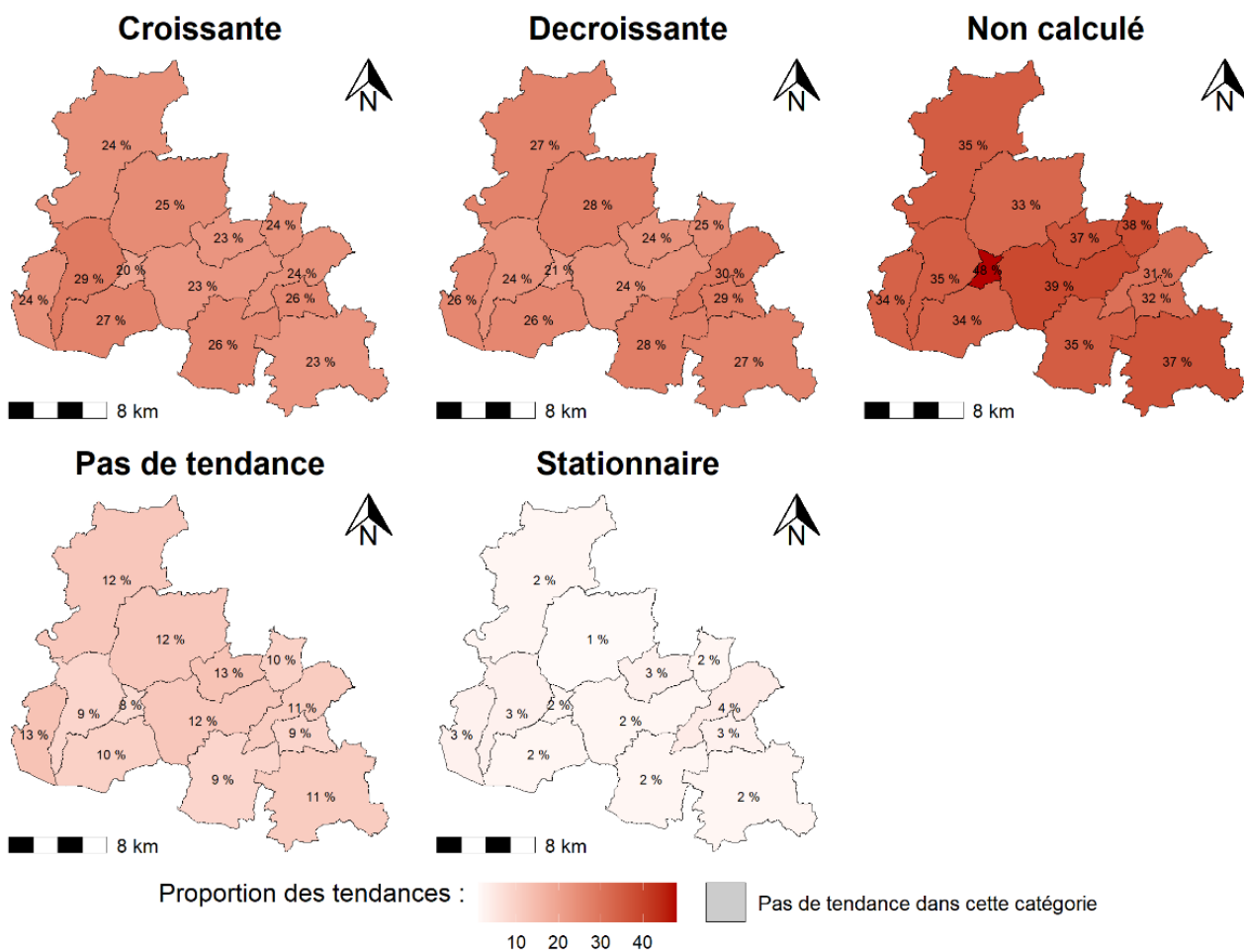
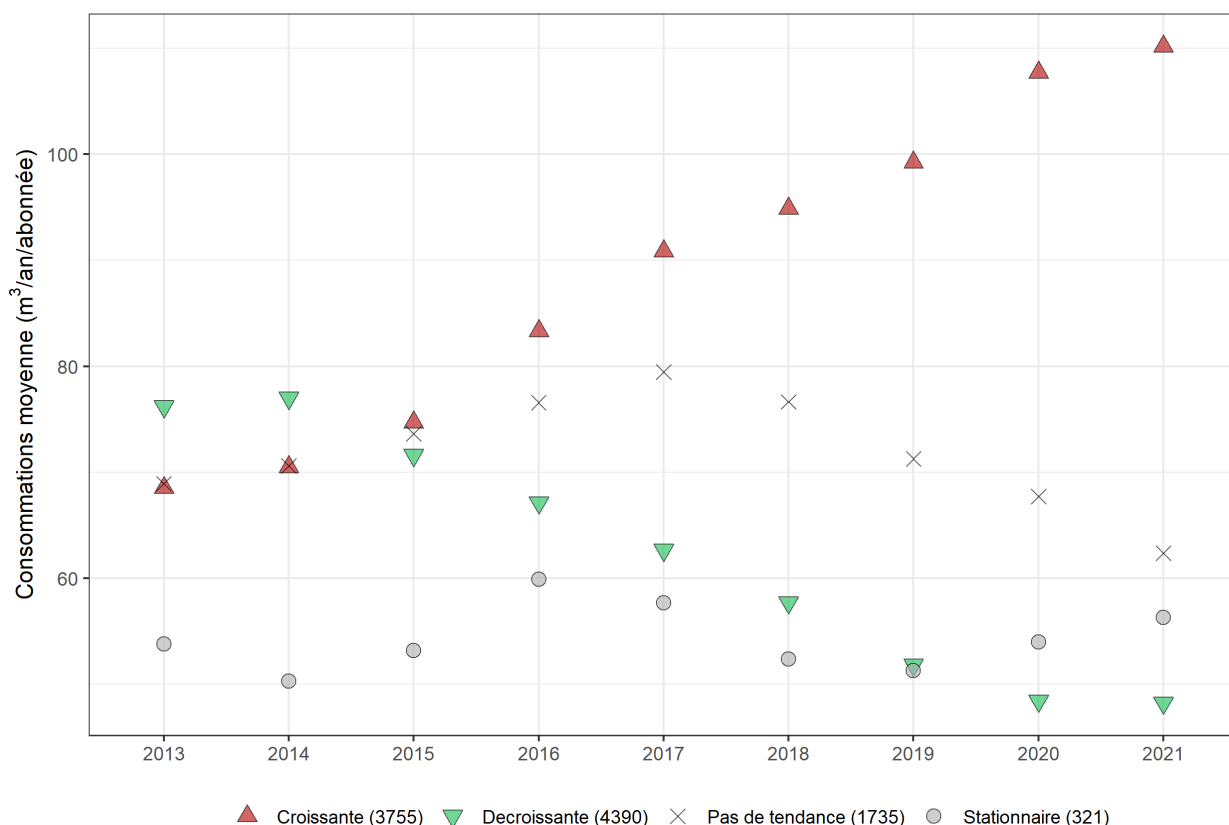


