

# Rapport de stage

-

## Calcul en temps réel des émissions de polluants issues du secteur résidentiel

Thibaud GENEL

Du 2 avril au 30 septembre 2024 (6 mois)

Maître de stage : Damien BOUCHARD, Ingénieur émissions & énergie

Dans le cadre du Master 2 Mathématiques Appliquées, Statistique, parcours Data Science  
d'Aix-Marseille Université

# **Remerciements**

Merci à Damien Bouchard, maître de stage, pour l'opportunité, pour l'accueil, pour sa bienveillance, et pour son accompagnement au quotidien.

Merci à Grégory Maillard, tuteur du stage.

Merci à Damien Piga, Directeur Relations Extérieures et Innovation, qui fut un soutien de poids dans le projet de ce stage.

Merci à tous les employés d'AtmoSud pour leur bonne humeur et leur bienveillance.

# Table des matières

Introduction.....	5
Contexte .....	5
AtmoSud .....	6
Organisation du stage.....	7
Objectifs et Attentes.....	8
I - Exploration .....	9
1.1 - Contours du projet et Inventaire .....	9
1.2 - Relation entre température et consommation .....	11
1.3 - Recherche de données en temps réel .....	14
II – Test de calcul sur les métropoles.....	15
2.1 - Objectif .....	15
2.2 - Présentation générale de la méthode.....	16
2.3 - Extraction des données de consommation.....	16
2.4 – Ratio de consommation.....	18
2.5 – Correction selon la température .....	18
2.6 – Facteurs d’émission .....	23
2.7 – Emissions finales .....	24
2.8 - Validation des résultats.....	25
III - Estimation de la consommation électrique résidentielle à partir de la température (Machine Learning).....	26
3.1 – Motivation et objectif.....	26
3.2 – Pertinence d’un modèle sur les métropoles.....	26
3.3 - Données d’entraînement .....	28
3.3.1 Consommation .....	28
3.3.2 Température .....	29
3.3.3 Consommation et Température .....	29
3.4 - Modélisation .....	30
3.4.1 Régression polynomiale .....	30
3.4.2 D’autres modèles ?.....	31
3.4.3 Gestion des extrêmes (Edge-Case) .....	31
3.5 - Export du modèle pour utilisation .....	34
3.6 - Test du modèle.....	34
IV - Méthode de calcul des émissions en temps réel du secteur Résidentiel .....	36
4.1 - Objectif .....	36

4.2 - Problématique.....	36
4.3 - Présentation générale de la méthode.....	38
4.4 - Extraction des températures communales journalières.....	38
4.5 - Consommation électrique à partir de la température .....	40
4.6 - Décomposition de la consommation électrique .....	42
4.6.1 Détermination de BASE_CONSO .....	43
4.6.2 Correction de l'ECS .....	43
4.6.3 Identification du chauffage et de la climatisation .....	44
4.7 - Consommations par usage et énergie .....	45
4.7.1 Pour les usages fixes .....	45
4.7.2 Pour les usages thermosensibles .....	45
4.8 - Emissions de polluants par usage et par énergie.....	47
4.9 - Validation de la méthode.....	49
4.10 - Visualisation .....	51
4.11 – Limites et Perspectives.....	52
Conclusion .....	53
Références.....	54
Annexes .....	55

# Introduction

## Contexte

Chaque année en France, près de 40 000 décès prématurés sont dus à la pollution de l'air. De plus, toujours en France, le coût de la pollution atmosphérique est estimé à environ 100 milliards d'euros annuels. Il est ainsi nécessaire d'agir pour limiter au maximum l'exposition des populations et des écosystèmes à des concentrations trop importantes de polluants atmosphériques. Par analogie, le problème est sensiblement le même en ce qui concerne les gaz à effets de serre, qui eux n'ont pas nécessairement d'effets directs d'exposition sur la santé, mais coûtent et vont coûter encore plus cher aux sociétés humaines, car contribuant directement au changement climatique. Il faut agir pour limiter les émissions de gaz à effet de serre dans l'atmosphère.

Et pour être en capacité d'agir, il faut d'abord surveiller. C'est là qu'interviennent les Associations Agréées de Surveillance de la Qualité de l'Air (AASQA), ainsi que d'autres observatoires air-climat-énergie. Ces organismes ont la mission de rendre compte, au quotidien, de la qualité de l'air et de l'évolution des émissions de gaz à effet de serre auprès du public, des autorités, des collectivités territoriales, et de tous ceux en capacité de prendre des décisions pour réduire l'impact négatif de la pollution. Il s'agit en réalité d'un travail de fourmis : dans chaque région française, plusieurs équipes d'ingénieurs travaillent chaque jour pour recenser toutes les émissions de polluants et gaz (dans ce qu'on appelle l'inventaire des émissions) ou pour modéliser les concentrations de chaque polluant sur l'intégralité des territoires, entre autres choses.

Plus concrètement, l'idée va être de pouvoir informer les collectivités territoriales et les pouvoirs publics afin que les décideurs prennent des décisions éclairées. Par exemple, les modélisations réalisées sont capables de prédire un pic de pollution 2 jours avant son occurrence, de quoi prendre des mesures de restriction de circulation pour un préfet, qui réduiraient l'importance du pic, ou bien diffuser à un maximum de monde de limiter son activité physique durant cet épisode pour préserver ses poumons. L'idée va être d'identifier des pratiques excessivement émettrices de polluants ou de gaz à effets de serre, pour inciter à diminuer ces pratiques et à les faire évoluer, soit via de la communication, soit via la loi. Par exemple, on sait maintenant que le chauffage au bois représente une part essentielle des émissions de particules fines, il s'agit donc de communiquer là-dessus en vue d'une prise de conscience du public et des décideurs.

Mon stage s'effectue au sein du pôle Systèmes d'Information d'AtmoSud (présentée ci-après), dans une équipe chargée de réaliser les inventaires des émissions et des consommations, qui recensent les niveaux d'émission de chaque source (et secteur) pour chaque polluant en région Provence-Alpes-Côte d'Azur. Ces inventaires sont calculés au pas de temps annuel, avec un délai de 2 ans du fait des délais de disponibilité de la donnée source et des divers étapes de calcul et validation de la donnée. Il serait forcément intéressant et utile de les rendre plus proche de l'action en produisant des données et indicateurs "en temps réel". C'est l'objet de ce stage de 6 mois, en se concentrant sur le secteur résidentiel, ainsi que tertiaire car relativement similaire du point de vue des pratiques émettrices.

## AtmoSud

AtmoSud est l'observatoire de surveillance de la qualité de l'air de la région Provence-Alpes-Côte d'Azur, agréé par le ministère en charge de l'environnement. Il s'agit d'une Association Agréée de Surveillance de la Qualité de l'Air (AASQA).



### Historique d'AtmoSud :

1972	Création d'AIRFOBEP, association chargée de surveiller la qualité de l'air de l'Ouest des Bouches-du-Rhône
1982	Création d'AIRMARAIX, association chargée de surveiller la qualité de l'air de l'Est des Bouches-du-Rhône, du Var et du Vaucluse
1989	Création de Qualit'Air, association chargée de surveiller la qualité de l'air des Alpes-de-Haute-Provence, des Hautes-Alpes et des Alpes-Maritimes
1996	Loi sur l'Air et Utilisation Rationnelle de l'Energie (LAURE)
2006	Création d'Atmo PACA (fusion de Qualitair et AIRMARAIX), association chargée de surveiller la qualité de l'air des Alpes-de-Haute-Provence, des Hautes-Alpes, des Alpes-Maritimes, de l'Est des Bouches-du-Rhône, du Var et du Vaucluse
2012	Régionalisation : AIRFOBEP et Atmo PACA s'unissent pour former Air PACA, l'observatoire de la qualité de l'air de la région Provence-Alpes-Côte d'Azur
2018	Évolution du nom Air PACA en AtmoSud

Ses missions d'intérêt général participent à l'amélioration de la qualité de vie et du bien-être de chacun, ainsi qu'à la sauvegarde de la biodiversité. Sa structuration collégiale en fait un acteur indépendant qui œuvre au côté de tous les acteurs - publics, acteurs économiques, associatifs – pour favoriser leurs engagements dans la préservation de l'air et du climat.

AtmoSud c'est l'organisme régional de référence sur l'air et le climat :

- 50 ans d'expertise technique et scientifique
- 150 adhérents (principalement des collectivités territoriales) œuvrant pour la qualité de l'air et le climat
- 60 collaborateurs engagés pour la préservation de l'environnement

AtmoSud est indépendante grâce à une gouvernance partagée, proche des acteurs, investie dans la transversalité des connaissances, impliquée dans le travail en réseau, engagée dans sa mission d'intérêt général.

## Organisation du stage

Ce stage se déroule du 2 avril au 30 septembre 2024 dans le siège d'AtmoSud au 146 rue Paradis à Marseille.

**Horaires** : Le contrat indique 35 heures hebdomadaires à effectuer, soit 7h par jour en moyenne, avec une certaine flexibilité.

**Equipe** : Les collègues avec lesquels je peux être amené à travailler sont les suivants :

- Damien Bouchard, mon “tuteur”. Il est inventeur chez AtmoSud, c’est lui qui gère l’équipe “Emissions” (transversale dans plusieurs pôles au sein de la structure) incluant ceux qui recensent et calculent toutes les émissions et bilans énergétiques dans la région. Il me donne les directions à suivre dans mon travail, me fait des retours réguliers, et est disponible au quotidien si j’ai besoin de son avis. Il apporte à mon travail une vraie expertise dans la méthodologie de calcul comme dans toutes les notions à connaître sur la qualité de l’air et le climat ; il connaît très bien les données que je manipule et leur organisation.
- Damien Piga, directeur de l’Innovation. C’est lui qui donne les directions principales et les grands objectifs à cocher, notamment en début de stage, puisque c’est l’un des principaux commanditaires du projet. Il s’attache aussi à mener les discussions avec des organismes extérieurs, par exemple lorsque l’on a dû demander des données en “CloseData” à Enedis.
- Emeline Carre, développeuse et inventeur. C’est elle qui a réalisé les derniers inventaires annuels des émissions du secteur résidentiel. Il m’est nécessaire d’échanger avec elle pour bien comprendre sa méthodologie puisque ma mission est d’utiliser les données qu’elle a obtenu en vue de produire une donnée en temps réel. Elle est également de bon conseil pour des aspects techniques autour de l’automatisation (API, machines virtuelles etc..).
- Morgane Imbertesche, Julien Poullidor, Younes Abenna, Benjamin Rocher, Ludovic Lelandais, étant tous dans l’équipe Emissions et de bon conseil en cas de besoin.
- Mathieu Izard, ingénieur d’études, pour qui j’ai eu l’occasion d’écrire un sujet de « Bulle d’Air », une publication hebdomadaire résumant la situation en termes de qualité de l’air ainsi que l’avancée de projets en cours.

**Environnement technique** : Chez AtmoSud, les données internes sont notamment stockées en bases de données PostgreSQL sur des serveurs. Pour les manipuler, nous utilisons le logiciel de gestion de bases de données DBeaver, ainsi que le langage de programmation SQL adapté. En dehors de la manipulation spécifique de ces données, je dispose d’une liberté importante dans le choix des technologies à utiliser. En outre, seront également utilisés Python avec VSCode comme éditeur, Power BI pour la validation et la visualisation, MobaXTerm pour l’automatisation avec des machines virtuelles.

Pour ce qui est des outils moins « tech », la gestion de mon projet (Tâches, Problèmes, Questions, Notes) se fera sur Google Docs, la télécommunication sur Zoom et Teams.

## Objectifs et Attentes

En quelques mots, l'objectif du stage est de mettre au point une méthode pour estimer ou calculer en temps réel les consommations énergétiques et les émissions de polluants et gaz à effet de serre issues du secteur résidentiel, c'est-à-dire des habitations. Par temps réel, on entend un pas de temps journalier idéalement. L'échelle spatiale sera idéalement celle de la commune. Ce projet s'inscrit dans une démarche globale d'amélioration des inventaires et des indicateurs proposés par AtmoSud.

En effet, les inventaires des émissions, qui recensent toutes les émissions du territoire selon leurs sources de manière très précise, à l'échelle communale et au pas de temps annuel, sont utiles pour identifier de manière précise et fiable les émissions de polluants, moins pour permettre aux acteurs d'interagir avec la réalité, étant donné qu'ils sont diffusés avec un décalage minimal de 2 ans. Par exemple, à l'été 2024, les inventaristes finalisent la production des inventaires des émissions jusqu'en 2022, car les données utilisées sont très variées et mettent du temps à être collectées, vérifiées, remontées.

Ainsi, AtmoSud a lancé des projets dans plusieurs secteurs comme le maritime ou le routier pour calculer en temps réel, ou au moins en temps rapproché, l'inventaire des émissions. L'idée est de gagner sur l'aspect temporel quitte à perdre en fiabilité par rapport aux inventaires annuels - qui resteront en place -, afin de répondre aux problématiques des acteurs de la qualité de l'air, du climat et de l'énergie.

