

INRAE : Institut National de Recherche pour l'Agriculture, l'Alimentation  
et l'Environnement

# Traitement et analyse de données de suivi de la Prime Air Bois

STAGE M1 MIASHS SSD

Gabriel Macé  
27/06/2024

## Remerciements

Je tiens tout d'abord à remercier toutes les personnes qui ont permis la réalisation de ce stage. J'adresse mes remerciements à mes encadrants de stage au LESSEM, Madame Sandrine ALLAIN et monsieur Frédéric BRAY, qui me suivent tout au long de mon stage, en me guidant et me soutenant pour mener à bien ce projet. Je tiens aussi à remercier les membres du jury et l'attention qu'ils porteront sur la lecture de mon rapport.

## Résumé global

Dans le cadre du Master 1 Statistiques et Sciences des données à l'Université Grenoble Alpes (UGA), je réalise un stage au sein de l'Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE), et plus précisément dans le Laboratoire EcoSystèmes et Sociétés En Montagne (LESSEM), sous la direction de Frédéric Bray et Sandrine Allain.

Le bois est une énergie renouvelable mais aussi la principale source d'émissions de particules fines. Une politique mise en place pour remédier à ce problème est la prime Air-Bois, permettant une aide aux particuliers changeant leurs modes de chauffage au bois par un matériel plus performant et moins polluant. Dans le Pays Voironnais, le Grésivaudan, et Grenoble Alpes métropole, cette aide est mise en place depuis 2015 et a engendré une collecte de données liées aux dossiers de suivis des demandeurs. Un travail exploratoire mené en amont de mon stage a permis d'émettre l'hypothèse qu'il existe deux modes d'usages du chauffage bois-bûche : l'archétype « Ressource de proximité », et l'archétype « Marchandise prête à brûler », et que la prime air-bois favorise ce deuxième type. L'objectif de mon stage est de répondre à ces hypothèses en analysant les données issues des dossiers de suivis de la Prime Air-Bois, tout en procurant différents indicateurs aux acteurs du territoire.

La moitié du stage n'étant pas encore réalisée, la réponse à ces hypothèses n'est pour l'instant pas aboutie, mais des premiers indicateurs ont pu être produits.

## Abstract

As part of the Master 1 in Statistics and Data Science at the University of Grenoble Alpes (UGA), I'm doing an internship at the « Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE) », and more specifically in the « Laboratoire EcoSystèmes et Sociétés En Montagne (LESSEM) », under the supervision of Frédéric Bray and Sandrine Allain.

Wood log is a renewable energy source, but it is also the main source of fine particle emissions. One policy that has been put in place to remedy this problem is the « Prime Air-Bois », which provides financial assistance to individuals who change their wood-burning methods for more efficient, less polluting equipment. In Pays Voironnais, Grésivaudan and Grenoble Alpes Métropole, this aid has been in place since 2015, and has led to the collection of data linked to applicants' monitoring files. Exploratory work carried out before my internship led to the hypothesis that there are two types of use for wood-log heating: the "local resource" archetype, and the "ready-to-burn merchandise" archetype, and that the « Prime Air-Bois » favours the second type. The aim of my placement is to respond to these hypotheses by analysing the data from the « Prime Air-Bois » allowance monitoring files, while at the same time providing various indicators for local players.

As half of the internship has not yet been completed, the answer to these hypotheses is not yet complete, but the first indicators have been identified.

## Table des matières

Introduction.....	3
I – Présentation du stage.....	4
A – Laboratoire EcoSystèmes et Sociétés En Montagne (LESSEM) .....	4
B – Contexte du stage.....	4
C – État de l’art .....	4
D – Objectifs du stage.....	5
E – Présentation des données .....	6
1) Source de données .....	6
2) Données supplémentaires mobilisées .....	6
II – Matériel et méthodes.....	7
A - Les principales étapes du projet .....	7
B – Outils de travail .....	8
III – Premiers Résultats.....	9
A – Analyse descriptive.....	9
1) La prime et ses bénéficiaires .....	9
2) Mode d’usage .....	14
B – Analyse inférentielle.....	16
1 – À l’échelle des individus.....	16
2 – À l’échelle communale .....	18
IV – Les prochaines étapes .....	21
V – Impact sociétal et environnemental de mon stage .....	21
Conclusion .....	22
Bibliographie.....	23
Annexes .....	23

## Introduction

En Auvergne Rhône Alpes, sur la période 2016-2018, 4 300 décès sont attribuables, chaque année, à une exposition aux particules fines (PM<sub>2,5</sub>), sur cette même période, 67% des émissions de particules fines (PM 2.5) sont dû au chauffage individuel au bois. Dans ce contexte, réduire les émissions du au chauffage individuel au bois devient un enjeu de santé publique, cependant le bois est une énergie renouvelable relativement bon marché pour les consommateurs. Face à cette dualité, l'ALEC et l'AGEDEN ont mis en place une prime aidant les particuliers se chauffant au bois à changer leur appareil, pour un appareil labélisé « Flamme Verte », qui sera plus efficace, et émettra moins de particules fines dans l'air, et ceci sur trois territoires depuis 2015 : Le Pays Voironnais, le Grésivaudan, et Grenoble Alpes métropole. Ces organismes ont récolté des données liées aux suivis des dossiers de cette prime. En amont de mon stage, un travail exploratoire mené en partenariat entre le LESSEM et le PNR de Chartreuse, a mis en avant, à travers d'entretiens qualitatifs auprès d'acteurs, producteurs, et usagers, les hypothèses qu'il existe deux modes d'usages du chauffage bois-bûche : l'archétype « Ressource de proximité », et l'archétype « Marchandise prête à bruler », et que les politiques publiques telles que la prime air-bois favorise ce deuxième type. Un travail complémentaire répondant à ces hypothèses à travers l'analyse des données récoltées par l'AGEDEN et l'ALEC est donc tout à fait pertinent, c'est l'objet de mon stage.

La première partie de ce rapport introduit le contexte, l'état de l'art, les objectifs du stage et les données disponibles. La deuxième partie, décrit les méthodes, les outils, et les techniques mises en place pour répondre aux objectifs. Dans la troisième partie, nous trouverons les premiers résultats obtenus. Ensuite, la quatrième partie comporte les prochaines étapes. Enfin, dans la dernière partie nous trouverons l'impact sociétal et environnemental provisoire de mon stage.

## I – Présentation du stage

### A – Laboratoire EcoSystèmes et Sociétés En Montagne (LESSEM)

INRAE est un Etablissement Public Scientifique et Technologique (EPST), sous la double tutelle des ministères chargés de la recherche et de l'agriculture, dont les missions sont « de réaliser, d'organiser et de coordonner, sur son initiative ou à la demande de l'État, tous travaux de recherche scientifique et technologique dans les domaines de l'agriculture, de l'alimentation, de la forêt, de l'environnement, de l'eau, de la biodiversité, de la bioéconomie, de l'économie circulaire, de la gestion durable des territoires et des risques dans les champs de compétence précités [1] ».

Qui a pour ambition de relever les défis de l'agriculture, de l'alimentation et de l'environnement de demain. Né le 1<sup>er</sup> janvier 2020 de la fusion entre l'Inra et Irstea, INRAE est le 1<sup>er</sup> organisme de recherche finalisée au monde sur le continuum agriculture-alimentation-environnement.

Au classement en termes de publications scientifiques, l'institut se classe 3<sup>ème</sup> dans les sciences agricoles, 4<sup>ème</sup> pour les sciences de l'animal et du végétal ainsi que pour les sciences de l'aliment, et 10<sup>ème</sup> en Écologie – Environnement.

Parmi ses centres, on compte le centre Lyon-Grenoble dont fait partie l'unité de recherche LESSEM, qui réunit des chercheurs en écologie et en sciences humaines et sociales. Le LESSEM développe des recherches sur les dynamiques des socio-écosystèmes en montagne en visant l'équilibre entre approfondissement disciplinaire et développement de recherches interdisciplinaires, entre apports thématiques et perspectives méthodologiques.

Mon stage est co-encadré par Sandrine Allain, chargée de recherche en économie écologique (équipe ASTRRE), et Frédéric Bray, ingénieur en systèmes d'informations territoriaux (équipe COMPET).

### B – Contexte du stage

Le bois est une énergie renouvelable très présente sur notre territoire d'étude (Le Pays Voironnais, le Grésivaudan, et Grenoble Alpes métropole), cependant, mal utilisé, le bois bûche peut impacter la qualité de l'air (principale source de pollution aux particules fines). Des politiques de transition énergétique ont donc été mises en place pour reconquérir la qualité de l'air. La Prime Air Bois en fait partie : depuis 2015, cette prime aide les particuliers à financer le remplacement d'anciens appareils de chauffage au bois par des appareils labélisés « Flamme Verte », qui seront plus performants et qui émettront moins de microparticules dans l'air pour un usage similaire. Le suivi des dossiers et l'octroi des primes est assuré par des organismes conventionnés, notamment l'ALEC et l'AGEDEN.

Ces organismes ont pu collecter des données liées à ces dossiers, et ceci sur trois territoires et depuis 2015 : Le Pays Voironnais, le Grésivaudan, et Grenoble Alpes métropole.