Figure 22 : Proportion d'abonnés présentant une tendance croissante, décroissante, ou stationnaire ou l'absence de tendance, représentées par commune. Calculé pour 20 470 abonnés entre 2013 et 2020. Source : INRAe-Brgm.



Collectivité: EdM - secteur centre Morbihan
Exploitant: SAUR
Traitement: INRAE-Brgm

Figure 23 : Consommation moyenne des abonnés domestiques présentant une tendance croissante, décroissante ou stationnaire sur la période 2013-2021.. Source : INRAE-Brgm.

3.5. ANALYSES COMPLÉMENTAIRES

À l'issue de la classification automatique, le gestionnaire peut réutiliser le fichier pour approfondir sa connaissance de certains usagers.

Dans le cas du secteur Centre Morbihan (géré par la SAUR, pour EDM), le BRGM a souhaité comprendre comment expliquer la variabilité de la consommation en eau observée dans les campings. Pour cela, il a été réalisé les étapes suivantes :

- il a d'abord été sélectionné les 29 compteurs du fichier pour lesquels l'usage a été classé « camping » ;
- une vérification manuelle montre que certains campings ont plusieurs compteurs, ces lignes ont donc été sommées pour n'avoir qu'une donnée de consommation par camping ;
- il a ensuite été recherché sur internet le nombre d'emplacements (tentes) et le nombre de locations (mobile homes) de chaque camping ;
- Enfin, il a été calculé un ratio de consommation par équivalent logement (emplacement tente + Mobil home).