Le projet se déroulera en plusieurs phases :

- Conception d'une méthode de calcul
- Déploiement de la méthode de calcul dans les bases de données internes, automatisation
- Démonstration des résultats auprès des acteurs internes voire externes

Ainsi, tout l'intérêt pour moi est de mettre en œuvre un projet Data du début à la fin, de sa conception avec un papier et un stylo à sa défense devant des acteurs en s'appuyant sur des visualisations poussées. J'attends donc de ce stage qu'il me permette de saisir profondément tout ce qu'implique un projet de genre, de renforcer mes compétences en ingénierie des données (ou data engineering), en gestion de projet, en statistiques et machine learning, en communication et en visualisation des données.

C'est également un sujet à la fois intéressant, faisant appel à des données variées et concrètes, et valorisant où la finalité sert le bien commun.



# I - Exploration

## 1.1 - Contours du projet et Inventaire

Chaque émission du jour doit être caractérisée selon la commune où elle est émise, le polluant émis, l'usage qui en est la source (Chauffage, Cuisson, Climatisation etc..) et l'énergie qui en est la source (Electricité, Bois, Gaz etc..). En effet, AtmoSud calcule chaque année un inventaire des émissions qui recense sur tout le territoire les émissions de polluants et leurs sources. De façon légèrement simplifiée, l'inventaire présente des valeurs d'émission de nombreux polluants

- Annuelles
- A l'échelle communale
- Différenciées par secteur

Les principaux secteurs considérés dans l'inventaire sont l'Industrie, le Résidentiel, le Maritime, le Routier, le Tertiaire etc..

Secteur d'activité	Contenu
Transports routier	Ensemble des émissions liées au transport sur route, pour tout type de véhicules et de motorisation, y compris pour l'usage personnel
Transport maritime	ensemble des émissions des navires dans le périmètre des ports régionaux
Transport aérien	ensemble des émissions engendrées par les aéronefs dans le périmètre des aéroports régionaux. N'inclut pas les émissions liées au fonctionnement de l'aéroport.
Transport ferroviaire	Ensemble des émissions liées à la circulation des trains et des transports en commun sur rail. N'inclut pas les émissions liées aux bâtiments ou aux bureaux.
Transport fluvial	Ensemble des émissions engendrées par la navigation fluviale
Agriculture	Ensemble des émissions engendrées par les activités de culture, élevage, sylviculture et des engins/structures associées à ces activités
Résidentiel	Ensemble des émissions liées à l'habitat pour les particuliers. Inclut le bâtiments, les engins spéciaux mais pas le transport
Tertiaire	Ensemble des émissions liées aux activités de service. Inclut le bâtiments, les engins spéciaux mais pas le transport
Biogénique	Ensemble des émissions naturelles liées à la végétation, aux zones humides ou encore aux animaux sauvages. Toutes émissions pour lesquelles il n'y a, à priori, pas de responsabilité de l'activité humaine. N'inclut pas les feux de forêt qui font l'objet d'un secteur à part entière.
Industrie	Ensemble des émissions de l'industrie manufacturière et du secteur de la construction. N'inclut pas les activités de production d'énergie ou de traitement de déchets
Production d'énergie	Emissions liées à la production, à la transformation ou à la distribution d'énergie (raffinage, centrales thermiques, stations de compression, etc.) Inclut également la valorisation énergétique des unités d'incinérations.
Déchets	Ensemble des émissions liées au traitement des déchets, hors valorisation énergétique
Incendies de forêts	Ensemble des émissions dues aux incendies de massifs forestiers
Séquestration carbone	Bilan des flux d'absorption/émissions de carbone liés à la gestion et à l'utilisation des sols

- Différenciées par usage

Chaque émission doit être identifiée selon l'activité qui en est la source. Par exemple dans le secteur Résidentiel, les usages importants sont le Chauffage, l'Eau chaude, la Cuisson, la Climatisation, l'Eclairage, le Froid, le Lavage, et un usage Autre.

- Différenciées par énergie

On connaît également l'énergie utilisée lors de chaque activité émettrice.

Catégorie d'énergie	Explication du contenu
Gaz Naturel	Non renouvelable, n'inclut donc pas le biogaz.
Produits pétroliers	Hydrocarbures : gazole, essence, fioul, kerosène, etc.

Electricité	Electricité consommée par l'utilisateur final.
Chaleur et froid issus de réseau	Chaleur/froid produit dans les réseaux de chaleur et consommés par l'utilisateur final
Bois-énergie (EnR)	Utilisation de bois pour la combustion
Autres énergies renouvelables (EnR)	Biocarburants, fraction organiques des ordures ménagères, etc.
Autres non renouvelables	Gaz résiduaire de process, plastiques, pneumatiques, déchets non organiques, etc.

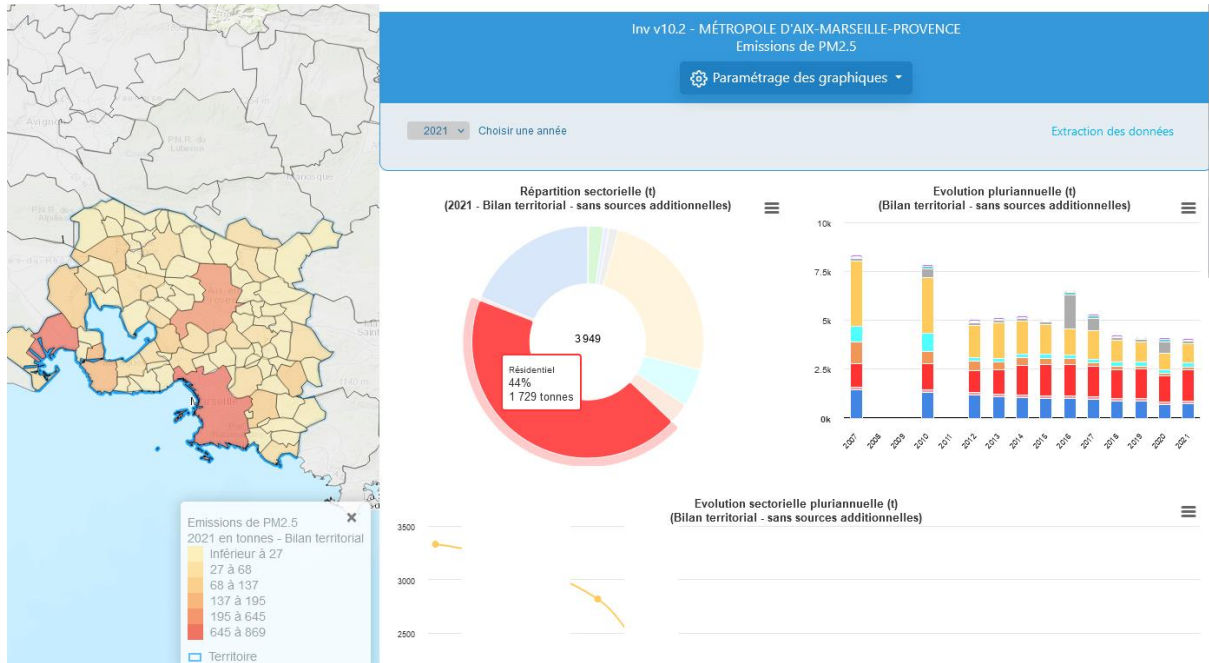


Illustration 1 : Visualisation de l'inventaire sur la plateforme CIGALE

A noter : L'inventaire est calculé avec un décalage de minimum 2 ans. Pendant mon stage à partir d'avril 2024, je travaille donc principalement avec l'inventaire 2021.

Il m'a donc fallu me renseigner sur toutes ces notions essentielles à l'inventaire des émissions. Naturellement, il a même fallu comprendre le plus en profondeur possible toutes les subtilités de l'inventaire annuel des émissions déjà au point chez AtmoSud. Dans cette mesure, j'ai réalisé plusieurs tâches.

D'abord, mon tuteur m'a fourni une liste de requêtes à effectuer pour manipuler l'inventaire :

S'entraîner à récupérer des données dans la table `total.bilan_comm_v10_diffusion`  
**Attention, uniquement des « select » (pas de drop, delete, insert)**

- Sélectionner les émissions totales de NOx du secteur « Résidentiel » de la région, en 2021 (afficher l'unité dans un champ)
  - Pour toutes les lignes qu'il y a dans la table (juste afficher le total)
  - Pour toute la région uniquement (la région contient les départements 4,5,6,13,83,84) → astuce : pour la commune 13035 par exemple, si je fais `id_comm/1000` j'obtiens 13
  - Pour le département 13 uniquement, avec le détail par catégorie d'énergie (avoir le bon libellé)
  - Par commune de l'EPCI suivant : CA Sud Sainte Baume
- Pour la commune 6088, la part des émissions de PM10 liée à la classe d'usage « Chauffage », pour toutes les années disponibles (tous secteurs confondus)

Aide :  
 Dans le schéma « commun », il y a de nombreuses tables dictionnaires « tpk », elles pourront être utiles pour faire des jointures pour les polluants, les secteurs, les communes, etc.  
 Idem, il y a des tables tpk dans les schémas « total » et « transversal »  
 Pour le choix du secteur, utiliser le champ `id_secteur_detail`  
 Pour la commune, la table `tpk_comm_evol`  
 Faire apparaître l'unité dans un champ à part

Pour le 1. tu peux vérifier les résultats avec ce qui est affiché sur CIGALE

Illustration 2 : Exercice SQL sur l'inventaire

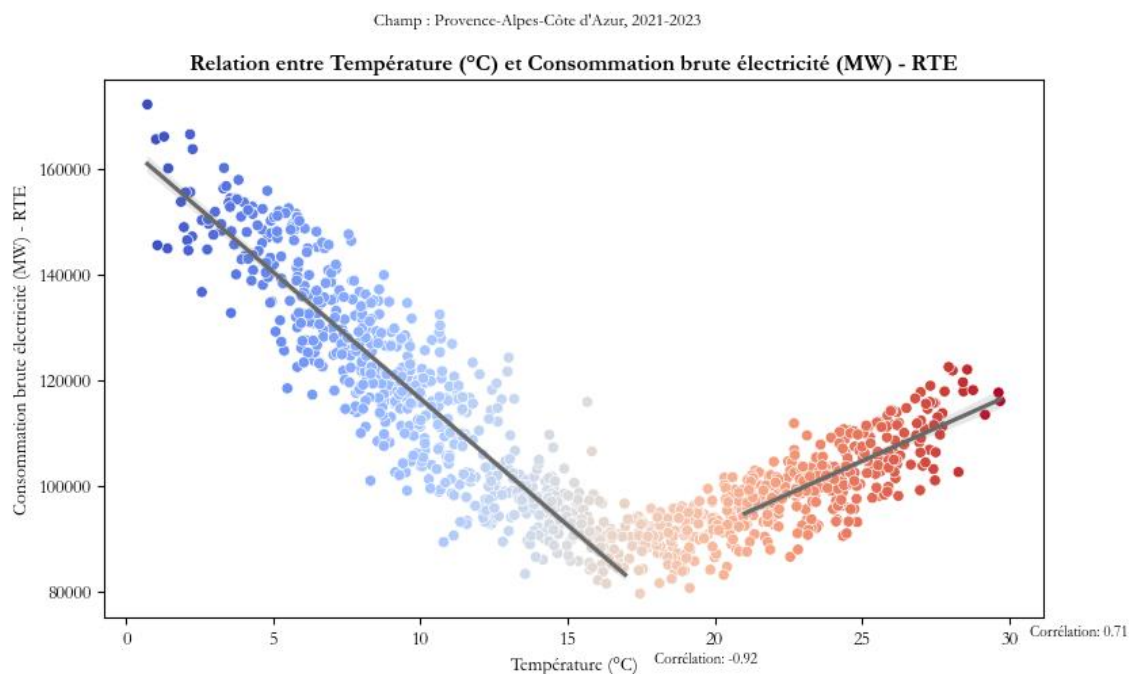
Une bonne idée me permettant de me familiariser avec la table inventaire, ses nombreuses colonnes, ce que représente une ligne, et de mémoriser des conventions de nommage dont j'aurai besoin ensuite tout au long de mon stage.

Ensuite, j'eus accès à la documentation de l'inventaire annuel des émissions. Il me fallut bien la lire, pour comprendre comment AtmoSud fait au pas de temps annuel ce que je devrai faire au pas de temps journalier.

J'ai alors eu la chance de discuter avec Emeline, autrice de cette documentation, pour de meilleures explications, mais aussi avec Sonia qui m'a parlé du lien entre l'inventaire et les modélisations qui en découlent, ou encore avec Morgane qui m'a parlé du cadastre, une version encore plus précise que l'inventaire en terme d'échelle spatiale.

## 1.2 - Relation entre température et consommation

Passer l'inventaire annuel en temps réel (journalier) nécessite d'inclure de l'information sur chaque jour pour différencier les jours entre eux. Soyons plus concrets : on ne peut pas annoncer des émissions (et donc, à la base, des consommations) liées à l'usage Chauffage identiques pour un jour d'hiver et un jour d'été. On pourrait donc essayer de se servir de la température du jour pour moduler les consommations d'usages thermosensibles comme le Chauffage, la Climatisation ou encore l'ECS (Eau Chaude Sanitaire). Regardons d'abord si la relation entre Température et Consommation est suffisamment forte :



Source : OPENDATA RESEAUX-ENERGIES, Données SYNOP essentielles OMM

Figure 1 : Relation entre température et consommation électrique quotidienne en région PACA

Sur des données de consommation électrique tous secteurs confondus comme ici issues de RTE, on observe une relation très forte avec la température moyenne en région. Lorsqu'il fait froid, plus il fait froid plus la consommation est élevée, notamment car les gens utilisent plus de chauffage. C'est

lorsqu'il ne fait ni chaud ni froid que la consommation est la plus faible. Et lorsqu'il fait chaud, plus il fait chaud plus la consommation est élevée car les gens utilisent plus de climatisation.