Ce sont ces données que je suis amené à analyser.

### C – État de l'art

Atmo Auvergne-Rhône-Alpes a mené une « évaluation quantitative d'impact sur la santé (EQIS) de la pollution de l'air ambiant en région Auvergne-Rhône-Alpes » [2], sur les années 2016 à 2018. Ce travail permet de mettre en évidence l'impact des polluants atmosphériques sur la santé humaine. Un des résultats est que sur la période 2016-2018, 4 300 décès sont attribuables, chaque année, à une exposition aux particules fines (PM<sub>2,5</sub>) en Auvergne Rhône Alpes. Aussi, sur cette

même période, 67% des émissions de particules fines (PM 2.5) sont dues au chauffage individuel au bois (Fig. I.C.1).

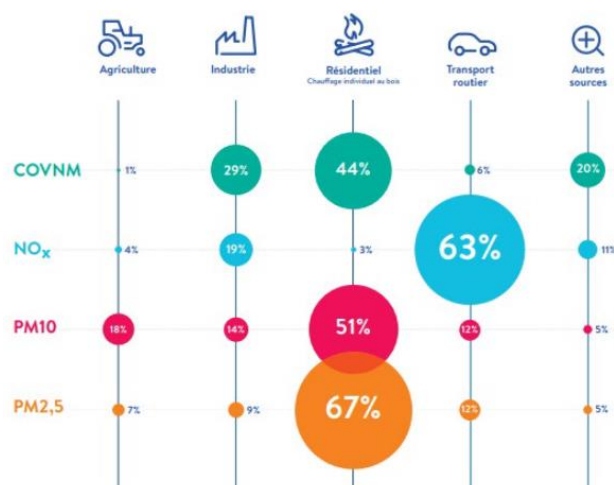


Figure I.C.1 : Contribution des secteurs d'activités humaines aux émissions de polluants de l'air ambiant en Auvergne-Rhône-Alpes, 2019, source : Atmo Auvergne-Rhône-Alpes

En août 2018, l'ADEME publie une enquête sur les pratiques d'utilisation des équipements domestiques de chauffage au bois [3]. L'objectif fut de « comprendre comment sont utilisés les appareils domestiques de chauffage au bois, particulièrement afin d'appréhender l'impact des pratiques sur la performance des équipements et sur la pollution de l'air ». Cette enquête a permis de dresser de nombreux indicateurs quant aux pratiques liées aux chauffages au bois, ainsi qu'une typologie spécifique pour « développer une campagne efficace d'information/sensibilisation sur la réduction de la pollution de l'air engendrée par le chauffage au bois domestique », voici les grandes catégories de ménages qu'il faudrait viser : les ménages vivant en milieu rural, et/ou se situant dans la tranche d'âge 35 à 49 ans, et/ou disposant de revenus inférieurs à 1 500 €/mois, et/ou les ménages se chauffant à l'électricité.

En 2023, un projet [4] de recherche exploratoire sur le bois-bûche a eu lieu en amont de mon stage, en partenariat entre le LESSEM et le PNR de Chartreuse dont le but était, à partir du cas d'étude de la Chartreuse, de comprendre comment et pourquoi cette filière évolue conjointement mais distinctement d'autres formes de bois énergie. Ce projet a mis en avant que d'une part le statut d'énergie renouvelable, bon marché, mais aussi de source polluante du bois-bûche le positionne aux prises de mutations politiques et économiques. Aussi, les ajustements économiques classiques sont contraints par le fait que les producteurs de bûches travaillent par habitude et proximité. Enfin, seule une partie des usagers porte des exigences nouvelles quant à la qualité, la quantité et la praticité de leur bois. Encouragés par des politiques publiques, ces nouveaux comportements constituent un moteur de changement de la filière bûche.

## D – Objectifs du stage

Le travail exploratoire mené en amont de mon stage comprenait principalement des entretiens qualitatifs auprès d'acteurs, producteurs, et usagers. L'hypothèse principale qui en est sortie étant qu'il existe deux modes d'usages du chauffage bois-bûche : l'archétype « Ressource de proximité », et l'archétype « Marchandise prête à brûler », et que les politiques publiques telles que la prime air-bois favorise ce deuxième type. Mon stage est donc complémentaire à ce travail,

l'objectif est de répondre à ces hypothèses en analysant les données issues des dossiers de suivis de la Prime Air-Bois, tout procurant différents indicateurs aux acteurs du territoire.

## E – Présentation des données

### 1) Source de données

Les données sont fournies par l'ALEC et l'AGEDEN, et sont de plusieurs origines : Les bases de données d'instruction présentant les données de gestion de 8542 dossiers de demande de la Prime Air Bois, qu'on peut relier, grâce aux numéros de dossier, au questionnaire en amont du projet, présentant 7447 individus et 40 variables décrivant les bénéficiaires, leurs matériels, et leurs comportements quant au chauffage au bois. Puis, environ 550 bénéficiaires ont rempli un questionnaire facultatif en aval du projet, où les questions étaient plutôt portées sur leur satisfaction et leur nouveau mode d'utilisation de l'appareil. Toutes ces données ont été collectées de novembre 2015 à mai 2024, et ceci sur trois territoires : Le Pays Voironnais, le Grésivaudan, et Grenoble Alpes métropole.

La base de données principale à analyser est le questionnaire en amont du projet, à laquelle nous avons ajouté plusieurs variables, telles que les différentes dates de demande et de versement de la prime, le coût total TTC des travaux, le montant d'aide demandé, le statu majoré ou non de la prime, le coût de l'appareil, ou encore la commune du demandeur.

### 2) Données supplémentaires mobilisées

Afin de compléter ces données, j'ai pu en récupérer d'autres, notamment au niveau communal. Grâce à l'outil SIDDT de INRAE, j'ai pu récupérer des variables socio-démographiques, telles que le nombre de maisons, le nombre de ménage par catégorie socioprofessionnelle, la densité de population au km<sup>2</sup>, ou encore la part de forêt dans la commune. Grâce à l'INSEE, j'ai pu récupérer la médiane du niveau de vie par commune Insee [5], et le type de chaque commune [6]. Aussi, monsieur Hervé Chanut, de l'observatoire de la qualité de l'air « Atmo Auvergne-Rhône-Alpes », m'a fourni des données indiquant par commune le nombre d'appareils anciens au 31 décembre 2022, donc ceux à changer, ainsi que le nombre de ces appareils changés en 2023.

## II – Matériel et méthodes

### A - Les principales étapes du projet

La première étape du projet fut la familiarisation avec le sujet et les enjeux relatifs au chauffage au bois. Pour cela, mes tuteurs m'ont fourni un certain nombre de documents (rapports, transcriptions de réunion), j'ai pu assister à des réunions, et interagir avec des acteurs du territoire qui, eux aussi, m'ont fourni de la documentation. Ces différentes interactions m'ont permis d'identifier en partie les besoins, et ce que l'analyse des données pouvait apporter, mais aussi l'origine des données et leurs limites.

La seconde étape fut la **mise en place de l'analyse**. J'ai d'abord découvert les données non mises à jour, et modifier et compléter les scripts R fournis par mes tuteurs et permettant de rendre les données exploitables, les compléter, et créer des premiers graphiques univariés ainsi qu'une cartographie du nombre de demandes par commune. Dans les données mises à jour, j'ai pu constater des doublons dans les numéros de dossiers, mais après consultation avec les fournisseurs des bases, il s'agit d'erreurs de frappe lors de l'entrée des numéros de dossiers, et donc il s'agit bien d'individus différents. Nous pouvons donc les exploiter tant que nous n'avons pas besoin de faire de lien vers une autre base via le numéro de dossier. Une des difficultés fut de quantifier la quantité de bois consommé par les demandeurs, puisqu'ils indiquaient des unités de mesure différentes, certaines n'étant pas exploitables, et qu'il y a deux types de combustibles : bois-bûche et granulés. Pour certaines variables les individus ont pu cocher plusieurs réponses qui étaient toutes renseignées de manière condensée dans la base de données. Afin d'exploiter ces différentes variables, les solutions furent de convertir les consommations en kwh puis en équivalent stère. Cependant, ces variables contiennent beaucoup de données non renseignées. Aussi, j'ai créé une variable binaire pour chaque réponse possible. De plus, certaines variables ne sont pas pertinentes à analyser, soit parce qu'elles ne présentent pas assez de réponses (par exemple la classe énergétique du logement, ou bien toutes les questions où il est demandé de préciser les réponses « autres »), soit parce que les réponses ne sont pas robustes (par exemple, pour l'origine du bois 750 individus ont indiqué que leur bois été labélisé « Rhône-Alpes-Bois-Bûche », or il n'existe que deux producteurs ayant ce label sur le territoire, ce qui montre que beaucoup de personnes ont mal compris cette modalité). Je ne les ai donc pas prises en compte lors de l'analyse.