Les résultats montrent des valeurs variant de 7 à 60 m³/emplacement et par an. Ces différences peuvent s'expliquer par la présence ou non d'un forage privé qui serait utilisé pour les usages extérieurs (arrosage), sanitaires (WC) et éventuellement pour les piscines. Des investigations de terrain pourraient être conduites sur la base de ces résultats. Une démarche identique pourrait être réalisée pour les exploitations agricoles dont une partie utilise aussi des forages individuels.

4. Conclusion

4.1. UNE SOURCE D'INFORMATIONS À VALORISER

Les fichiers de facturation sont une source d'informations riche, mais sous-exploitée. Ce rapport vise à montrer que leur analyse peut permettre :

- de mieux appréhender la part de la consommation totale imputable à différents types d'abonnés ;
- de calculer des ratios unitaires de consommation pour une grande diversité d'utilisateurs ;
- d'analyser l'hétérogénéité spatiale de ces ratios de consommation ;
- de caractériser les tendances d'évolution de la consommation des différents types d'abonnés.

L'analyse de ces fichiers peut s'avérer complexe et peut être sous-traitée à des consultants spécialistes de l'analyse de données ou exigée des délégataires dans le cadre des contrats.

4.2. UNE ANALYSE À RÉALISER EN ROUTINE

Cette analyse peut être réalisée comme une étude ponctuelle, mobilisant les données de facturation de 5 à 10 ans. Elle aidera dans ce cas à dresser un état des lieux de la situation, par exemple au moment de la réalisation d'un schéma directeur.

Elle peut également être réalisée de manière régulière (annuellement), devenant alors un outil de gestion (Figure 24). L'analyse a alors plusieurs finalités :

- elle permet d'alimenter un observatoire des consommations, en révisant chaque année les courbes, graphiques et cartes présentés dans la section précédente (section 1) (1) ;
- elle permet par ailleurs de produire un fichier de données clientèle enrichies des résultats de l'analyse (notamment le type de consommateur, la tendance détectée avec le test de Mann Kendal) (2) Ce fichier peut ensuite être exploité par le service pour approfondir une question particulière relative à un type précis de consommateurs (3) ;
- enfin, les données relatives à la typologie peuvent être réintégrées dans le fichier clientèle, afin de permettre une vérification manuelle par les agents chargés de la relève par exemple (4), permettant ainsi de progressivement fiabiliser la classification et de réduire le nombre d'abonnés « non classés » par la procédure automatique.

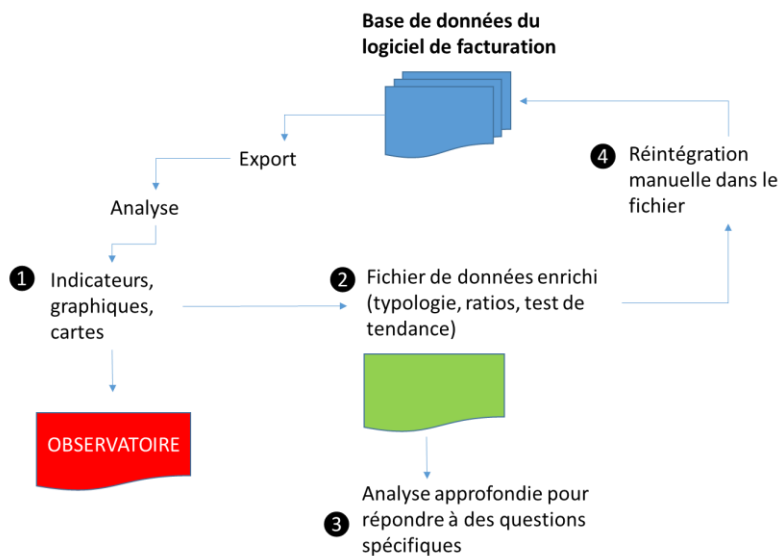


Figure 24 : L'analyse des fichiers de facturation intégrée dans une stratégie de gestion du fichier clientèle et facturation.

5. Références bibliographiques

Boisson, A., Drzewiecki, P., Bourgeois, C., Neverre, N., Rinaudo, J-D., Jehanno, F., (EDM56), Cauet, Y., et Rivallan, J., (SDAEP 22), Deconchy, A., et Ainaoui, M., (SMG Eau 35), Goarnisson, J-M. et Barraï, N. (CD 29)

Favre, M., Montginoul, M., 2018. Observatoire des consommations d'eau urbaine. Application à Nantes Métropole. Veolia et Irstea - UMR G-eau, Nantes, 60 p.