Une autre façon de visualiser cette relation est de regarder l'évolution des consommations et des températures au fil de l'année :

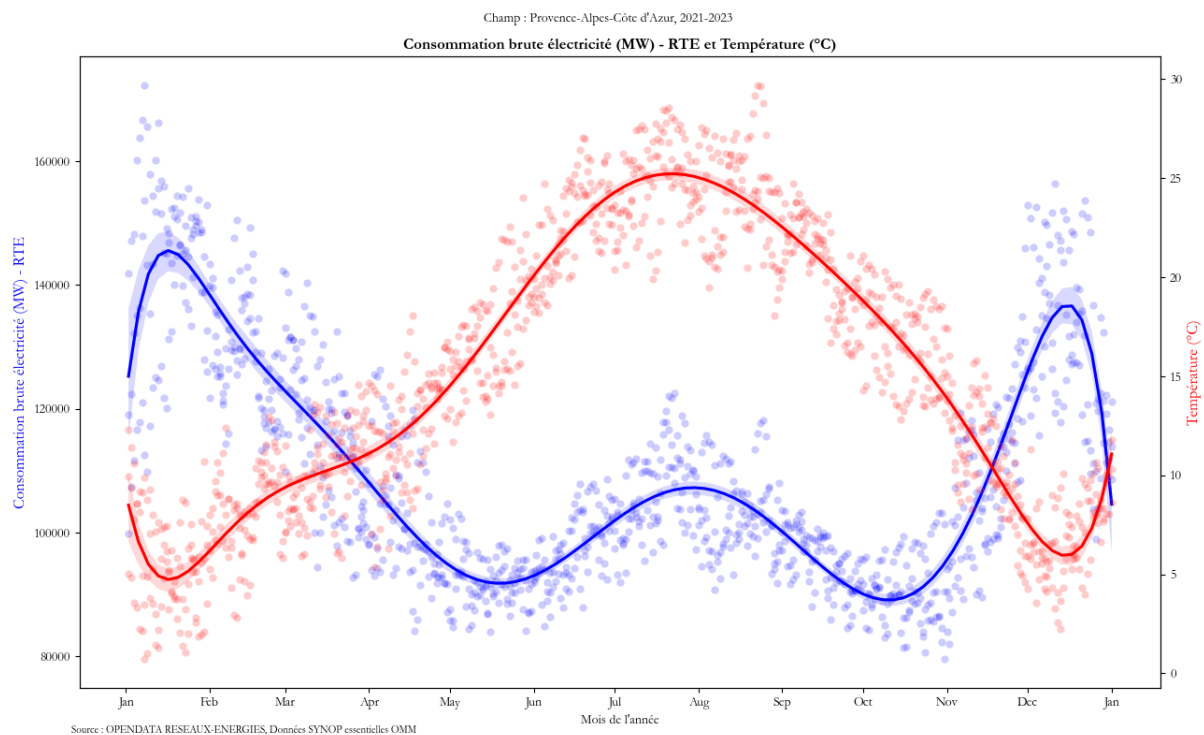


Figure 2 : Température et Consommation électrique en région PACA au fil de l'année

L'hiver, la température est basse et la consommation élevée en raison du chauffage notamment. Lorsqu'arrive le printemps, la température monte et la consommation diminue jusqu'à son niveau minimal, avant de remonter en été lorsqu'il fait chaud, notamment à cause de la climatisation.

Plus tard, on constatera que cette relation Consommations-Températures fonctionne au pas de temps journalier mais pas horaire :

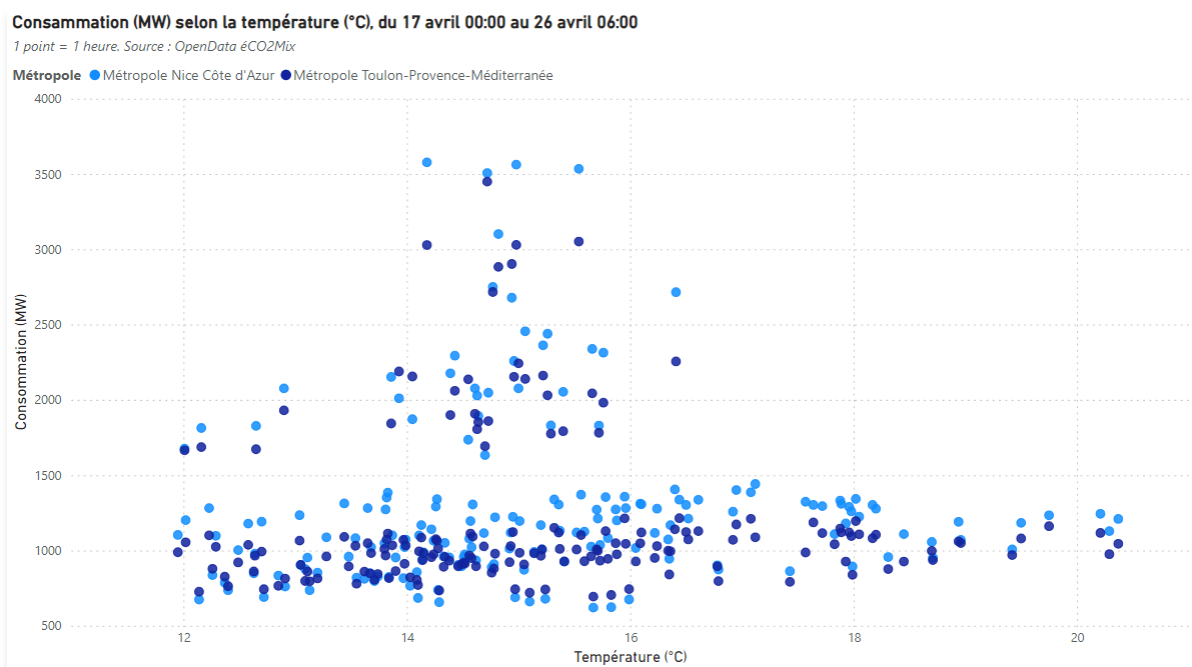


Figure 3 : Relation entre température et consommation horaire dans 2 métropoles

Les températures les plus basses surviennent la nuit, et pour d'autres raisons (les gens dorment), la consommation y est faible. La consommation est la plus haute le matin et le soir, à des températures intermédiaires. Ceci tend à diriger davantage le projet vers un pas de temps journalier et non horaire.

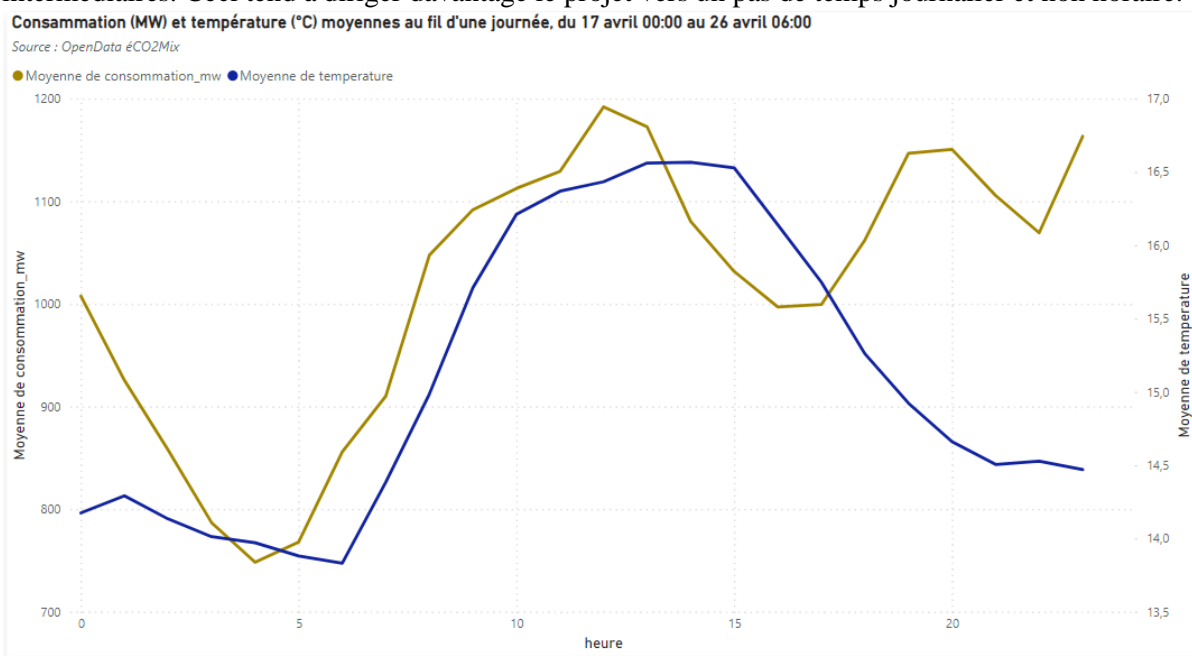


Figure 4 : Evolution de la température et de la Consommation horaire à Nice et Toulon

En fin de compte, la relation entre température et consommation électrique, lorsqu'observée au pas de temps journalier, paraît à ce stade très intéressante pour notre projet.

## 1.3 - Recherche de données en temps réel

Idéalement dans ce projet, nous aurions besoin de données :

- Accessibles via une API
- en « temps réel », c'est-à-dire chaque jour
- à l'échelle communale
- par secteur, pour filtrer sur le Résidentiel

D'abord, nous cherchons cela dans les sources OpenData.

#	Titre	Energie	Maille géographique	Pas temporel	MAJ	Lien	Par secteur ?	Infos supplémentaires	Commentaire
1	Données #CO2mix régionales temps réel	Electricité	Région	1/4 d'heure	1 par heure	<a href="https://odre.opds">https://odre.opds</a>	Non	découpage selon l'origine de production disponible	Bon jeu de données sur la conso électrique, mais données régionales = trop large
2	Consommation d'électricité des grandes Métropoles françaises temps réel	Electricité	EPCI (seulement métropoles)	1/4 d'heure	4 par jour	<a href="https://odre.opds">https://odre.opds</a>	Non	-	Ressemble au 1. en un peu + localisé
3	Consommation annuelle résidentielle par adresse	Electricité	Adresse (de 10 logements ou +)	Annuel (2018-2022)	31/12/2022	<a href="https://www.de">https://www.de</a>	Résidentiel	-	Bien localisé mais annuel, pas d'intérêt ?
4	Puissances Électriques au pas demi-heure	Electricité	Nationale (tout Enedis)	1/2 heure	1 par jour	<a href="https://openda">https://openda</a>	Non	Température	
5	Énergies consommées par commune - indicateur	Bois	Commune	Annuel (2016)	16/09/2019	<a href="https://www.d">https://www.d</a>	Chaudières bois	-	Chaudières collectives seulement, il y en a très peu
6	Consommation annuelle d'électricité et gaz par commune et par secteur d'activité (jusqu'en 2021)	Electricité, Gaz	Commune	Annuel (2011-2021)	1 par an	<a href="https://openda">https://openda</a>	Oui	-	
7	Consommation quotidienne régionalisée & définitive de gaz des clients industriels raccordés sur le réseau de transport de gaz	Gaz	Région	Horaire	1 par mois	<a href="https://odre.opds">https://odre.opds</a>	Industriel	-	
8	Consommation quotidienne définitive & régionalisée de gaz des distributions publiques à l'interface avec le réseau de transport de gaz	Gaz	Région	Horaire	1 par mois	<a href="https://odre.opds">https://odre.opds</a>	Non	-	
9	Extremes de consommation quotidienne définitive de gaz à la maille régionale relevée sur le réseau de GRT gaz	Gaz	Région	Journalier	1 par mois	<a href="https://odre.opds">https://odre.opds</a>	Non	-	
10	Consommation quotidienne brute régionale	Electricité, Gaz	Région	1/2 heure	1 par mois	<a href="https://odre.opds">https://odre.opds</a>	Non	-	Ressemble au 1. avec gaz

Tableau 1 : Jeux de données Opendata trouvés

On ne trouve pas tout à fait ce que l'on souhaite. En raison de règles de confidentialité et de quotas minimaux de points de données dans cette optique, en général plus l'échelle géographique est précise, plus le pas temporel est grand (exemple : données à l'adresse mais annuelles), et plus le pas temporel est resserré, moins l'échelle géographique est précise (données horaires mais sur toute la région).

Ainsi, le jeu de données qui nous semble le moins éloigné de l'objectif est ici le 2°, des données de consommation électrique quart-horaires pour les grandes métropoles (dont les 3 de la région, Nice Toulon et Aix-Marseille). Mais ces données sont tous secteurs confondus, on ne travaille donc pas sur le secteur résidentiel avec ces données. Par exemple, l'Industrie qui est un secteur majeur de consommation, fausserait nos analyses restreintes au Résidentiel.