Le début de l'analyse a consisté en une **analyse descriptive** des données, afin de comprendre les tendances et indicateurs, et caractériser l'évolution des comportements : création de graphique représentant la répartition uni et bi variées des individus, analyse temporelle de l'évolution de la demande de prime, et analyse spatiale au niveau communal. Puis j'ai produit des graphiques plus spécifiques en fonctions des besoins des acteurs, notamment des graphiques temporels et des cartographies. La suite de cette étape, fut un temps de travail avec les acteurs du territoire, dans lequel je leur ai présenté mes résultats, nous avons pu en discuter, m'ont donné leurs interprétations, des pistes d'améliorations et de nouvelles pistes d'exploration et d'analyse.

L'étape suivante fut une **analyse inférentielle**, l'analyse des liens statistiques entre les différents **individus** et les différentes variables, afin de trouver des groupes d'individus. Pour cela, j'ai d'abord étudié les liens inter-variables, à l'échelle de l'individu, à l'aide de tests, tels que les tests de régression linéaire, de corrélation de Spearman, de Chi2 d'indépendance, et des tests d'égalité de moyennes et de variances. Pour tous ces tests, j'ai choisi 5% comme seuil de significativité, et étant donné le nombre de tests, j'ai utilisé la correction de Benjamini-Hochberg et dans certains cas celle de Bonferroni pour corriger les p-valeurs. Puis, étant donné le nombre de variables (*et d'individus*), j'ai utilisé une analyse factorielle de données mixtes (AFDM) en sélectionnant préalablement les



variables pertinentes, afin de concentrer l'information portée par les données dans un nombre moindre de dimensions.

Étant donné les résultats de la réduction de dimension au niveau de l'individu, et étant donné que l'une des hypothèses est que le comportement des bénéficiaires change en fonction de la proximité à la ressource et des caractéristiques socio-économique du lieu de vie, j'ai continué l'analyse mais au **niveau communal** cette fois-ci. Pour cela, j'ai transformé les variables individuelles pertinentes, en les regroupant par communes : en utilisant la moyenne communale pour les variables qualitatives, et en créant une variable représentant le taux par communes de chaque modalité des variables quantitatives (par exemple, le taux de primes majorées par commune). Pour vérifier cette hypothèse, j'ai tout d'abord résumé l'information portée par ces variables, en leur appliquant une analyse en composante principale (ACP). Ensuite, j'ai appliqué des méthodes de clustering sur les données transformées par ACP : un clustering ascendant hiérarchique, permettant aussi de trouver le nombre optimal de cluster, et une méthode k-means, afin de trouver les éventuels groupes de communes. Puis, pour comprendre sur quelles variables se différencient les clusters de communes, j'ai réalisé des tests non paramétriques d'égalité de distributions de Wilcoxon des deux clusters et analysé les différences de moyennes inter-clusters dans les variables issues des dossiers, avant transformation par ACP. Afin de rendre compte de la proximité à la ressource et des caractéristiques socio-économique des communes, j'ai ajouté des variables communales, telles que le taux de forêt par commune, la médiane de niveau de vie, la densité de population, ou encore l'évolution du nombre de logements entre 2015 et 2020, mais aussi le nombre de dossiers, le nombre d'appareils « anciens » a changé fin 2022, et le nombre d'appareils changés en 2023, et ceci aussi ramené au nombre d'habitants. Afin d'essayer de comprendre l'origine des différences entre les deux groupes de communes, j'ai là aussi utilisé le test de Wilcoxon et calculé les différences de moyennes entre les deux groupes pour chaque variable « ajoutée », que l'on peut aussi nommer « explicative ».

## B – Outils de travail

Au cours de mon stage, j'utilise plusieurs outils afin de faciliter mon travail. Premièrement, afin de partager nos codes, et d'assurer la reproductibilité de mon projet, mes tuteurs et moi-même utilisons Gitlab, et en particulier celui de INRAE. J'ai aussi accès à un outil appelé « SIDDT », permettant notamment d'extraire des données par commune, en choisissant des zones EPCI, par exemple nos trois territoires. Pour mon analyse, j'utilise principalement trois outils, qui sont R, Rstudio, et Rmarkdown. Plusieurs librairie R me permettent d'améliorer l'efficacité de mon analyse, j'utilise plus particulièrement « readxl » et « writexl » afin d'exploiter les bases de données, « ggplot2 » et « shiny » afin de faciliter la lecture des représentations graphiques, « FactoMineR » et « factoextra » pour les réductions de dimensions, et « tidyr » et « dplyr » pour faciliter l'exploitation des données.

### III – Premiers Résultats

#### A – Analyse descriptive

##### 1) La prime et ses bénéficiaires

L'analyse des variables socio-professionnelles montre qu'un bénéficiaire moyen, a plus de 35 ans, aide non majorée, a un revenu annuel entre 30 000 et 50 000 €, et est retraité, ou cadre, voir employé.

Les motivations principales de changement d'appareil sont d'économiser de l'énergie, gagner en confort, améliorer la qualité de l'air, et souvent ces trois réponses ensemble (Fig.III.A.1).

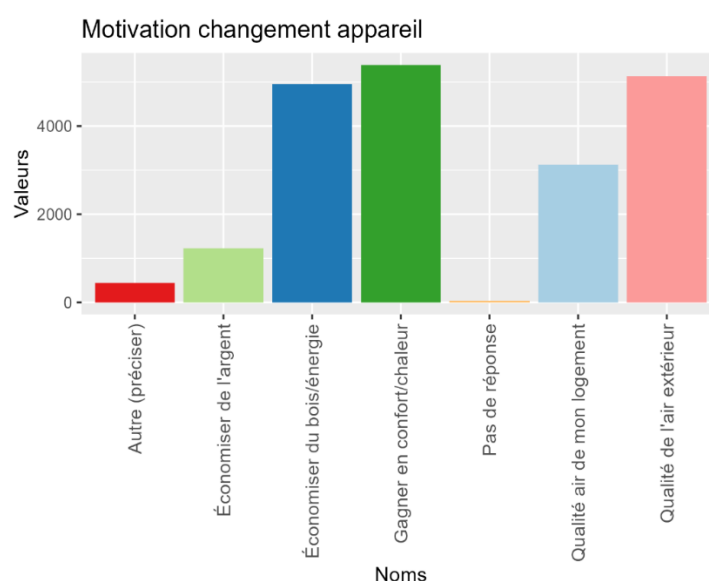


Figure III.A.1.1 : Nombre totale d'itérations de chaque motivation de changement d'appareil

J'ai pu comprendre qu'il était important pour les élus de savoir à quelles catégories de ménages était versées les primes. L'analyse de l'évolution de la part de prime majorée par territoire et par an montre que le Pays Voironnais accorde proportionnellement plus de rimes majorées, et que ce taux atteint un creux en 2019, avant de recroître. Pour les deux autres territoires, il semble plutôt stable (Fig.III.A.1.2). D'après les acteurs du territoire, cela est dû au fait qu'il y a plus de propriétaires modestes de maisons que dans les autres territoires.

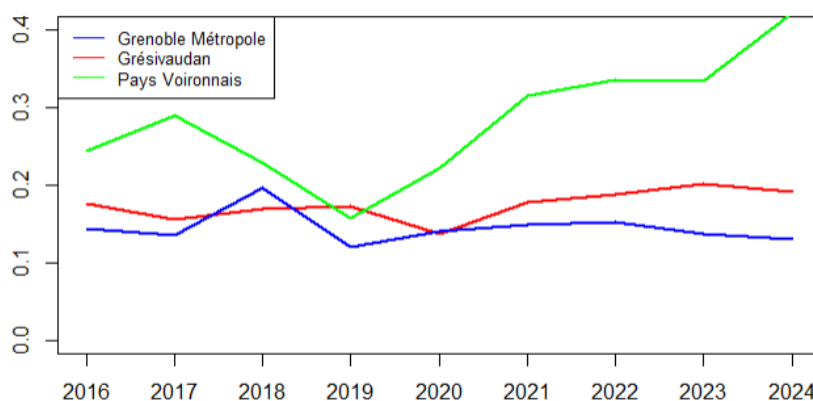


Figure III.A.1.2 : Évolution de la part de prime majorées par an et par territoire

Globalement, les parts de dossier par tranche de revenus sont stables, les plus importantes étant celles entre 20000 et 60000 euros par an. Mais ces dernières années, le taux de bénéficiaires ayant un revenu supérieur à 100 000 euros par an semble augmenter (Fig.III.A.1.3).