Favre, M., Montginoul, M., 2020. Comment observer de façon optimale la consommation en eau potable ? Le cas de Nantes Métropole. TSM (Techniques, Sciences et Méthodes) 1/2, 29-42. <https://astee-tsm.fr/numeros/tsm-1-2-2020/favre/> rinaudo prévision de la consommation en eau potable

Montginoul, M., 2013. La consommation d'eau en France : historique, tendances contemporaines, déterminants. Revue SET 10, 68-73.

Montginoul, M., Desprats, J.-F., Rinaudo, J.-D., 2013a. La consommation en eau potable. Le cas de la CAM (Communauté d'Agglomération de Montpellier) et /ou du Syndicat Garrigues Campagne. Irstea, Montpellier, 30 p.

Montginoul, M., Desprats, J.-F., Rinaudo, J.-D., 2013b. La consommation en eau potable. Le cas de la PMCA (Communauté d'Agglomération de Perpignan Méditerranée). Irstea, Montpellier, 27 p.

Montginoul, M., Even, L., Verdon, D., 2010. Comprendre l'évolution de la consommation d'eau potable dans l'agglomération nantaise. Cemagref et Nantes Métropole, 43 p.

Montginoul, M., Even, L., Verdon, D., 2013. Focus : Le cas de Nantes Métropole : un cas à part ? Revue SET 10.

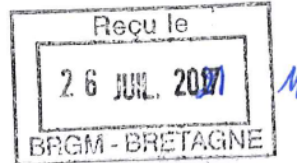
Montginoul, M., Rinaudo, J.-D., Desprats, J.-F., 2017. La consommation en eau urbaine dans les villes françaises : tendances actuelles sur deux métropoles méditerranéennes. TSM (Techniques, Sciences et Méthodes) 3, 12-24.

Rinaudo JD (2013) Prévoir la demande en eau potable : une comparaison des méthodes utilisées en France et en Californie. Revue Sciences Eau et Société, 10, pp. 78-86

Rivallan, J., Jehanno, F., Cauet, Y., Ainaoui, M., Barraï, N., Berrehouc, G., 2022. Connaissance de la consommation d'eau potable distribuée par les réseaux publics. De l'Eau Pour Demain

Annexe 1

Exemple de convention



CONVENTION

Echange et traitement de données personnelles sur les volumes d'eau potable consommés par les abonnés

Entre les soussignés :

Le BRGM, Bureau de Recherches Géologiques et Minières, établissement public à caractère industriel et commercial, dont le siège se trouve 3, avenue Claude-Guillemin, BP 36009, 45060 Orléans Cedex 02, immatriculé au Registre du Commerce et des Sociétés d'Orléans sous le numéro 582 056 149, représenté par Mélanie Bardeau, Directrice régionale de Bretagne, ayant tous pouvoirs à cet effet.

Ci-après désigné par le BRGM

Et

Le **Syndicat Mixte Eau du Morbihan**, représenté par son Président en exercice, dûment autorisé par délibération, demeurant au 27 rue de Luscanen à Vannes
ci-après désigné Eau du Morbihan

Le BRGM et Eau du Morbihan sont collectivement désignés par « les parties »

Visas :

Vu les statuts en vigueur de Eau du Morbihan ;

Vu le cahier des charges techniques du programme De l'Eau pour Demain ;

Vu la délibération de financement du projet De l'Eau pour Demain de la région Bretagne (Délibération n°20_0501-08 du la Commission Permanente du 30/11/20) ;

Vu la délibération de financement du projet Eau pour Demain de l'Agence de l'eau Loire Bretagne (Convention n°21C012 du 31/05/2021) ;

Vu le protocole de coopération relatif au programme De l'Eau pour Demain, signé le 15/04/2021 par les partenaires du projet De l'Eau pour Demain, incluant le BRGM et EDM,

ETANT PREALABLEMENT EXPOSE QUE :

- Dans le cadre du projet De l'Eau pour Demain, le BRGM et Eau du Morbihan ont convenu de réaliser une analyse des consommations d'eau potable, à partir des données individuelles issues du fichier de facturation du service. Cette analyse doit permettre d'évaluer des ratios de consommation par catégorie d'abonné et de caractériser les tendances d'évolution de ces ratios sur une période de 5 à 10 ans.
- La méthodologie proposée par le BRGM est issue de travaux antérieurs réalisés dans le cadre de projets de recherche scientifique (projet ANR Eau&3^E). Son application aux données de Eau du Morbihan permettra d'améliorer cette méthode et d'en valider la pertinence.
- Les résultats de cette analyse permettront à Eau du Morbihan d'améliorer sa gestion du service, en particulier en ce qui concerne la prévision de l'évolution future de la consommation, la mise en œuvre de mesures d'économie d'eau ou la tarification.