Le constat est rapide : il nous faut des données plus précises. Pour cela, AtmoSud peut contacter Enedis et lui proposer un partenariat afin d'avoir accès à des données confidentielles. Faute de succès rapide dans cette entreprise, il nous faudra trouver une autre méthode provisoire pour remplacer les données réelles par des estimations.

## II – Test de calcul sur les métropoles

### 2.1 - Objectif

Dans le cadre du projet visant à estimer les émissions du secteur résidentiel/tertiaire en temps réel, il s'agit ici d'une phase de test avec des données d'input « temps réel » insuffisantes ; en effet nous travaillerons ici à partir de données de consommation électrique quart-horaires disponibles seulement pour les métropoles (Nice, Toulon, et Aix-Marseille), **sans différenciation par secteur**. L'idée est donc de tester et de mettre en œuvre une méthode de calcul imaginée, pour être prêts lorsque des données plus pertinentes (par commune, par secteur) nous seront accessibles.

Ainsi, on souhaite calculer, à la date d'aujourd'hui dans la commune c (en l'occurrence la métropole m dans cette phase de test), la quantité émise du polluant p pour l'usage u via l'énergie e. Le résultat final ressemblera alors à ceci :

	date	123 id_metropole	123 id_polluant	123 id_usage	123 code_cat_energie	123 consommation_usage_energie_mwh	123 facteur_emission_kg_by_mwh	123 emission_tps_reel_kg
1	2024-05-26	6100	2	0	1	2 641,2767222365	0,0000000465	0,0001228134
2	2024-05-26	6100	2	0	2	1 048,8011476154	0,0000000066	0,0000068966
3	2024-05-26	6100	2	0	4	763,5172127737	0,0000203181	0,0155132386
4	2024-05-26	6100	2	1	1	3 128,995153882	0,0000000459	0,0001437222
5	2024-05-26	6100	2	1	2	216,3708391942	0,0000000046	0,0000010003
6	2024-05-26	6100	2	1	4	3,6038841089	0,0000407363	0,000146809
7	2024-05-26	6100	2	2	1	1 937,1189003413	0,0000000453	0,0000878063
8	2024-05-26	6100	2	2	2	173,6678926999	0,0000000003	0,0000000593
9	2024-05-26	6100	2	2	4	13,3897408947	0,0000407376	0,0005454665
10	2024-05-26	6100	2	35	2	41,8017672005	0,0000000252	0,0000010559
11	2024-05-26	6100	2	999	1	599,4965206499	0,0000000529	0,0000317128
12	2024-05-26	6100	2	999	2	278,8923753163	0,0000000065	0,0000018207
13	2024-05-26	6100	2	999	4	5,5793781034	0,0000407589	0,0002274095
14	2024-05-26	6100	3	0	1	2 641,2767222365	0,0006652319	1,7570615984
15	2024-05-26	6100	3	0	2	1 048,8011476154	0,0001901639	0,1994441533
16	2024-05-26	6100	3	0	4	763,5172127737	0,092898071	70,929276269

Tableau 2 : Résultat final pour les métropoles (head)

Nous détaillerons toutes les étapes sous-jacentes du calcul au fil de cette partie II.

## 2.2 - Présentation générale de la méthode

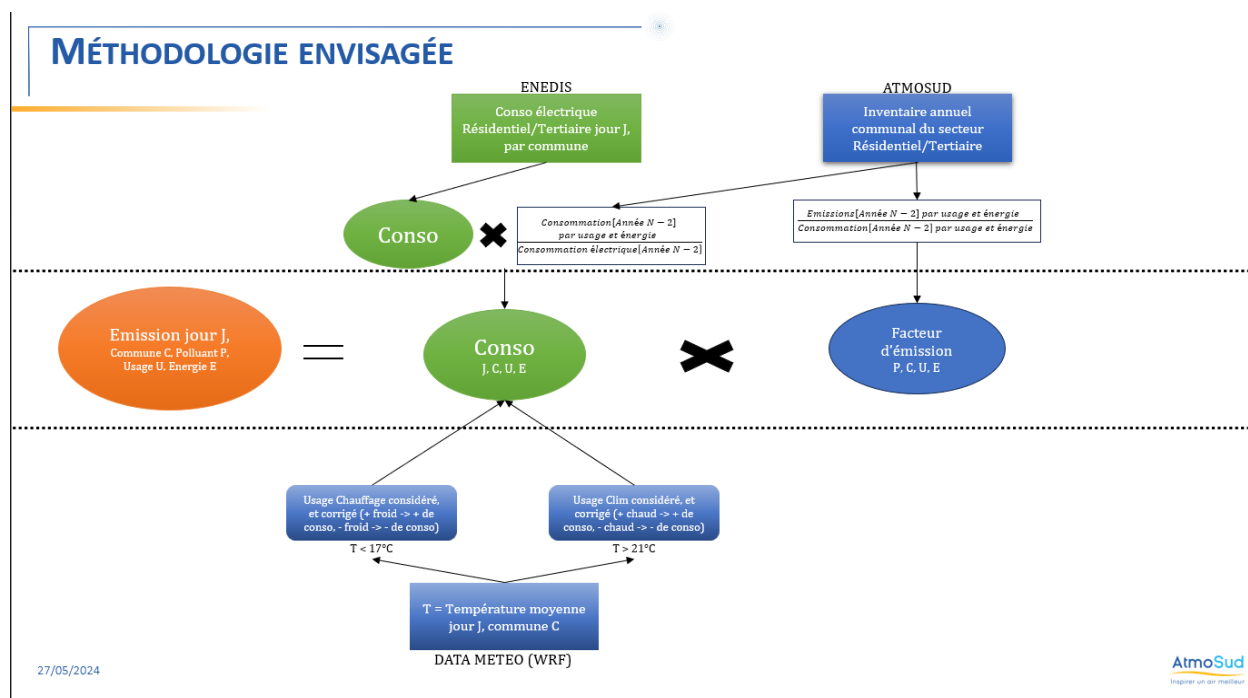
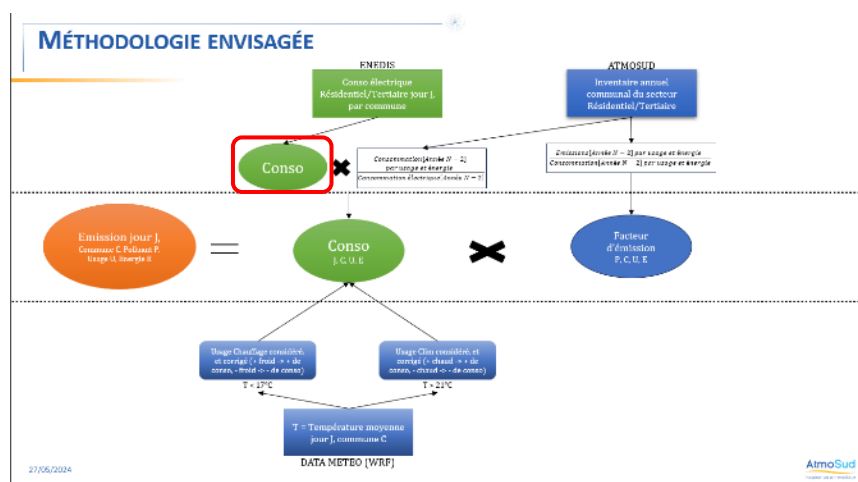


Illustration 3 : Méthode envisagée

Les rectangles tout en haut et tout en bas représentent les données sources sur lesquelles s'appuie le calcul. Sur la ligne du milieu on retrouve l'équation principale : Emission (kg) = Consommation (MWh) \* Facteur d'émission (kg / MWh)

Comme on peut le voir, le calcul d'une consommation journalière par commune, usage, et énergie, corrigé de la température selon les usages, représente la partie la plus complexe. L'application d'un facteur d'émission adapté à cette consommation, pour passer d'une consommation à une émission, est ensuite triviale.

## 2.3 - Extraction des données de consommation





Source : [https://odre.opendatasoft.com/explore/dataset/eco2mix-metropoles-tr/api/?disjunctive.libelle\\_metropole&disjunctive.nature](https://odre.opendatasoft.com/explore/dataset/eco2mix-metropoles-tr/api/?disjunctive.libelle_metropole&disjunctive.nature) ; Données de consommation électrique (appel instantané) en MW à chaque quart d'heure (on fait donc la moyenne horaire pour passer en MWh) dans la métropole de Nice, Toulon, Aix-Marseille. MAJ toutes les 6h. Accessible via une API.

L'extraction se fait via un script Python qui utilise l'API mis à disposition, traite les données (colonnes, formats etc..) et les insère dans une table PostgreSQL sur le serveur interne d'AtmoSud.

	date_heure	date	heure	code_insee_epci	libelle_metropole	consommation_mw
1	2024-07-01 07:15:00.000	2024-07-01	07:15	248 300 543	Métropole Toulon-Provence-Méditerranée	222
2	2024-07-01 07:00:00.000	2024-07-01	07:00	248 300 543	Métropole Toulon-Provence-Méditerranée	210
3	2024-07-01 06:45:00.000	2024-07-01	06:45	200 030 195	Métropole Nice Côte d'Azur	164
4	2024-07-01 06:45:00.000	2024-07-01	06:45	248 300 543	Métropole Toulon-Provence-Méditerranée	210
5	2024-07-01 06:30:00.000	2024-07-01	06:30	200 030 195	Métropole Nice Côte d'Azur	157
6	2024-07-01 06:30:00.000	2024-07-01	06:30	248 300 543	Métropole Toulon-Provence-Méditerranée	206
7	2024-07-01 06:15:00.000	2024-07-01	06:15	200 030 195	Métropole Nice Côte d'Azur	157
8	2024-07-01 06:15:00.000	2024-07-01	06:15	248 300 543	Métropole Toulon-Provence-Méditerranée	197
9	2024-07-01 06:00:00.000	2024-07-01	06:00	200 030 195	Métropole Nice Côte d'Azur	150
10	2024-07-01 06:00:00.000	2024-07-01	06:00	248 300 543	Métropole Toulon-Provence-Méditerranée	194
11	2024-07-01 05:45:00.000	2024-07-01	05:45	200 054 807	Métropole d'Aix-Marseille-Provence	1406
12	2024-07-01 05:45:00.000	2024-07-01	05:45	200 030 195	Métropole Nice Côte d'Azur	154
13	2024-07-01 05:45:00.000	2024-07-01	05:45	248 300 543	Métropole Toulon-Provence-Méditerranée	191
14	2024-07-01 05:30:00.000	2024-07-01	05:30	200 054 807	Métropole d'Aix-Marseille-Provence	1438
15	2024-07-01 05:30:00.000	2024-07-01	05:30	200 030 195	Métropole Nice Côte d'Azur	143

Tableau 3 : Données de puissance quart-horaires (head)

On transforme ensuite cette table pour passer en pas de temps journalier. Pour cela :

- ➔ On calcule des consommations horaires en MWh, en prenant la moyenne pour chaque heure des consommations instantanées quart-horaires en MW
- ➔ On somme les 24 consommations horaires en MWh pour avoir une consommation journalière en MWh

	date	libelle_metropole	id_metropole	consommation_mwh	missing_percentage
1	2024-07-01	Métropole d'Aix-Marseille-Provence	13000	8702	75
2	2024-07-01	Métropole Nice Côte d'Azur	6100	1167	70,83
3	2024-07-01	Métropole Toulon-Provence-Méditerranée	83100	1633	66,67
4	2024-06-30	Métropole d'Aix-Marseille-Provence	13000	38405	0
5	2024-06-30	Métropole Nice Côte d'Azur	6100	5260	0
6	2024-06-30	Métropole Toulon-Provence-Méditerranée	83100	5720	0
7	2024-06-29	Métropole d'Aix-Marseille-Provence	13000	39435	0
8	2024-06-29	Métropole Nice Côte d'Azur	6100	5471	0
9	2024-06-29	Métropole Toulon-Provence-Méditerranée	83100	5866	0
10	2024-06-28	Métropole d'Aix-Marseille-Provence	13000	40930	0
11	2024-06-28	Métropole Nice Côte d'Azur	6100	5720	0
12	2024-06-28	Métropole Toulon-Provence-Méditerranée	83100	6009	0

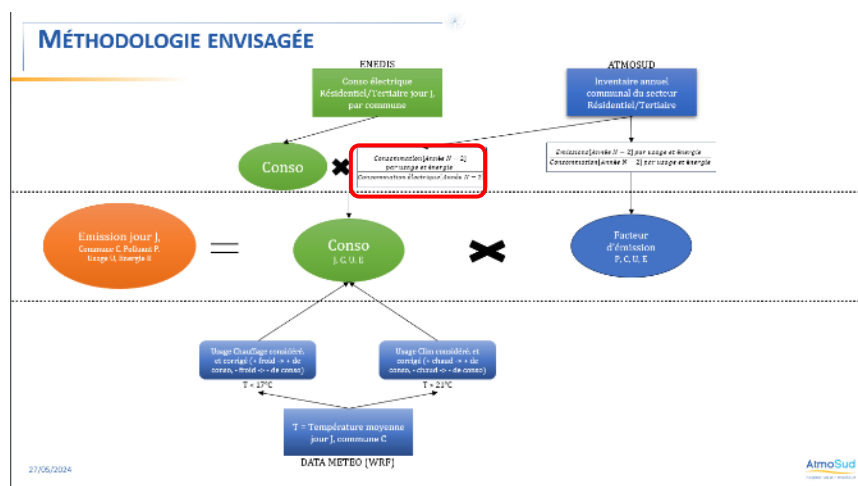
Tableau 4 : Consommations journalières en temps réel

Notons :

- Ces données de consommation reflètent **tous les secteurs confondus**. Elles ne sont pas par secteur. Nous ne sommes malheureusement pas restreints au Résidentiel ici.
- La présence d'une colonne *missing\_percentage* qui indique le pourcentage d'heures manquantes [nombre d'heures manquantes / 24] au sein d'une journée. C'est d'abord utile pour

rendre compte de la (non) pertinence des données d'aujourd'hui, étant donné qu'elles ne sont pas complètes.