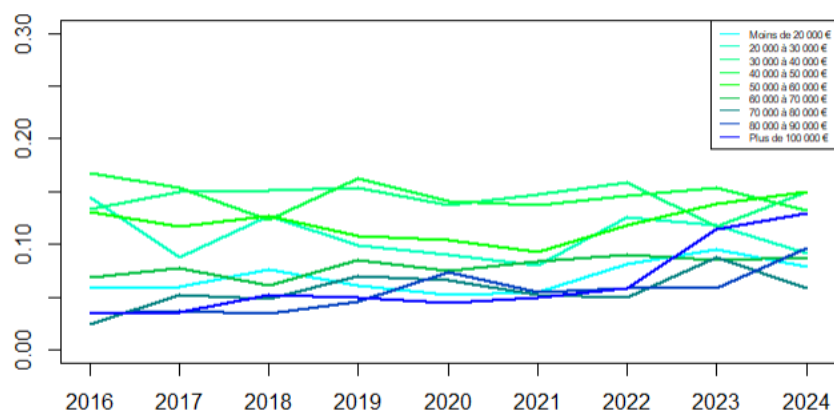


Figure III.A.1.3 : Évolution de la part dossier par tranche de revenus par an sur tout le territoire

La répartition des aides montre que chaque territoire a ses modalités de primes, et qu'elles évoluent dans le temps. Pour les territoires du Grésivaudan et Grenoble Alpes Métropole, les primes initiales étaient de 1200 € pour les majorées, et de 800 € pour les autres, en 2019 elles ont toutes deux augmentées de 800 €. Pour le Pays Voironnais, les primes étaient de 400 et 800 €, avant d'augmenter en 2021 de 400 €. Les bénéficiaires du voironnais reçoivent donc des aides plus faibles, et les bénéficiaires du Grésivaudan reçoivent proportionnellement plus de primes minorées que ceux de Grenoble Alpes Métropole (Fig.III.A.1.4).

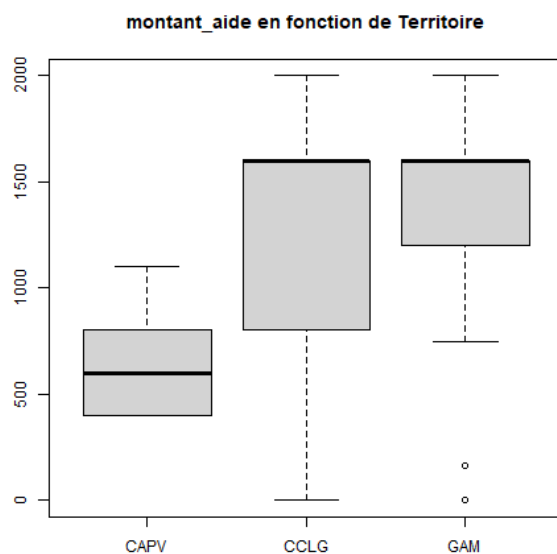


Figure III.A.1.4 : Répartition des montants des aides en fonction du territoire

On peut aussi remarquer une augmentation du coût moyen TTC des travaux (Fig III.A.1.5).

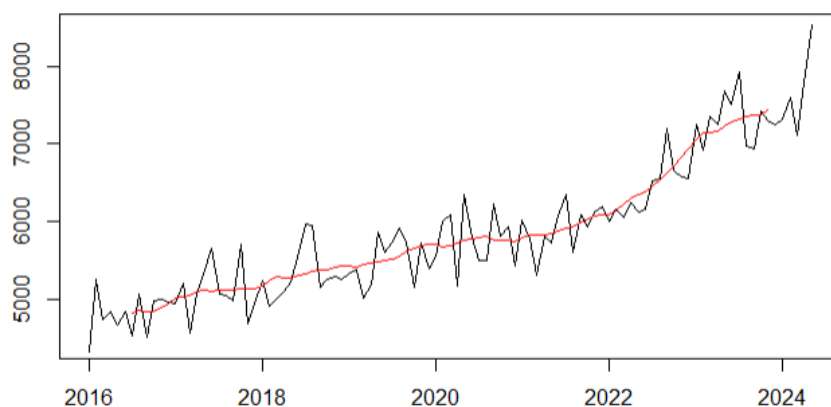


Figure III.A.1.5 : Évolution du coût moyen TTC des travaux

L'analyse temporelle de l'évolution du nombre de demandes de primes pour tous les territoires et par mois, montre que tout d'abord, la série ne semble pas stationnaire (Fig.III.A.1.5). Il semble y avoir deux pics : en 2019 et en 2022, suivies de deux chutes les années suivantes (Fig.III.A.1.6). Le premier pic est dû à l'augmentation des primes au Grésivaudan et à Grenoble Alpes Métropole, c'est pour cela que l'on ne l'observe pas au Pays Voironnais (Fig.III.A.1.7), puis la crise du covid fait chuter les demandes en 2020. Ensuite, les demandes reprennent fortement, notamment grâce à l'augmentation des primes dans le Pays Voironnais, mais la crise énergétique de 2022 est responsable de la deuxième chute. Il semble y avoir un effet saisonnier assez fort (Fig.III.A.1.6), il y a plus de demande durant les mois d'automne, avant la saison de chauffe.

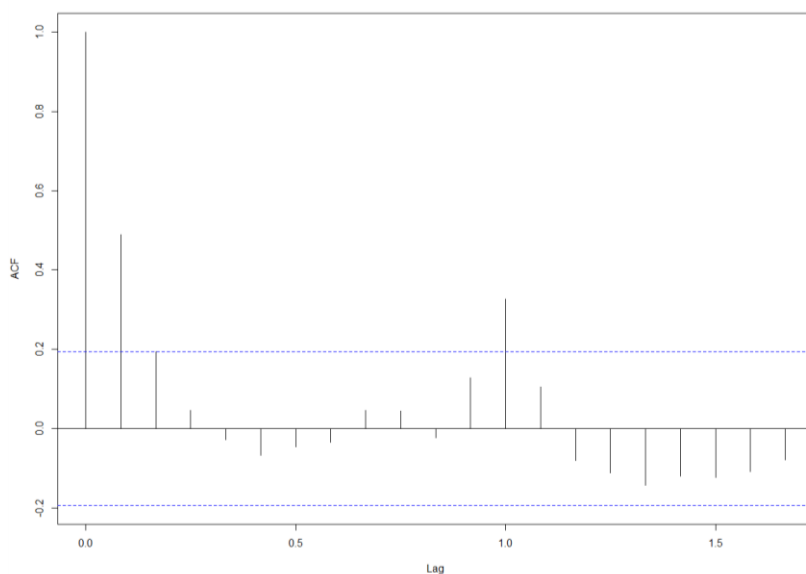


Figure III.A.1.5 : Autocorrélation du nombre total de demande de primes par mois

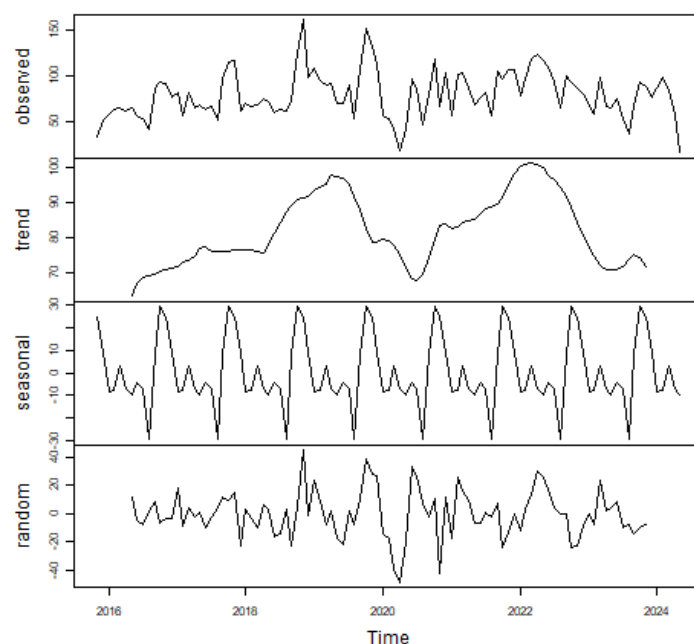


Figure III.A.1.6 : Décomposition de la série représentant le nombre total de demande de primes par mois

La demande est la plus forte pour le territoire Grenoble Alpes Métropole, et la moins forte au Pays Voironnais, ce qui est corrélé avec le nombre d'habitants. Aussi, les demandes au Grésivaudan et à Grenoble Alpes Métropole semblent évoluer ensemble (Fig.III.A.1.7).

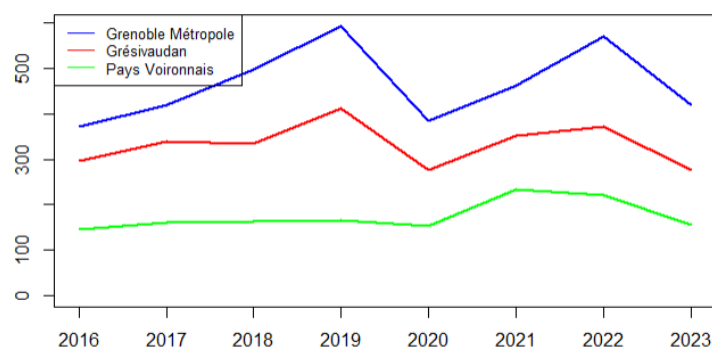


Figure III.A.1.7 : Évolution du nombre de demande par mois par territoire

La cartographie montre que la répartition varie de 0 à plus de 300 dossiers par commune, et que les plus importantes sont l'Ouest du Grésivaudan et Grenoble (Fig.III.A.1.8). J'ai aussi pu proposer cette carte sous forme interactive grâce à « Plotly » (1). Le Plateau des Petites Roches a un nombre très fort de demande de prime, malgré une population pas très élevée (moins de 1000 ménages fiscaux). On remarque aussi des bénéficiaires hors zones d'intérêt, il s'agit de propriétaires ayant un logement sur un des 3 territoires mais ne vivant pas dans ces territoires. Le nombre de dossiers par habitant met en avant l'importance de la demande de prime au Plateau des Petites Roches, mais remet en perspective le nombre de demandes à Grenoble, qui est très faible comparé à la taille de la population (Fig.III.A.1.9).