Il est convenu ce qui suit :

Article 1 : Objet de la convention

1.1. La présente convention a pour objet de définir les conditions dans lesquelles le BRGM s'engage à effectuer une analyse des données de consommation d'eau potable des abonnés de 2 secteurs du service de Distribution de Eau du Morbihan :

- Périmètre « b » Centre Morbihan Communauté
- Périmètre « d » Arc Sud Bretagne pour partie, Redon agglomération pour partie, Questembert communauté.

Article 2 : Données mises à disposition

2.1. Nature des données mises à disposition

Eau du Morbihan s'engage à mettre à disposition du BRGM la base de données servant à la facturation des consommations d'eau, qui permet de connaître le volume d'eau potable consommé par chaque abonné, pour chaque période de relève et sur plusieurs années ainsi que tous les renseignements liés à chaque client et chaque branchement. Ces données seront fournies sous forme nominative, les champs « noms » et « adresse » de la base étant nécessaires pour réaliser une classification automatique des abonnés par catégorie d'usagers. Eau du Morbihan se chargera de collecter les fichiers auprès de ses délégataires.

2.2. Respect de la réglementation

Les parties s'engagent à respecter la réglementation en vigueur applicable au traitement de données à caractère personnel et, en particulier le règlement (UE) 2016/679 du 27 avril 2016. Les mesures prises pour respecter la réglementation sont détaillées dans l'article 4.

Article 3 : Objectif de l'analyse des données

3.1. Le BRGM réalisera un traitement des données décrites à l'article 2 dans le but :

- De classer chaque abonné dans une catégories d'utilisateurs (domestiques, commerçants-artisans, agriculteurs, industriels, collectivités ...), en utilisation la typologie développée par la FNCCR ;
- De calculer, pour chaque abonné, une consommation annuelle et de caractériser la tendance d'évolution sur plusieurs années (test non paramétrique) ;
- De calculer des ratios de consommation unitaire (en volume par an) pour chaque catégorie d'abonné ;
- D'analyser les tendances évolutions de ces ratios de consommation par catégorie d'utilisateur et valoriser les résultats dans le cadre de ses missions (assistance technique auprès des collectivités de production, élaboration du schéma départemental d'alimentation en eau potable, ...) ;
- D'identifier, uniquement pour les catégories d'utilisateurs professionnels, les plus gros consommateurs et/ou ceux dont les consommations ont fortement augmenté afin de leur proposer des actions d'économie d'eau adaptées.

Article 3 : Coût

3.1. Les échanges d'information prévus dans la présente convention sont réalisés à titre gratuit. Dans le cadre du projet Eau pour Demain, le BRGM prend en charge les coûts liés à son intervention définie dans la présente convention.

Article 4 : Obligations du BRGM relatives aux données

4.1. Utilisation des données

En mettant à disposition les données décrites à l'article 2, Eau du Morbihan autorise le BRGM à :

- Les stocker dans sa base de données interne sécurisée, à accès restreint aux personnels en charge du traitement des données
- S'en servir, dans le cadre du projet Eau pour Demain, pour accomplir les missions citées dans l'article 3 de la convention. Sauf accord avec Eau du Morbihan, le BRGM s'interdit de traiter les données à d'autres finalités.

4.2. Confidentialité

Le BRGM s'engage à garantir la confidentialité des données à caractère personnel traitées dans le cadre du présent contrat et veille à ce que toutes les personnes autorisées à traiter les données à caractère personnel :

- Soient informées du caractère confidentiel des données ;
 - Aient reçu une formation appropriée concernant la législation applicable en matière de protection des données à caractère personnel ;
- Soient soumises à des procédures d'authentification d'utilisateur et de connexion pour accéder aux données.