## 2.4 – Ratio de consommation



Le ratio de consommation a le rôle de « transformer » une consommation électrique en une consommation par usage (Chauffage, Cuisson, Froid etc..) et par énergie (Bois, Gaz, Electricité).

Un ratio de consommation est donc calculé de la façon suivante pour chaque combinaison de [métropole, usage, énergie] :

$$ratio\_conso = \frac{Conso_{métropole,usage,energie \text{ de l'année } N-2}}{Conso_{métropole,energie=Electricité \text{ de l'année } N-2}}$$

A noter que le calcul des ratios de consommation prend sa source dans l'inventaire AtmoSud de l'année N-2 (la plus récente, 2021 ici), et ne nécessite pas d'autre source.

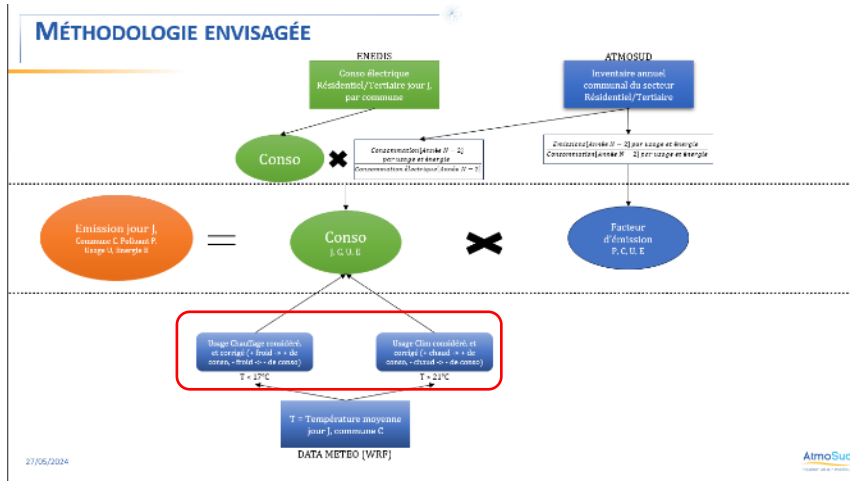
w_ratios_conso_metropoles 1 (2) ×									
drop table if exists w_ratios_conso; create temp table w_ratios_conso (									
	id_metropole	id_usage	usage_old	code_cat_energie	energie_old	consommation	id_unite	consommation_elec	ratio_conso
1	6100	0	Chauffage	1	Gaz Naturel	741 603,9699686237	-30007	3 132 355,172646486	0,2367560283
2	6100	0	Chauffage	2	Produits pétroliers	294 476,9429992227	-30007	3 132 355,172646486	0,0940113514
3	6100	0	Chauffage	4	Bois-énergie	214 376,4003844571	-30007	3 132 355,172646486	0,0684393655
4	6100	0	Chauffage	6	Chaleur et froid	56 360,9611407004	-30007	3 132 355,172646486	0,0179931579
5	6100	0	Chauffage	8	Electricité	377 224,2483546642	-30007	3 132 355,172646486	0,1204283127
6	6100	1	ECS	1	Gaz Naturel	366 138,2816200972	-30007	3 132 355,172646486	0,1168891334
7	6100	1	ECS	2	Produits pétroliers	25 318,5586295842	-30007	3 132 355,172646486	0,0080829144
8	6100	1	ECS	4	Bois-énergie	421,70724782	-30007	3 132 355,172646486	0,0001346294
9	6100	1	ECS	8	Electricité	472 999,4391329248	-30007	3 132 355,172646486	0,1510044082
10	6100	2	Cuisson	1	Gaz Naturel	226 671,2956026214	-30007	3 132 355,172646486	0,0723644935
11	6100	2	Cuisson	2	Produits pétroliers	20 321,6881709832	-30007	3 132 355,172646486	0,0064876705
12	6100	2	Cuisson	4	Bois-énergie	1 566,7958822019	-30007	3 132 355,172646486	0,0005001974
13	6100	2	Cuisson	8	Electricité	203 440,8502089938	-30007	3 132 355,172646486	0,0649482064
14	6100	3	Electricité spécifique	8	Electricité	678 955,09872777	-30007	3 132 355,172646486	0,216755464
15	6100	4	Chauffage urbain	6	Chaleur et froid	22 251,3903524072	-30007	3 132 355,172646486	0,0071037252

Tableau 5 : ratios de consommation

## 2.5 – Correction selon la température

A ce stade, nous sommes en mesure d'avoir une consommation quotidienne par usage et énergie de chaque métropole. Cependant, la ventilation d'une consommation journalière par usage se fait à partir de la répartition par usage dans l'inventaire annuel d'AtmoSud. Par exemple, si l'usage Chauffage

représentait sur l'année 2021 30% des consommations, alors on attribuerait aujourd'hui 30% de la consommation du jour à l'usage Chauffage, que l'on soit en hiver ou en été. Ce n'est pas correct. Il nous faut donc corriger les consommations par usage du jour, et pour cela la température du jour s'avère utile.



Concentrons-nous sur la correction de l'usage Chauffage. La méthode sera la même pour la Climatisation. On fixe une Date J et une Commune C.

## Résumé

$$SUM_{chauff1} = \sum_{usage = Chauffage} ratio$$

$$\overline{SUM_{chauff1}} = \sum_{usage \neq Chauffage} ratio$$

Correction des ratios liés au Chauffage :

$$ratio_{Corrige_{[U=Chauffage]}} = ratio_{[U=Chauffage]} * \min \left( \frac{DJ17}{Base_{DJ17}(C)} ; \frac{\overline{SUM_{chauff1}}}{SUM_{chauff1}} + 1 - 0.1 \right)$$

$$Diff_{Correction} = SUM_{chauff2} - SUM_{chauff1}$$

Rééquilibrage avec les ratios non liés au chauffage :

$$ratio_{Corrige_{U \neq Chauffage}} = ratio_{U \neq Chauffage} - \frac{ratio_{U \neq Chauffage}}{\overline{SUM_{chauff1}}} * Diff_{Correction}$$

- ➔ On effectue cette correction simultanément pour le Chauffage et la Climatisation :  
En utilisant  $DJ21 = \max(0 ; \text{Temperature} - 21)$  et  $Base\_DJ21$ .  
-On corrige les ratios de Chauffage et les ratios de Climatisation  
-On identifie  $Diff\_Correction$  par la différence entre la somme des ratios et la somme des ratios corrigés  
-On rééquilibre avec tous les ratios qui ne sont liés ni au Chauffage ni à la Climatisation
- ➔ On effectue cette correction d'abord uniquement sur les ratios liés à l'énergie Electricité. Ceci permet d'encapsuler les ratios de consommation liés à l'Electricité (dont la somme doit faire 1 car c'est l'énergie référence) pour que leur somme reste la même avant et après correction.
- ➔ Dans un second temps, on relance cette correction sur les ratios liés à une autre énergie que l'Electricité.

## Explications et détail

Ici on concentre l'explication sur la correction pour le Chauffage, omettant la Climatisation qui s'inclura simplement dans la même méthode. On prend aussi toutes les énergies confondues, alors qu'en pratique on fera la correction en 2 fois [ 1) Electricité 2) Autres ].

L'idée première et essentielle est de multiplier les ratios de consommation concernés par l'usage « Chauffage » par un facteur dépendant de la température, avec la correction suivante :

$$ratioCorrige_{usage=Chauffage} = ratio_{usage=Chauffage} * \frac{DJ17}{Base_{DJ17}(C)}$$

Où  $DJ17 = \text{Max}(0, 17 - \text{Température}^{\circ}\text{C})$  représente le besoin en chauffage

Et  $Base_{DJ17}(C)$  est la moyenne du besoin en chauffage lorsqu'il y a besoin de chauffage (Température  $< 17^{\circ}\text{C}$ ), dans la commune C sur l'année 2023.

Expliquons par l'exemple.

Si en 2023 quand il y a besoin de chauffage (température  $< 17^{\circ}\text{C}$ ) la moyenne des températures est de  $10^{\circ}\text{C}$ ,  $Base_{DJ17}$  vaut 7.

Ainsi, si aujourd'hui il fait également  $10^{\circ}\text{C}$  ce qui donne un DJ17 de 7,  $\frac{DJ17}{Base_{DJ17}(C)} = 1$  ; les ratios liés au chauffage sont donc inchangés.

Si aujourd'hui il fait  $20^{\circ}\text{C}$ , DJ17 vaut 0 et les ratios liés au chauffage sont annulés ; en effet on considère qu'à cette température il n'y a pas de chauffage résidentiel.

Si aujourd'hui il fait  $3^{\circ}\text{C}$ , DJ17 vaut 14, soit 2 fois plus que  $Base_{DJ17}$ . Les ratios liés au Chauffage seront donc multipliés par 2, étant donné que le besoin en Chauffage est ici plus élevé que la moyenne.

On doit maintenant s'arranger pour que ces corrections ne modifient pas le volume total de consommation que l'on considèrera pour la commune C à cette date J, c'est-à-dire que la somme des ratios de consommation doit être la même avant et après correction. Pour cela, si on augmente les ratios liés au Chauffage, il faudra donc diminuer en conséquence les ratios non liés au chauffage pour retomber sur la même somme. On procède comme suit :

$$SUMchauff1 = \sum_{usage=Chauffage} ratio \qquad \frac{SUMchauff1}{SUMchauff1} = \sum_{usage \neq Chauffage} ratio$$

$$SUMchauff2 = \sum_{usage=Chauffage} ratioCorrige$$

$$Diff_{Correction} = SUMchauff2 - SUMchauff1$$

### Rééquilibrage :

$$ratioCorrige_{usage \neq Chauffage} = ratio_{usage \neq Chauffage} - \frac{ratio_{usage \neq Chauffage}}{SUMchauff1} * Diff_{Correction}$$

Avec cette formule, on diminue chaque ratio non lié au chauffage d'une portion de  $Diff_{Correction}$ , cette portion étant la part de ce ratio parmi tous les ratios non liés au chauffage. Ainsi, une fois le

rééquilibrage appliqué sur tous les ratios non liés au chauffage, on aura soustrait l'intégralité de  $Diff_{Correction}$ , comme souhaité pour conserver la même somme de ratios.

#### Contrainte de positivité :

Or, les ratios de consommation ne peuvent par définition pas être négatifs, au risque de donner des consommations négatives puis des émissions négatives. Il faut donc veiller à ce qu'ils ne soient pas réduits de trop lors du rééquilibrage.

$$ratio_{Corrige_{usage \neq Chauffage}} > 0$$

$$\Leftrightarrow Diff_{Correction} < \underline{SUMchauff1}$$

Comment s'arranger pour contraindre  $Diff_{Correction} < \underline{SUMchauff1}$  ?

$$Diff_{Correction} = SUMchauff2 - SUMchauff1$$

$$\begin{aligned} &= \sum_{U=Chauffage} \square ratio * \frac{DJ17}{Base_{DJ17}} - \sum_{U=Chauffage} \square ratio \\ &= \sum_{U=Chauffage} \square ratio * \left( \frac{DJ17}{Base_{DJ17}} - 1 \right) \\ &= \left( \frac{DJ17}{Base_{DJ17}} - 1 \right) * \sum_{U=Chauffage} \square ratio \end{aligned}$$

D'où  $Diff_{Correction} < \underline{SUMchauff1}$  équivaut à :

$$\begin{aligned} &\left( \frac{DJ17}{Base_{DJ17}} - 1 \right) * \sum_{U=Chauffage} \square ratio < \sum_{U \neq Chauffage} \square ratio \\ &\Leftrightarrow \frac{DJ17}{Base_{DJ17}} - 1 < \frac{\sum_{U \neq Chauffage} \square ratio}{\sum_{U=Chauffage} \square ratio} \\ &\Leftrightarrow \frac{DJ17}{Base_{DJ17}} < \frac{SUMchauff1}{SUMchauff1} + 1 \end{aligned}$$

On a donc notre **condition de positivité** sur le facteur de correction :

$$\frac{DJ17}{Base_{DJ17}} < \frac{SUMchauff1}{SUMchauff1} + 1$$

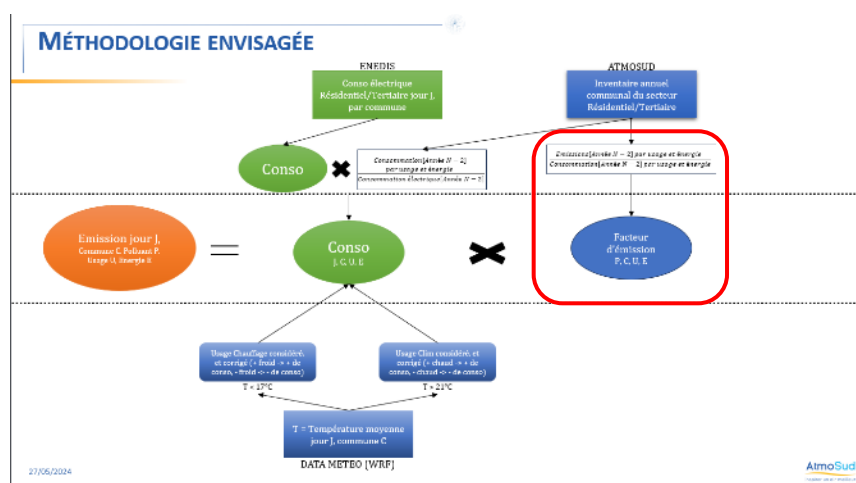
Si elle n'est pas respectée, alors les ratios non liés au chauffage seront négatifs pour compenser la hausse trop importante des ratios liés au chauffage. On doit donc limiter cette hausse.