(1) Disponible qu'en local pour l'instant

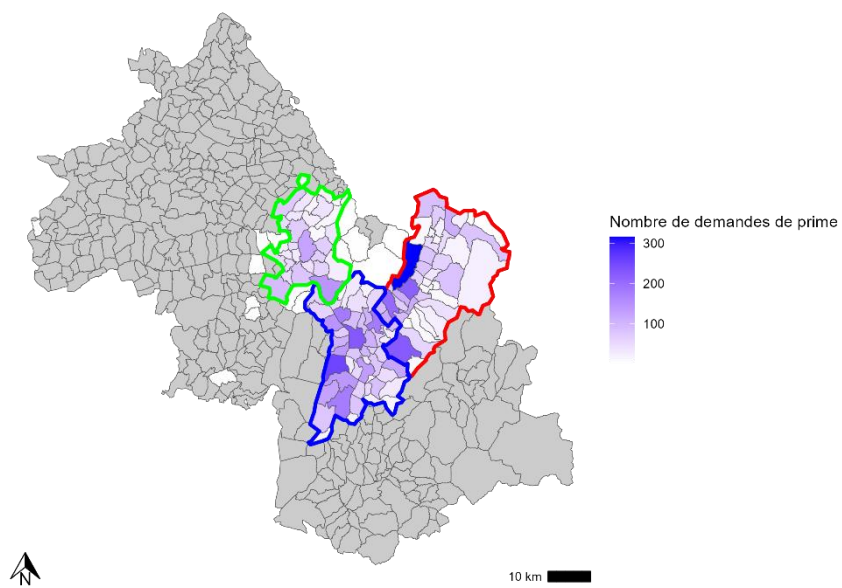


Figure III.A.1.8 : Nombre de dossiers par commune

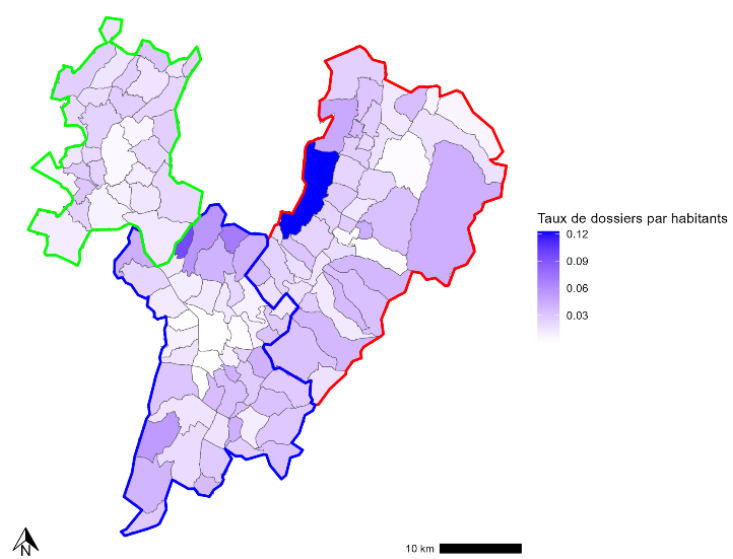


Figure III.B.1.9 : Nombre de dossiers par habitant et par commune

## 2) Mode d'usage

L'analyse des changements d'usage montre que les demandeurs de prime remplacent leur appareil par un poêle, alors que la moitié d'entre eux possèdent un insert ou un foyer fermé, et plus d'un quart un foyer ouvert (Fig.III.A.2.1). Les bénéficiaires se tournent vers un appareil plus performant, ce qui est le but de la prime Air-Bois.

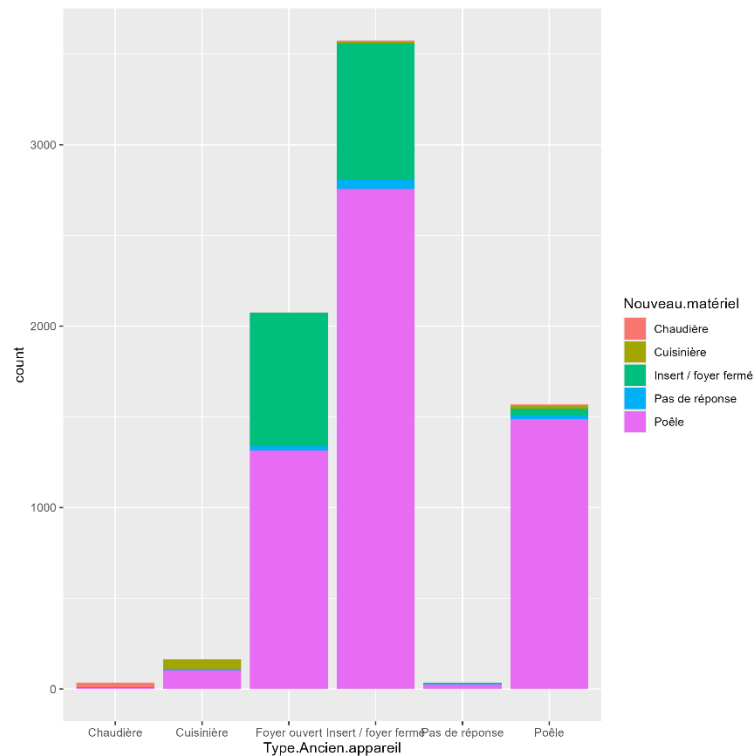


Figure III.A.2.1 : Changement de type d'appareil

Il existe bien des changements de comportements suite aux changements d'appareils. Les utilisateurs se servant de leur appareil comme chauffage principal n'ont principalement pas changé leur usage, mais un tiers des utilisateurs se servant de leur appareil comme chauffage d'appoint l'utilisent maintenant comme chauffage principal, et les utilisateurs s'en servant comme plaisir ou agrément ayant pratiquement disparu (Fig.III.A.2.2). La plupart des bénéficiaires répondant au questionnaire facultatif (76.5%) ayant déclaraient avoir une utilisation de leur nouvel appareil conforme à ce qui était prévu. Malgré ce plus fort usage de leur appareil, la consommation de bois en équivalent stère est globalement plus basse après changements (Fig.III.A.2.3), ce qui est dû à la meilleure performance des nouveaux appareils.

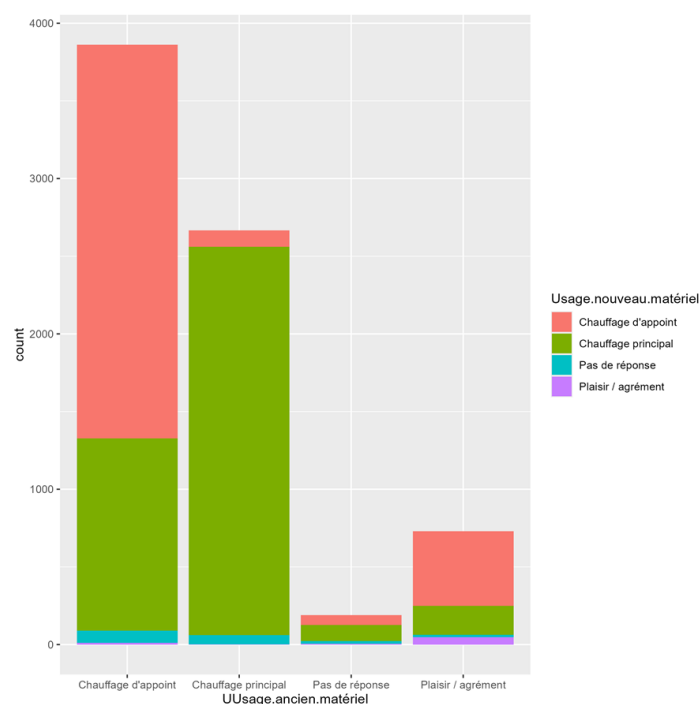


Figure III.A.2.2: Changement des usages du matériel

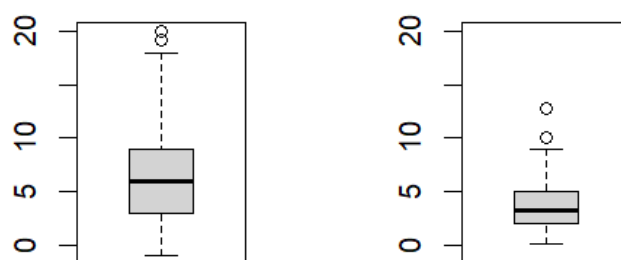


Figure III.A.2.3 : Évolution de la consommation en équivalent stères entre les dossiers d'instructions, donc avant changement (à droite) et les questionnaires facultatifs, donc après changement (à gauche)

L'analyse de l'évolution des taux de demandes de chaque type de combustible montre un léger rapprochement entre le taux de nouveau matériel fonctionnant à la bûche, et ceux fonctionnant au granulé de 2015 à 2021, puis une chute de ce deuxième en 2022 (Fig.III.A.2.2), dû à la crise énergétique augmentant drastiquement le coût du granulé. D'après les acteurs du territoire, il y a un effet crise de l'énergie, et que cela concorde avec les constats du côté des installateurs.