4.3. Accès aux données par des tiers

Le BRGM pourra mobiliser un ou plusieurs experts scientifiques pour réaliser les traitements prévus à l'article 3. Le BRGM précisera à Eau du Morbihan l'identité de ces experts, la nature technique de leur intervention ainsi que la période pendant laquelle ils auront accès aux données. Ces experts seront tenus de respecter les obligations de la présente convention. Il appartient au BRGM de s'assurer que les experts présentent les garanties suffisantes quant à la mise en place de mesures techniques et organisationnelles répondant au règlement sur la protection des données personnelles.

4.4. Droit d'information des personnes concernées

Si nécessaire, il appartient à Eau du Morbihan, responsable de traitement, de fournir l'information aux personnes concernées par les opérations de traitement au moment de la collecte des données.

4.5. Sécurité

Compte tenu de la nature du traitement et des informations à sa disposition, le BRGM s'engage à prendre les mesures techniques et organisationnelles appropriées afin de garantir un niveau de sécurité adapté au risque et afin d'empêcher toute violation de la sécurité entraînant de manière accidentelle ou illicite, la destruction, la perte, l'altération, la divulgation non autorisée de données transmises, conservées ou traitées, ou l'accès non autorisé à de telles données.

4.6. Violation

En cas de violation des données à caractère personnel, le BRGM s'engage à signaler la violation à l'autorité de contrôle et/ou aux personnes concernées.

Il informe Eau du Morbihan dans les meilleurs délais après en avoir pris connaissance, en précisant au moins les informations suivantes : la nature de la violation, le contenu et la quantité des données concernées, les catégories et le nombre de personnes concernées, les conséquences probables, les mesures prises ou proposées afin de remédier à la violation des données y compris, le cas échéant, les mesures pour en atténuer les éventuelles conséquences négatives, la date et l'heure de l'incident et de la détection de l'incident ainsi que le nom et les coordonnées du délégué à la protection des données ou d'un autre point de contact auprès duquel des informations supplémentaires peuvent être obtenues.

Dans la mesure où il n'est pas possible de fournir ces informations simultanément, elles pourront être communiquées de manière échelonnée.

4.7. Devenir des données

Au terme de la présente convention, le BRGM s'engage à détruire toutes les données à caractère personnel. Une fois détruites, le BRGM doit justifier par écrit de leur destruction.

4.8. Délégué à la protection des données

Le BRGM s'engage à communiquer à Eau du Morbihan le nom et les coordonnées de son délégué à la protection des données.

4.9. Registre des catégories d'activités de traitement

Le BRGM déclare tenir par écrit un registre de toutes les catégories d'activités de traitement effectuées pour le compte de Eau du Morbihan comprenant :

- Le nom et les coordonnées du responsable de traitement pour le compte duquel il agit, des éventuels sous-traitants et, le cas échéant, du délégué à la protection des données ;
- Les catégories de traitements effectués pour le compte du responsable du traitement ;
- Le cas échéant, les transferts de données à caractère personnel vers un pays tiers ou à une organisation internationale, y compris l'identification de ce pays tiers ou de cette organisation internationale et, dans le cas des transferts visés à l'article 49, paragraphe 1, deuxième alinéa du règlement européen sur la protection des données, les documents attestant de l'existence de garanties appropriées ;

Article 5 : Propriété intellectuelle

5.1. Eau du Morbihan reste l'unique propriétaire des données mises à disposition du BRGM dans le cadre de la présente convention.

5.2. Le BRGM détient les droits de propriété intellectuelle associés aux développements méthodologiques qui seront réalisés dans le cadre du projet De l'Eau pour Demain, y compris ceux basés sur l'étude de cas Eau du Morbihan.

Article 6 : Durée de la présente convention

La présente convention est signée pour une durée de deux ans.

Convention établie en 2 exemplaires*,

A Rennes, le 09/07/2021

La directrice du BRGM Bretagne

A Vannes, le 15/07/2021

Le Président de Eau du
Morbihan



**RÉPUBLIQUE
FRANÇAISE**

*Liberté
Égalité
Fraternité*

Centre scientifique et technique

3, avenue Claude-Guillemain
BP 36009
45060 – Orléans Cedex 2 – France
Tél. : 02 38 64 34 34

Direction régionale Bretagne

Rennes Atalante Beaulieu
2 rue Jouanet – 35700 Rennes
Tél : 02 99 84 26 70

www.brgm.fr



Géosciences pour une Terre durable

brgm