Le facteur de correction  $\frac{DJ17}{Base_{DJ17}}$  devient  $\min \left( \frac{DJ17}{Base_{DJ17}} ; \frac{SUMchauff1}{SUMchauff1} + 1 - \varepsilon \right)$

Avec  $\varepsilon = 0.1$  par défaut.  $\varepsilon$  sert à éviter d'être exactement à la limite maximale d'augmentation du chauffage qui forcerait tous les ratios non liés au chauffage à valoir 0.

Ainsi, on se retrouve avec des ratios de consommation corrigés par usage en fonction de la température, qui nous permettent, par multiplication de la consommation d'électricité du jour par ces ratios, d'avoir des **consommations journalières par usage et par énergie adaptées**.

## 2.6 – Facteurs d’émission



Une fois que l’on a nos consommations du jour par commune, usage et énergie adaptées selon la température, il ne reste plus qu’à les multiplier par des facteurs d’émission pour savoir quelles émissions en résultent.

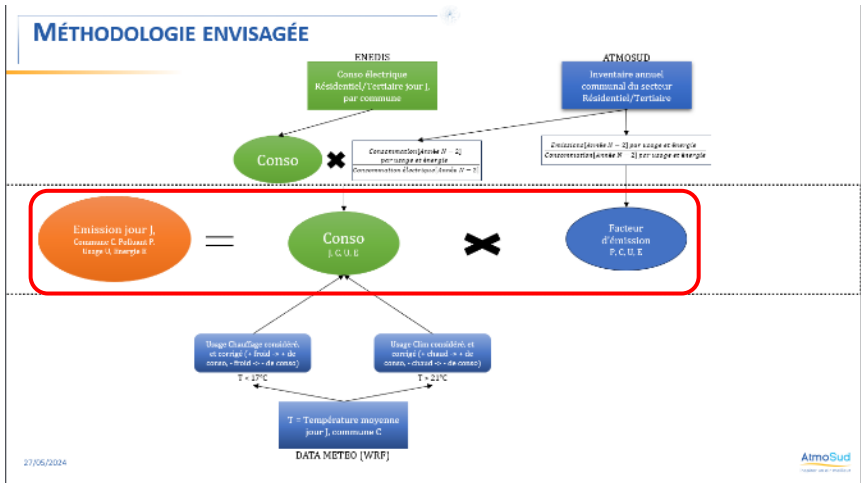
En se basant sur les données de l’inventaire 2021 d’AtmoSud, on calcule un facteur d’émission induit par polluant, commune, usage et énergie de la façon suivante :

$$FE_{Polluant;Commune;Usage;Energie} = \frac{Emissions[Année N - 2] (P, C, U, E)}{Consommation[Année N - 2] (P, C, U, E)}$$

w_facteurs_emiss_metropoles 1									
	id_metropole	id_polluant	nom_court_polluant	id_usage	code_cat_energie	emission_kg	consommation_mwh	facteur_emission_kg_by_mwh	
1	6100	2	As	0	1	0,0386483292	741 603,969986237	0,0000000521	
2	6100	2	As	0	2	0,0019363864	294 476,9429992226	0,0000000066	
3	6100	2	As	0	4	6,1298081752	214 376,4003844571	0,0000285937	
4	6100	2	As	1	1	0,0183582379	366 138,2816200972	0,0000000501	
5	6100	2	As	1	2	0,0001170472	25 318,5586295842	0,0000000046	
6	6100	2	As	1	4	0,6733459613	421,70724782	0,0015967142	
7	6100	2	As	2	1	0,0102746234	226 671,2956026215	0,0000000453	
8	6100	2	As	2	2	0,0000069336	20 321,6881709832	0,0000000003	
9	6100	2	As	2	4	0,0638275754	1 566,7958822019	0,0000407376	
10	6100	2	As	8	1	0,0336690515	147 073,5154749403	0,0000002289	
11	6100	2	As	8	2	0,0593370573	60 087,8044873038	0,0000009875	
12	6100	2	As	8	4	0,5694666298	16 862,1508209307	0,0000337719	
13	6100	2	As	0	2	0,002081022	507 309,120550222	0,0000000122	

Tableau 6 : facteurs d’émission (head)

## 2.7 – Emissions finales



Les émissions par jour, commune, polluant, usage et énergie se calculent en multipliant la consommation par le facteur d'émission correspondant.

w\_emissions\_tps\_reel\_test 1

Entrez une expression SQL pour filtrer les résultats (utilisez Ctrl+Espace)

	date	123_id_metropole	123_id_polluant	123_id_usage	123_cat_energie	123_consommation_usage_energie_mwh	123_facteur_emission_kg_by_mwh	123_emission_tps_reel_kg
1	2024-07-01	6100	2107	0107	1107	0	0,0000000521	0
2	2024-07-01	6100	2107	0107	2107	0	0,0000000066	0
3	2024-07-01	6100	2107	0107	4107	0	0,0000285937	0
4	2024-07-01	6100	2107	1107	1107	171,3034562122	0,0000000501	0,0000085892
5	2024-07-01	6100	2107	1107	2107	11,8456791253	0,0000000046	0,0000000548
6	2024-07-01	6100	2107	1107	4107	0,1973022562	0,0015967142	0,0003150353
7	2024-07-01	6100	2107	2107	1107	106,0516703935	0,0000000453	0,0000048071
8	2024-07-01	6100	2107	2107	2107	9,5078160206	0,0000000003	0,0000000032
9	2024-07-01	6100	2107	2107	4107	0,733049679	0,0000407376	0,0000298627
10	2024-07-01	6100	2107	8107	1107	68,810618236	0,0000002289	0,0000157526
11	2024-07-01	6100	2107	8107	2107	28,1130083949	0,0000009875	0,0000277618
12	2024-07-01	6100	2107	8107	4107	7,9842265667	0,0000337719	0,0002696424
13	2024-07-01	6100	2107	9107	2107	279,4135257044	0,0000001133	0,0000037233
14	2024-07-01	6100	2107	15107	1107	5,3639880913	0,0000000433	0,0000002325
15	2024-07-01	6100	2107	15107	2107	4,889796985	0,0000000013	0,0000000062
16	2024-07-01	6100	2107	16107	2107	1,6211510284	0,0000000012	0,0000000194
17	2024-07-01	6100	2107	23107	2107	1,2542489271	0,0000000086	0,0000000108

Tableau 7 : émissions (head)

La table complète retrace l'intégralité du calcul :

w\_emissions\_tps\_reel\_test 1

Entrez une expression SQL pour filtrer les résultats (utilisez Ctrl+Espace)

	date	123_id_metropole	123_id_polluant	123_id_usage	123_cat_energie	123_ratio_conso	123_conso_provisoire	123_temperature	123_ratio_conge	123_consommation_usage_energie_mwh	123_facteur_emission_kg_by_mwh	123_emission_tps_reel_kg
1	2024-07-01	6100	1107	2107	0107	1,07	0,28790033	276,294209514	21,50954888	0	0,0000000521	0
2	2024-07-01	6100	1107	2107	0107	2,07	0,940173514	109,712474142	21,50954888	0	0,0000000066	0
3	2024-07-01	6100	1107	2107	0107	4,07	0,940483915	79,887784810	21,50954888	0	0,0000285937	0
4	2024-07-01	6100	1107	2107	1107	1,07	0,116889134	136,409618735	21,50954888	0,14678994	0,0000000501	0,0000085892
5	2024-07-01	6100	1107	2107	1107	2,07	0,000009144	8,432701971	21,50954888	0,0101500391	0,0000000046	0,0000000548
6	2024-07-01	6100	1107	2107	1107	4,07	0,000134024	0,157112509	21,50954888	0,0001690679	0,0015967142	0,0003150353
7	2024-07-01	6100	1107	2107	2107	1,07	0,072344935	84,440034413	21,50954888	0,000774673	0,0000000453	0,0000048071
8	2024-07-01	6100	1107	2107	2107	2,07	0,000487025	7,371711444	21,50954888	0,0001470208	0,0000000003	0,0000000032
9	2024-07-01	6100	1107	2107	4107	1,07	0,0000001974	0,581703542	21,50954888	0,0006201488	0,0000407376	0,0000298627
10	2024-07-01	6100	1107	2107	8107	1,07	0,0495930137	54,794870402	21,50954888	0,000638831	0,0000002289	0,0000157526
11	2024-07-01	6100	1107	2107	8107	2,07	0,019102474	22,384996055	21,50954888	0,024099815	0,0000009875	0,0000277618
12	2024-07-01	6100	1107	2107	8107	4,07	0,0054480473	6,357811455	21,50954888	0,00081668	0,0000337719	0,0002696424
13	2024-07-01	6100	1107	2107	9107	2,07	0,190358178	222,449039459	21,50954888	0,234248995	0,0000001133	0,0000037233
14	2024-07-01	6100	1107	2107	15107	1,07	0,0039801342	4,2713648	21,50954888	0,045963808	0,0000000433	0,0000002325
15	2024-07-01	6100	1107	2107	15107	2,07	0,003330593	3,993784684	21,50954888	0,0041900574	0,0000000013	0,0000000062
16	2024-07-01	6100	1107	2107	16107	2,07	0,0011001945	1,290028566	21,50954888	0,0013891811	0,0000000012	0,0000000194
17	2024-07-01	6100	1107	2107	23107	2,07	0,0000338384	0,996703378	21,50954888	0,0010747634	0,0000000086	0,0000000108
18	2024-07-01	6100	1107	2107	24107	2,07	0,0000017441	0,0020254025	21,50954888	0,0000021903	0,0000000006	0,0000000001

Tableau 8 : calcul des émissions (head)



## 2.8 - Validation des résultats

Pour valider les résultats, on fait la somme des consommations établies grâce à la méthode présentée, toujours par caractéristiques (Commune, Usage, Energie), sur l'année 2021. On peut alors comparer les valeurs obtenues avec l'inventaire 2021 qui recense également ces consommations :

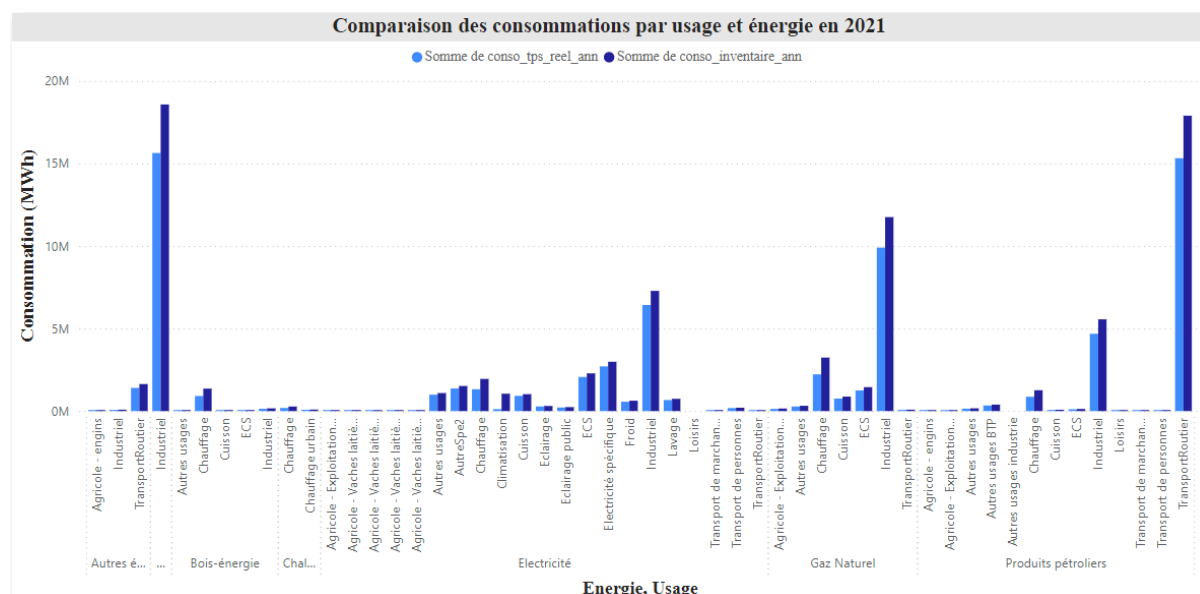


Figure 5 : Validation avec l'inventaire 2021 - Somme sur Aix-Marseille, Nice, Toulon