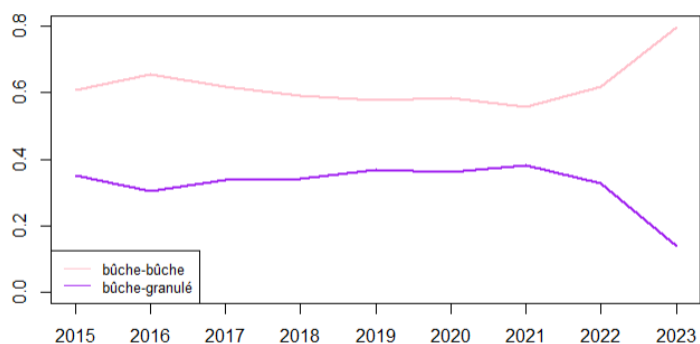


Figure III.B.1.6 : Évolution des taux de changements de type de combustible



## B – Analyse inférentielle

### 1 – À l'échelle des individus

L'analyse des liens statistiques entre les différentes variables au niveau de l'individu fût peu concluante. En effet, les conclusions des tests furent qu'il existe des liens entre les variables quantitatives, mais très peu de liens entre les variables qualitatives, et que les modalités des variables qualitatives forment des groupes d'individus assez distincts. Pour les variables quantitatives, seules deux variables semblent corrélées (Fig.III.B.1.1) : celle représentant le coût total TTC des travaux, et le coût de l'appareil (0.667), ce sont aussi les deux seules variables où l'on obtient un  $R^2$  supérieur à 0.2 : 0.445, c'est-à-dire que 44.5 % de la variance d'une variable peut être expliquée par celle de l'autre. Cependant, on rejette les deux hypothèses nulles : « X n'a pas d'effet sur Y » et « X et Y sont indépendantes », au seuil 5%, pour 8 des 15 couples de variables possibles (Fig.III.B.1.2), et les 5 variables sont présentes dans ces couples, elles semblent donc dépendantes les unes des autres.

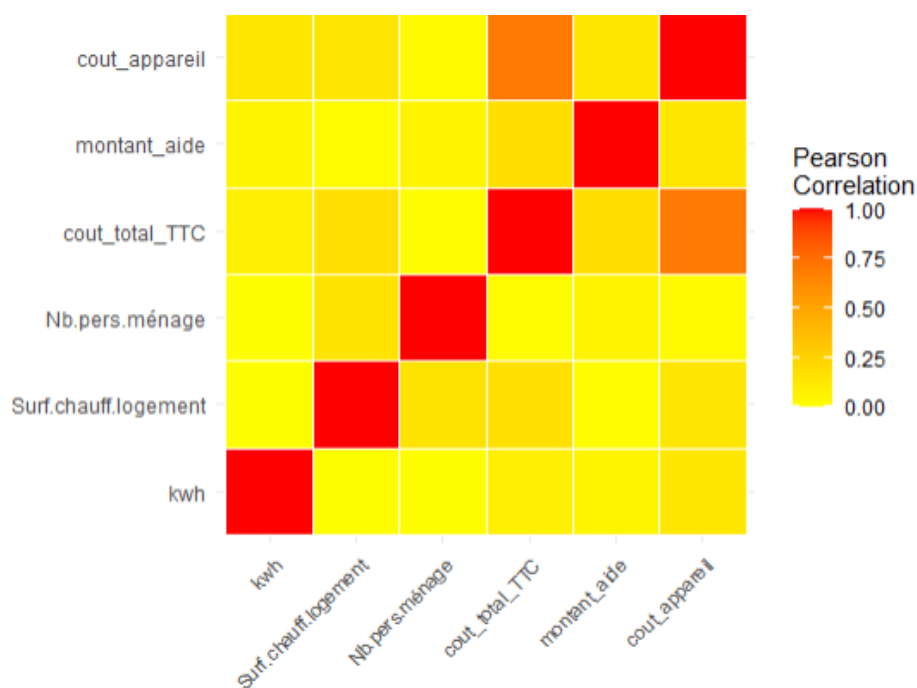


Figure III.B.1.1 : Corrélations entre les variables quantitatives

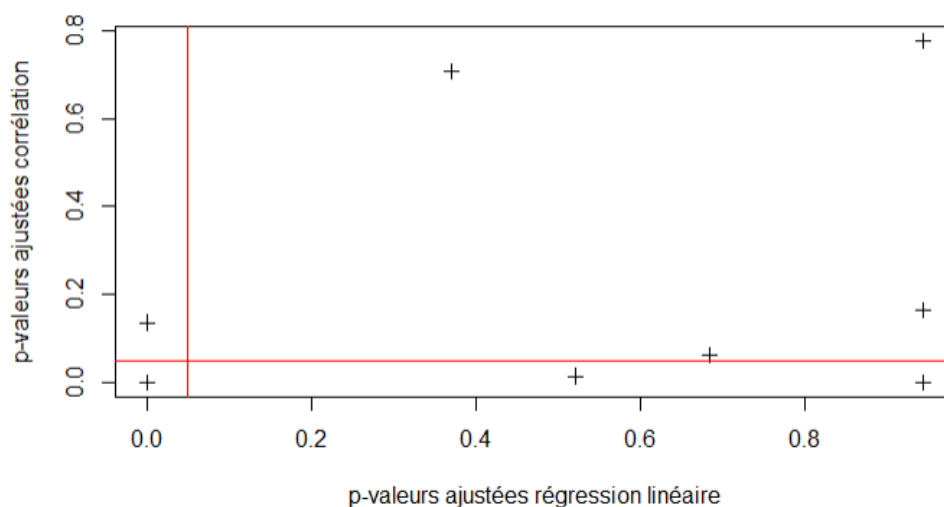


Figure III.B.1.2 : Liens inter variables quantitatives

Pour les variables qualitatives, il y a seulement moins d'un huitième (31 / 253) des tests de  $\chi^2$ , où l'on ne rejette pas, au seuil 5%, l'indépendance des variables (Fig.III.B.1.3), malgré la correction de Bonferroni qui est très forte. Il y a donc peu de dépendances entre ces variables.

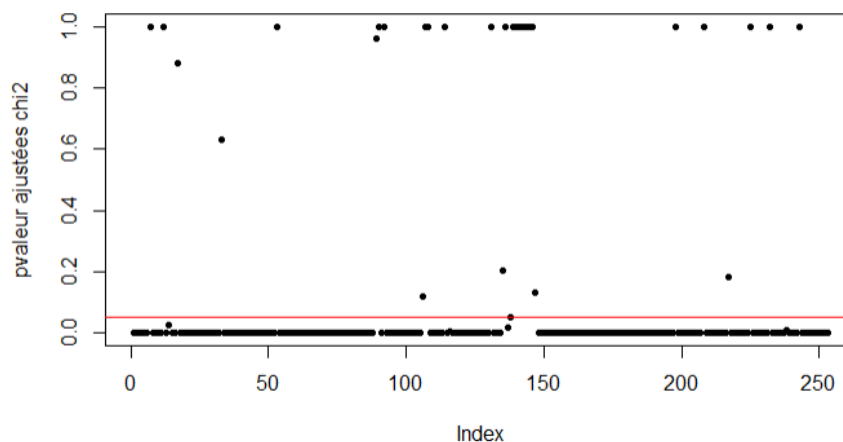


Figure III.B.1.3 : p-valeur des tests de  $\chi^2$  d'indépendance

Les résultats des tests d'égalité de moyennes et des tests d'égalité de variances, réalisés pour chaque variable quantitative (sauf la variable représentant la consommation de bois en équivalent stère car elle présente trop de valeurs manquantes pour être divisée en groupe) en fonction de chaque variable qualitative pertinente, furent qu'au seuil 5%, nous rejetons l'égalité des moyennes pour 80 des 115 tests, nous rejetons l'égalité des variances pour 103 tests, et nous rejetons aucune des deux pour 6 des 115 combinaisons de variables (Fig.III.B.1.4). Les modalités des variables qualitatives forment donc des groupes assez distincts lorsque nous les projetons sur les différentes variables quantitatives, ce qui nous encourage à résumer l'information par réduction de dimensions.

#### Tests d'égalité des variances et des moyennes

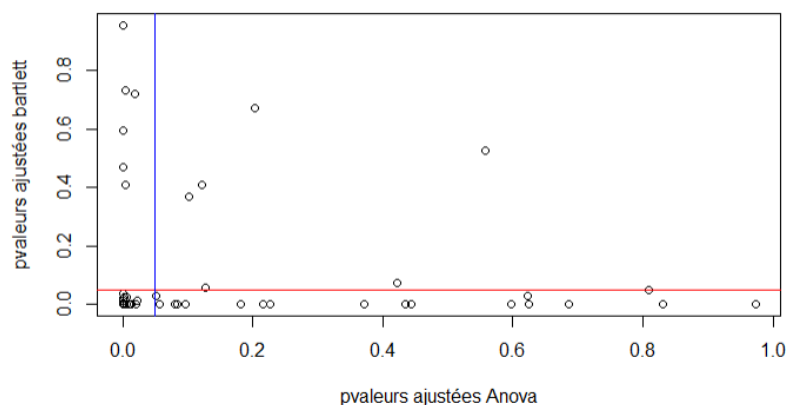


Figure III.B.1.4 : Titre

La réduction de dimensions fut peu analysable. En effet, même en sélectionnant strictement les variables à inclure, et en essayant de faire plusieurs AFDM en séparant les individus selon leur territoire, le statut de majoration de leur prime, ou encore le type de combustible de leur nouvel appareil, nous obtenons au maximum 15 % de variance expliquée par les deux premières dimensions, et cette valeur peine à augmenter en ajoutant des dimensions. J'ai donc choisi de ne pas analyser les résultats de cette réduction de dimensions.

## 2 – À l'échelle communale

À la suite de ce constat, nous avons choisi de se placer, non plus au niveau de l'individu, mais à celui de la commune. Les variables conservées sont le type de l'ancien et du nouvel appareil, leur usage (chauffage d'appoint, principal...), et leur type de combustible, ainsi que la motivation de changement d'appareil, le statut majoré ou non de la prime, et le coût total TTC des travaux. Les variables qualitatives sont transformées en plusieurs variables représentant le taux de dossier de chaque modalité, par commune. Avant d'effectuer une ACP sur ces données, et après en avoir discuté avec mes tuteurs, nous avons décidé de travailler uniquement sur les appareils de chauffage du logement, et pas sur les autres installations type cuisinière ou chaudière, et d'enlever les communes contenant moins de dix dossiers puisque leurs taux par modalités sont trop volatiles.

Cette fois, les premières dimensions de l'ACP expliquent 44% de la variance, et les dix premières : 80 % de la variance. J'ai donc tout d'abord appliqué un clustering ascendant hiérarchique à ces dix premières dimensions, dont le dendrogramme suggère plutôt de créer deux groupes (Fig.III.B.2.1).

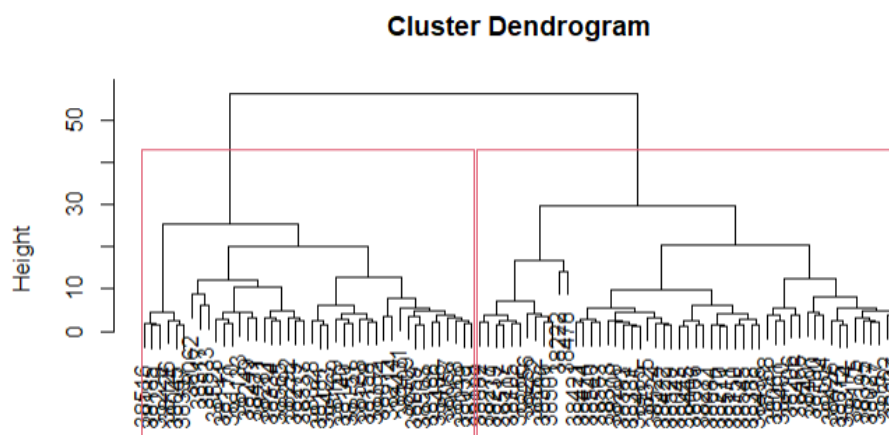


Figure III.B.2.1 : Dendrogramme issu du clustering ascendant hiérarchique

L'algorithme des k-means, semble bien séparer les deux groupes (respectivement 52 et 59 individus) selon les deux premières dimensions (Fig.III.B.2.2). La contribution aux deux premières dimensions étant attribuable à l'ensemble des variables, exceptées celles représentant les modalités peu représentées (« NA », « Autre », « Pas de réponses »), et le coût total TTC des travaux (Fig.III.B.2.3).

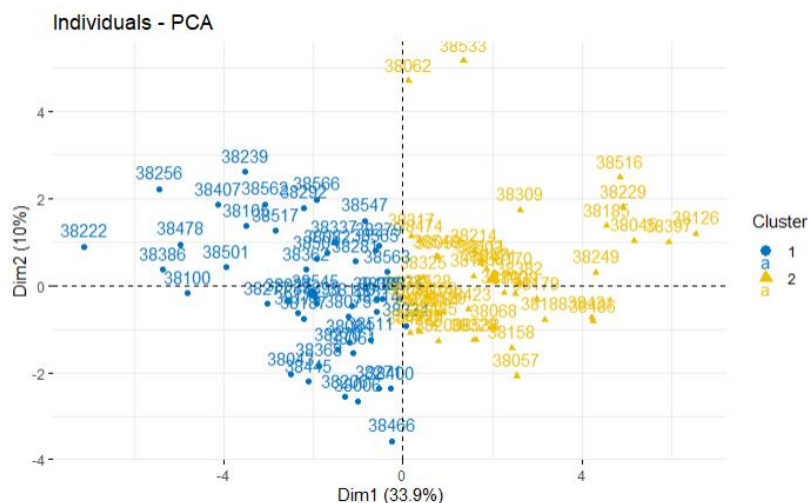


Figure III.B.2.2 : Les groupes issus des k-means, projetés sur les deux premières dimensions de l'ACP



Les variables supplémentaires qui semblent pouvoir expliquer la séparation des deux groupes sont surtout le nombre de dossiers par communes (13), et la part de maisons dans les résidences principales (11) (Fig.III.B.2.5 et annexe III.B.2.2). D'autres variables démographiques peuvent aussi avoir un rôle, notamment celles liées au nombre de logement et à la taille de la population. On observe aussi une différence significative dans la distribution des deux groupes des deux variables liées aux forêts (1 et 4), mais les différences de moyennes ne semblent pas être très fortes. Une des hypothèses étant que le comportement des bénéficiaires change en fonction de la proximité à la ressource, il se peut que les variables concernant les forêts soient incomplètes pour expliquer la proximité de la ressource, et qu'il faudrait que ce soit forêt de feuillus, ou bien producteur de bûche.

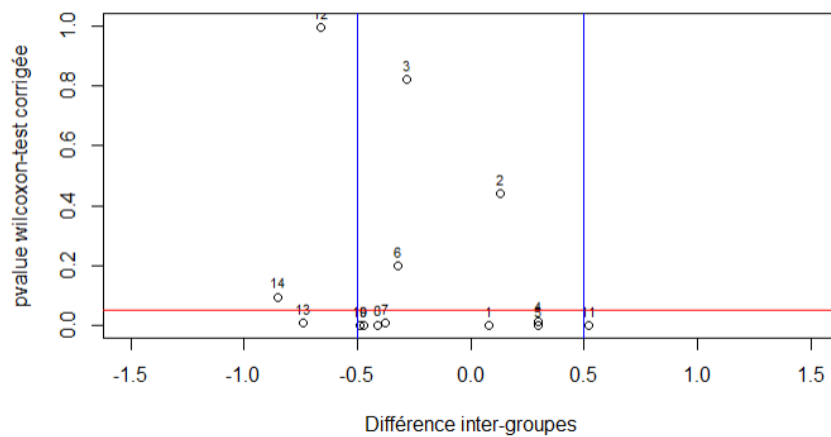


Figure III.B.2.5 : Différences inter-clusters, variables supplémentaires

## IV – Les prochaines étapes

Une des étapes qui va suivre est de revoir la typologie par individus et de l'approfondir.

Il faut aussi creuser les groupes communaux et leurs différences, notamment en ajoutant ou en changeant des variables explicatives, comme l'accès à la ressource, mais aussi en analysant plus précisément les résultats. Il faudrait aussi réessayer la réduction de dimensions « communale » sans les variables représentant les modalités de types « NA », « Autre » et « Pas de réponses ». On peut aussi réessayer l'analyse en prenant plus que deux clusters, afin d'être plus précis. Il faut aussi approfondir la détection d'outliers dans les communes, et la recherche de liens entre les variables supplémentaires communales, que j'ai commencées mais n'ai pas eu le temps d'approfondir.

Une étape importante est de produire des supports pour les acteurs du territoire et aussi d'assurer le relais des connaissances, notamment pour le projet de thèse. Notamment, il faut trouver comment rendre les cartes interactives facilement exploitables.

## V – Impact sociétal et environnemental de mon stage

Dans un contexte d'enjeux environnementaux et sociétaux, le bois-bûche présente une dualité entre énergie renouvelable et bon marché pour les consommateurs, mais aussi de polluants et principale source d'émissions de particules fines, en Auvergne Rhône Alpes du moins. Mon stage s'inscrit donc dans un projet global d'amélioration des matériels de chauffage au bois, permettant aux utilisateurs de continuer d'avoir accès à cette ressource renouvelable et relativement peu chère, tout en consommant moins de combustibles et en émettant moins de polluants.