Pour y voir plus clair, faisons la même chose mais cette fois avec des ratios de consommation calculés uniquement sur le Résidentiel. Seuls les ratios par usage et énergie ont alors changé mais on reste bien sur des données et des estimations tous secteurs confondus. On retire la Métropole Aix-Marseille Provence dont la composante industrie impacte considérablement l'évolution des consommations :

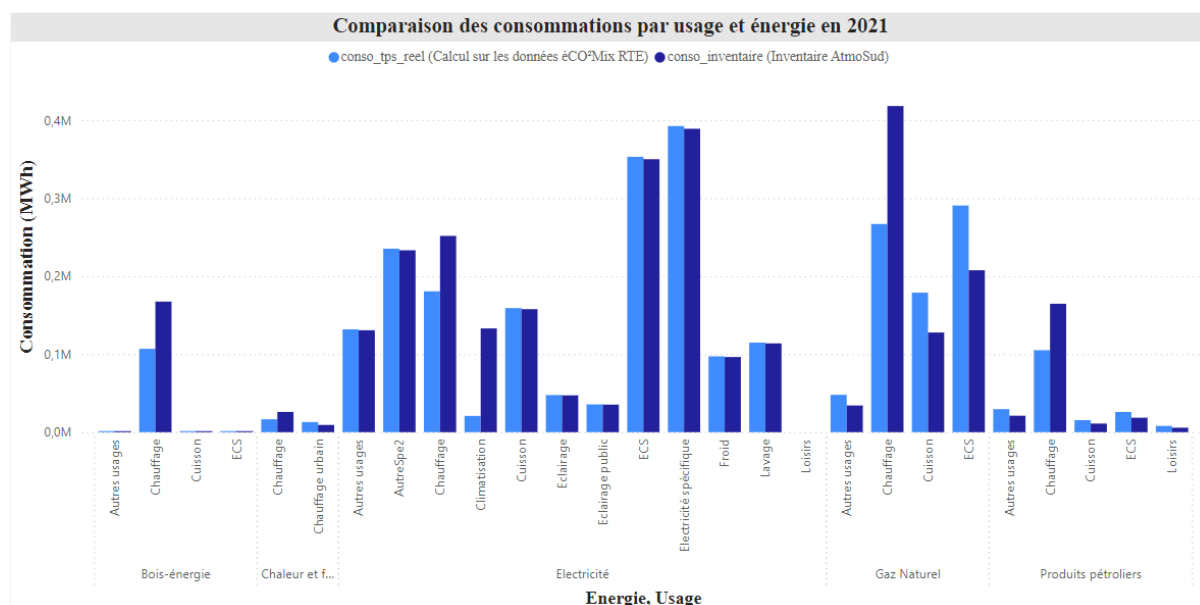


Figure 6 : Validation avec l'inventaire 2021, pour sur la Métropole Toulon Provence Méditerranée

# III - Estimation de la consommation électrique résidentielle à partir de la température (Machine Learning)

## 3.1 – Motivation et objectif

A cet instant du projet, nous avons donc développé une méthode de calcul des émissions en temps réel, que nous avons testé (en partie II) sur 3 métropoles avec des données de consommation électrique insuffisamment détaillées (tous secteurs confondus, maille géographique trop large). Il n'est donc pas question d'utiliser ces données en temps réel pour poursuivre le projet, seule la méthode développée conserve un intérêt pour la suite. Deux possibilités viendront alors :

- Accès à de meilleures données (par secteur, par commune) via Enedis ou autre énergéticien
- Pas de données de consommation électrique

Pour assurer nos arrières, il nous faut trouver un moyen de faire vivre le projet en l'absence de données de consommation électrique en temps réel pour le secteur résidentiel.

Comme **estimer la consommation électrique grâce à la température**. Il nous faudra toujours des données pour entraîner un tel modèle mais elles seront plus faciles à trouver car plus besoin de les avoir en temps réel. Un jeu de données fixe suffira pour entraîner un modèle que l'on pourra ensuite appliquer en temps réel sur la température du jour afin d'estimer la consommation électrique.

Cette partie III explique cette modélisation.

## 3.2 – Pertinence d'un modèle sur les métropoles

Dans un premier temps et pour effectuer la transition, l'on s'attache à « remplacer » les données de consommation des métropoles utilisées dans la partie II, on va donc modéliser la consommation électrique des métropoles tous secteurs confondus. Cela nous permettra de tester notre modèle en comparant les estimations avec les données de consommation des métropoles réellement observées. Dans ce rapport de stage, cette sous-partie est un résumé. Plus de détails seront fournis dans la modélisation finale sur le secteur résidentiel qui suivra à partir de 3.3.

On choisit ici d'utiliser 7 régressions polynomiales, 1 par jour de la semaine. On contrôle les valeurs aux extrêmes avec des régressions linéaires. Voici donc notre modélisation de la consommation électrique en région PACA tous secteurs confondus, en fonction de la température du jour.

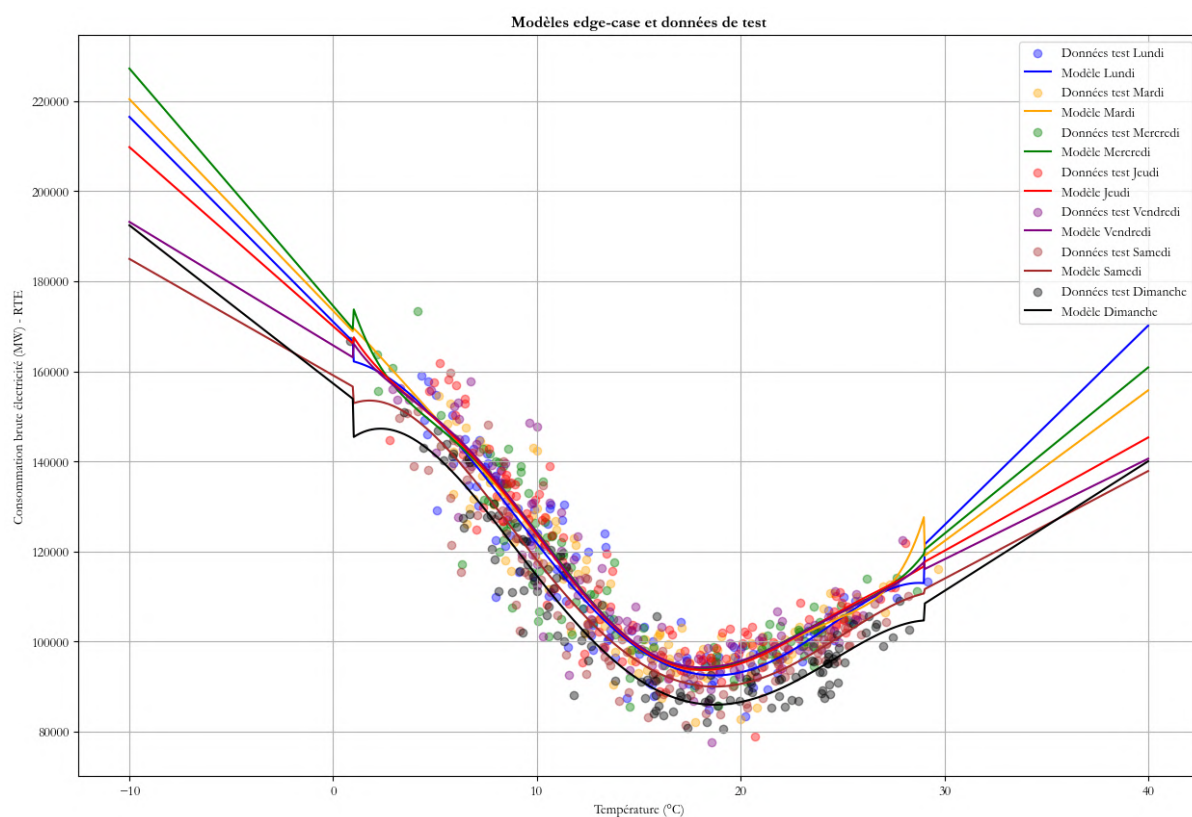
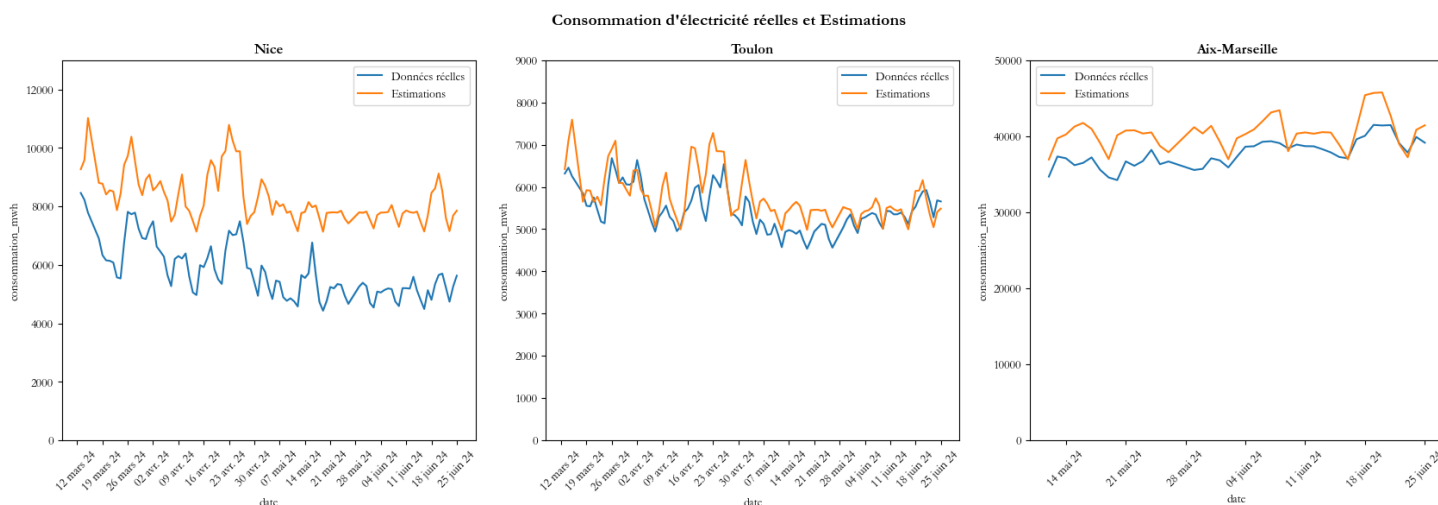


Figure 7 : Modélisation de la conso journalière en PACA, tous secteurs

On peut maintenant déployer ce modèle régional sur les communes/métropoles en appliquant un ratio de consommation communale.

Comparons maintenant les estimations de notre modèle en temps réel sur les 3 métropoles de la région avec les données réelles collectables en temps réel dans ces mêmes métropoles (source : éco<sup>2</sup>Mix – RTE).



Sources : éco<sup>2</sup>Mix RTE (Données réelles), table meteo\_routier\_tps\_reel (Estimations)

Figure 8 : Comparaison des données réelles éco<sup>2</sup>Mix et des estimations à Nice, Toulon, Aix-Marseille (métropoles)

On peut être plutôt très satisfaits par la capacité de nos estimations à bien suivre les tendances. Quant aux ordres de grandeur, pour ces données de test, on surestime légèrement dans les métropoles de Toulon et Aix-Marseille mais cela reste très satisfaisant compte tenu du fait que les données de test ne sont ici pas issues du même jeu de données que les données d'entraînement. A Nice, on surestime clairement la consommation.

Cela est dû au facteur *part\_conso\_metropole* (issu de l'inventaire) de Nice appliqué à la consommation en PACA, celui-ci est trop élevé pour faire correspondre les estimations avec les données en temps réel d'ÉCO<sup>2</sup>Mix.

En chiffres, on retrouve sur ces données test

- **une erreur RMSE** (Root Mean Squared Error) de 2103 MWh, soit **13.5%** du Root Mean Square des consommations.
- **Une erreur MAE** (Mean Absolute Error) de 1652 MWh, soit **15.9%** de la consommation moyenne.

En enlevant Nice qui surestime beaucoup :

- **une erreur RMSE** (Root Mean Squared Error) de 1681 MWh, soit **8.5%** du Root Mean Square des consommations.
- **Une erreur MAE** (Mean Absolute Error) de 1021 MWh, soit **7.4%** de la consommation moyenne.