J'ai aussi cherché à connaître mon impact environnemental personnel au cours de ce stage en mesurant mon empreinte carbone. Au niveau des déplacements, j'ai surtout utilisé mon vélo personnel (environ 280 km), mais quelques fois aussi le bus et la voiture (environ 30 km chacun). D'après le calculateur de l'ADEME [7], un trajet de 10 km à vélo n'émet aucun équivalent CO<sub>2</sub>, 30 km de bus (gnv) émettent 3.66 kg d'équivalents CO<sub>2</sub> (30 km \* 122 g /km), et 30 km de voiture émettent 6,54 kg d'équivalents CO<sub>2</sub> (218 g/km \* 30 km). J'ai aussi participé à des visio-conférences, environ 10 heures au total, une minute de visioconférence émettant environ 1g de CO<sub>2</sub> (Leboucq, 2020), j'ai donc émis 600g d'équivalents CO<sub>2</sub> par ce biais-là. J'ai aussi consommé une importante d'électricité à travers mon code informatique, que j'ai tenté d'estimer à travers le package python codecarbon. J'estime cette quantité à 1.8 kWh et donc à 180 g équivalents CO<sub>2</sub>.

Je suis aussi conscient que mon analyse n'est pas complète, je n'ai pas pris en compte mon temps passé à utiliser l'ordinateur sans utiliser R, notamment via des requêtes internet, et que cela a aussi une émission d'équivalents CO<sub>2</sub> non négligeable.

À la date du 27 juin, j'estime ma quantité d'équivalent CO<sub>2</sub> émise à :  $3.66 + 6.54 + 0.6 + 0.180 = 7.98$  équivalents CO<sub>2</sub>.

## Conclusion

Mon stage a pour objectif de répondre aux émises durant le travail exploratoire précédant mon stage, et de procurer différents indicateurs aux acteurs du territoire. J'ai utilisé 3 bases de données, dont deux étaient liées entre elles, et quelques données supplémentaires.

Après mettre familiarisé avec le contexte et les enjeux, et après avoir rendu les données exploitables, j'ai mis en place une analyse descriptive des données me permettant de mieux comprendre celle-ci, et de partager les tendances et indicateurs avec les acteurs du territoire.

Puis une première analyse inférentielle au niveau des individus, à travers les tests de régression linéaire, de corrélation de Spearman, de Chi2 d'indépendance, et des tests d'égalité de moyennes et de variances et à travers une analyse factorielle de données mixtes, a permis de montrer qu'il existait des liens entre les variables quantitatives, mais très peu de liens entre les variables qualitatives, et que les modalités des variables qualitatives forment des groupes d'individus assez distincts, et qu'il était compliqué de résumer l'information au niveau de l'individu.

Ensuite, une analyse inférentielle des variables au niveau communale, en utilisant des méthodes telles qu'une analyse en composante principale, un clustering ascendant hiérarchique, une méthodes k-means, ou encore des tests non paramétriques d'égalité de distributions de Wilcoxon ont permis de montrer que l'on pouvait créer deux groupes de communes, et qu'ils se différencient sur des variables liées à l'utilisation et au type du nouveau et de l'ancien matériel de chauffage, mais aussi au taux de majoration des primes, et que des variables notamment démographiques pourraient expliquer cela.

Sur le plan personnel, ce stage m'apporte de nouvelles compétences, notamment dans l'interaction avec les chercheurs et les acteurs du territoire, mais m'a aussi permis de développer mes compétences quant à l'exploitation de bases de données, et ma rigueur dans les calculs et les interprétations des résultats.

## Bibliographie

- [1] Décret n° 2019-1046 du 10 octobre 2019 relatif à l'organisation et au fonctionnement de l'Institut national de recherche pour l'agriculture, l'alimentation et l'environnement - Légifrance. (s. d.). <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000039207211>
- [2] Evaluation Quantitative d'Impact Sanitaire | Atmo Auvergne-Rhône-Alpes. (s. d.). <https://www.atmo-auvergnerhonealpes.fr/actualite/evaluation-quantitative-dimpact-sanitaire>
- [3] Connaissance des usages liés au chauffage domestique au bois en France. (2018). Dans Enquête Sur les Pratiques D'utilisation des Équipements Domestiques de Chauffage Au Bois. Consulté le 22 mai 2024, à l'adresse <https://bibliothèque.ademe.fr/ged/761/rapport-connaissances-usages-lies-chauffage-bois-domestique-en-france-2018.pdf>
- [4] Allain, S., & Lopez-Dovina, L. (2023). Mutations de la filière et de la gouvernance du bois-bûche : entre univers domestique, logique de marché et politiques de transition énergétique. Une étude de cas en Chartreuse.
- [5] Principaux indicateurs sur les revenus et la pauvreté aux niveaux national et local en 2021 – Principaux résultats sur les revenus et la pauvreté des ménages en 2021 | Insee. (s. d.). <https://www.insee.fr/fr/statistiques/7756729?sommaire=7756859>
- [6] Une nouvelle définition du rural pour mieux rendre compte des réalités des territoires et de leurs transformations – La France et ses territoires | Insee. (s. d.). <https://www.insee.fr/fr/statistiques/5039991?sommaire=5040030#onglet-1>
- [7] Calculer les émissions de carbone de vos trajets. (s. d.). Agir Pour la Transition Écologique / ADEME. <https://agirpourlatransition.ademe.fr/particuliers/bureau/calculer-emissions-carbone-trajets>

## Annexes

### III.B.2.1 : Tableau des différences inter-clusters, variables issues des dossiers

Variables	Différence de moyenne inter-groupes	pvalue wilcoxon-test	index
Type.combustible.nouveau.materiel_Bûche..Granulés	0	2,33E-10	1
Type.Ancien.appareil_Pas.de.réponse	-0,04	0,238186	2
Usage.nouveau.matériel_Pas.de.réponse	-0,05	0,046854	3
Type.combustible.nouveau.materiel_autre	-0,09	3,83E-08	4
Type.Ancien.appareil_Insert...foyer.fermé	0,11	4,36E-11	5
Usage.ancien.matériel_Pas.de.réponse	-0,19	2,23E-14	6
Type.combustible.nouveau.materiel_Pas.de.réponse	-0,21	0,146657	7
Type.combustible.nouveau.materiel_Bûche..autre	-0,27	1,69E-10	8
Nouveau.matériel_Pas.de.réponse	-0,29	6,43E-13	9
majoration_NA	0,37	0,006084	10
Économiser.de.l'argent	0,37	3,05E-13	11



Gagner.en.confort.chaleur	0,51	0,811539	12
Qualité.de.l.air.extérieur	0,58	2,93E-11	13
Qualité.air.de.mon.logement	0,66	0,025947	14
cout_total_TTC	0,7	0,190354	15
Usage.nouveau.matériel_Plaisir...agrément	-0,73	6,76E-12	16
Économiser.du.bois.énergie	0,79	0,010173	17
Type.Ancien.appareil_Poêle	1,02	2,57E-14	18
majoration_M	1,13	1,74E-14	19
Type.Ancien.appareil_Foyer.ouvert	-1,15	0,100141	20
Usage.ancien.matériel_Plaisir...agrément	-1,15	0,000181	21
majoration_NM	-1,15	1,58E-10	22
Usage.ancien.matériel_Chauffage.d.appoint	-1,24	1,58E-10	23
Type.combustible.nouveau.materiel_Bûche	-1,24	0,286272	24
Type.combustible.nouveau.materiel_Granulés	1,28	3,52E-05	25
Nouveau.matériel_Insert...foyer.fermé	-1,34	0,000486	26
Nouveau.matériel_Poêle	1,37	2,17E-05	27
Usage.ancien.matériel_Chauffage.principal	1,45	0,114571	28
Usage.nouveau.matériel_Chauffage.d.appoint	-1,45	0,001258	29
Usage.nouveau.matériel_Chauffage.principal	1,45	0,001355	30

### III.B.2.2 : Tableau des différences inter-clusters, variables supplémentaires

Variables	Différence de moyenne inter-groupes	pvalue wilcoxon-test	index
taux_forets	0,08	2,98E-05	1
taux_dossiers_habitant	0,13	0,439809	2
taux_changement : Nombre d'appareils changés en 2023 sur le nombre définit fin 2022 d'appareils à changer	-0,28	0,819396	3
superficie_foret	0,3	0,015258	4
objectifs_2022_par_maison : Nombre définit fin 2022 d'appareils à changer diviser par le nombre de maisons	0,3	0,000392	5
Part_résidences_principales_loc : Part de résidences principales en location	-0,32	0,198679	6
nb_ménages_fiscaux	-0,38	0,006884	7
nb_personnes_ménages_fiscaux	-0,41	0,00218	8
évolution_nb_logements : Différence entre le nombre de logements en 2015 et en 2020	-0,47	0,000664	9
Densité_pop	-0,49	0,000664	10
part_maisons (dans les résidences principales)	0,52	2,98E-05	11
objectifs_2022 : Nombre définit fin 2022 d'appareils à changer	-0,66	0,992927	12
n : nombre de dossiers par commune	-0,74	0,007759	13
mediane_niveau_vie	-0,85	0,095605	14