- ➔ Ces performances peuvent être considérées **satisfaisantes** étant donné :
- La simplicité du modèle (transposable, explicable, rapide à entraîner)
  - La différence entre les jeux de données d'entraînement et de test en pratique

Voilà pour le modèle visant à remplacer les données de consommation tout secteur en temps réel de ÉCO<sup>2</sup>Mix. Ceci nous permet de confirmer que bien qu'imparfaite, cette estimation des consommations par la température peut être satisfaisante et surtout utile dans notre contexte. On peut donc lancer une modélisation similaire pour le secteur résidentiel en trouvant d'autres données d'entraînement.

## 3.3 - Données d'entraînement

### 3.3.1 Consommation

On extrait d'abord des données de consommation bihoraires, issues du secteur résidentiel, en région PACA, dans le jeu de données d'Enedis intitulé : « *Agrégats segmentés de consommation électrique au pas 1/2 h des points de soutirage <= 36kVA - Maille Régionale* », à l'adresse suivante : <https://data.enedis.fr/explore/dataset/conso-inf36-region/information/>

Agréats segmentés de consommation électrique au pas 1/2 h des points de soutirage <= 36kVA - Maille Régionale

X  
in

Informations
Tableau
Analyse
Export
API

Horodate	Région	Code région	Profil	Plage de puissance souscrite	Nb points soutirage	Total énergie soutirée (Wh)
30 mai 2023 01:30	Provence-Alpes-Côte d'Azur	93	PRO2 (+ PRO2WE + PRO6)	P3: [6-9] kVA	14 482	3 688 102
30 mai 2023 01:30	Provence-Alpes-Côte d'Azur	93	PRO2 (+ PRO2WE + PRO6)	P7: [18-24] kVA	5 466	4 801 796
30 mai 2023 01:30	Provence-Alpes-Côte d'Azur	93	PRO2 (+ PRO2WE + PRO6)	P9: [30-36] kVA	12 186	15 786 657
30 mai 2023 01:30	Provence-Alpes-Côte d'Azur	93	PRO4	P7: [18-36] kVA	893	781 140
30 mai 2023 01:30	Provence-Alpes-Côte d'Azur	93	PRO5	P8: [24-30] kVA	89	
30 mai 2023 01:30	Provence-Alpes-Côte d'Azur	93	RES11 (+ RES11WE)	P3: [6-9] kVA	197 833	22 870 897
30 mai 2023 01:30	Provence-Alpes-Côte d'Azur	93	RES11 (+ RES11WE)	P9: [30-36] kVA	2 078	1 268 352
30 mai 2023 01:30	Provence-Alpes-Côte d'Azur	93	RES2 (+ RES5)	P4: [9-12] kVA	166 111	69 941 571
30 mai 2023 01:30	Provence-Alpes-Côte d'Azur	93	RES2WE	P0: Total <= 36 kVA	40 016	10 383 914
30 mai 2023 01:30	Provence-Alpes-Côte d'Azur	93	RES2WE	P1: [0-6] kVA	23 597	4 251 756
30 mai 2023 01:30	Provence-Alpes-Côte d'Azur	93	RES2WE	P6: [15-36] kVA	866	625 251
30 mai 2023 02:00	Provence-Alpes-Côte d'Azur	93	PRO1 (+ PRO1WE)	P0: Total <= 36 kVA	331 621	60 729 697
30 mai 2023 02:00	Provence-Alpes-Côte d'Azur	93	PRO1 (+ PRO1WE)	P2: [3-6] kVA	81 488	6 375 370
30 mai 2023 02:00	Provence-Alpes-Côte d'Azur	93	PRO2 (+ PRO2WE + PRO6)	P0: Total <= 36 kVA	86 443	45 638 905
30 mai 2023 02:00	Provence-Alpes-Côte d'Azur	93	PRO2 (+ PRO2WE + PRO6)	P5: [12-15] kVA	2 094	1 019 730
30 mai 2023 02:00	Provence-Alpes-Côte d'Azur	93	PRO3	P1: [0-9] kVA	274	

Illustration 4 : Présentation des données sur le site d'Enedis

Les filtres appliqués sont les suivants :

- Région PACA
- Profil commençant par 'RES' pour identifier le secteur Résidentiel
- Plage de puissance valant 'P0: Total <= 36 kVA'

On groupe ensuite par date en comptant la somme de la colonne `total_energie_soutiree`, que l'on convertira en MWh (en divisant les Wh par 1 million). On a donc une consommation (MWh) du secteur résidentiel pour la région PACA, à chaque date, du 1<sup>er</sup> avril 2022 au 30 juin 2024.

### 3.3.2 Température

En parallèle, on extrait des données de température quotidienne par commune dans les données internes d'AtmoSud.

Pour passer de données communales à une valeur en région PACA, on calcule simplement la moyenne des températures, sans la pondérer.

Des pondérations par la population des communes et par la consommation des communes ont été testées, et non retenues à la vue des résultats de validation du projet dans lequel s'inscrit ce modèle.

### 3.3.3 Consommation et Température

date	temperature	DJ17	DJ21	energie_soutiree_mwh
2022-04-03	2.084880	14.915120	0.0	60184.784975
2022-04-04	3.953556	13.046444	0.0	54886.855295
2022-04-05	6.622294	10.377706	0.0	48141.907184
2022-04-06	7.606582	9.393418	0.0	45816.916094
2022-04-07	8.895626	8.104374	0.0	42882.454936

Tableau 9 : Données utilisées (head)

Les données sans NA vont ainsi du 3 avril 2022 au 30 juin 2024 (avec un trou de 3 mois début 2023), et présentent pour chaque jour la consommation d'électricité en Région PACA ainsi que des données météo. Nous ajoutons une variable `WeekDay` pour informer du jour de la semaine, qui s'avère avoir une incidence sur les consommations.

## 3.4 - Modélisation

Voici la relation à modéliser :

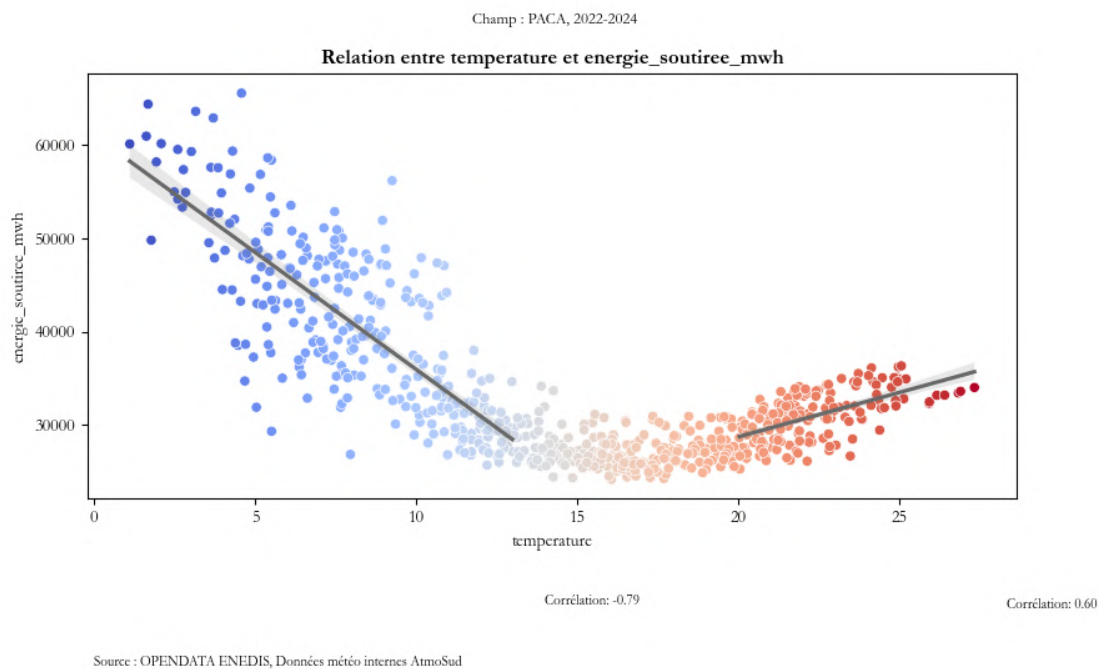


Figure 9 : Nuage de points des températures et consommations

Nous essayons différents modèles de Machine Learning pour prédire la consommation en fonction de la température. Comme pour toute méthode de Machine Learning, on sépare les données aléatoirement en 3 ensembles : un ensemble d'entraînement (70% des lignes), qui sert à ajuster un modèle  $m$ , un ensemble de validation (15% des lignes) utile pour sélectionner le meilleur modèle, et un ensemble de test (15% des lignes) qui sert enfin à évaluer la performance du modèle  $m$  sur des données non rencontrées lors de l'entraînement.

### 3.4.1 Régression polynomiale

L'idée est d'entraîner une régression polynomiale, c'est-à-dire un polynôme de la forme  $Cons_{est} = a + b1 * tempe + b2 * tempe^2 + \dots + bp * tempe^p$ ,

Lorsque l'on travaillait sur tous les secteurs confondus, il était intéressant de prendre en compte le jour de la semaine, par exemple en entraînant 7 polynômes différents. Ce n'est plus le cas dans le secteur résidentiel, où les jours ne sont pas significativement différents en terme de consommation.

Pour connaître le meilleur degré  $p$  à utiliser, on effectue une validation croisée à 5 folds sur notre ensemble d'entraînement. En d'autres termes, on coupe ces données en 5 « plis », on entraîne notre régression polynomiale de degré  $p$  sur 4 plis et on le teste sur le 5<sup>e</sup> pli pour obtenir un score d'erreur. On répète ceci 5 fois de sorte à ce que chaque pli soit l'échantillon test 1 fois et 1 seule fois. On a donc 5 scores d'erreurs, la moyenne de ces 5 scores nous donne une estimation assez robuste de l'erreur du modèle de degré  $p$ .

On fait cette validation croisée pour chaque degré  $p$  de 1 à 7 pour identifier le degré qui commet le moins d'erreurs.

On trouve  $p = 6$ .

On peut désormais tester notre modèle sur les données de validation.

RMSE (racine carrée de la moyenne des erreurs quadratiques) = 3 748.9 MWh

Visualisons le modèle ainsi qu'un nuage de points test (non rencontré par le modèle lors de l'entraînement) :

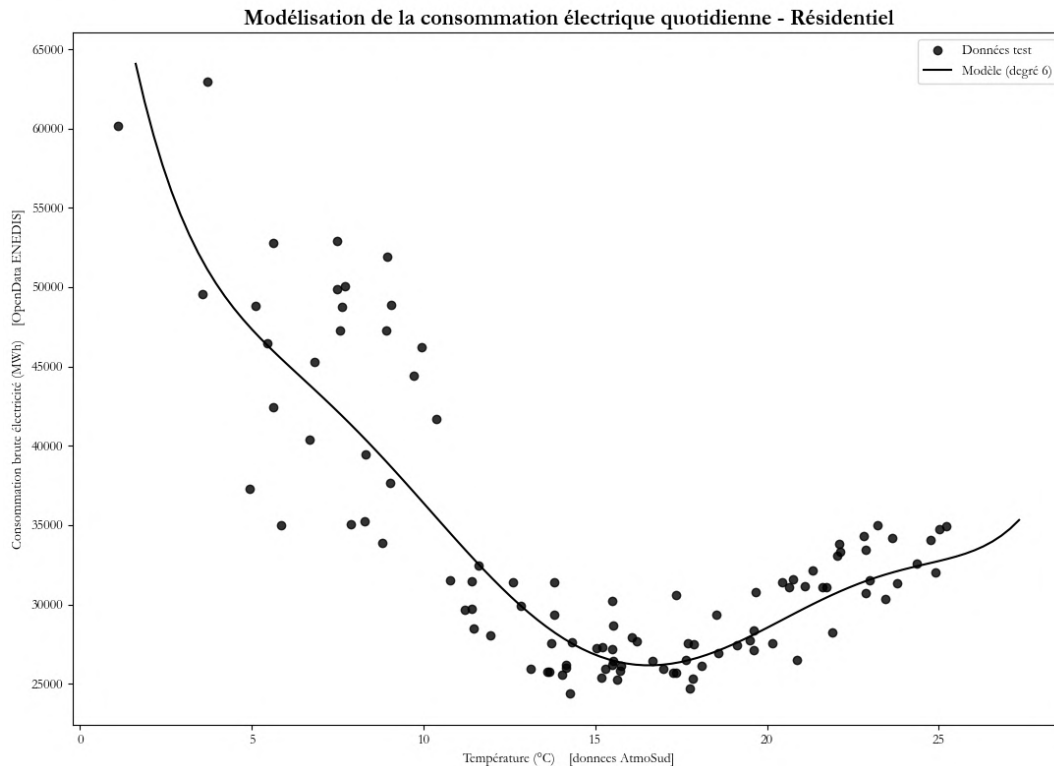


Figure 10 : Modèle de régression polynomiale

### 3.4.2 D'autres modèles ?

Essayons maintenant d'autres types de modèles de régression, pour voir s'il est possible de faire mieux :

Tree	Random Forest	SVM
4687.601721	4146.698904	8851.951821

Illustration 5 : Autres modèles, RMSE

➔ Les RMSE de ces 3 algorithmes appliqués sur l'ensemble de validation sont moins bons que notre régression polynomiale.

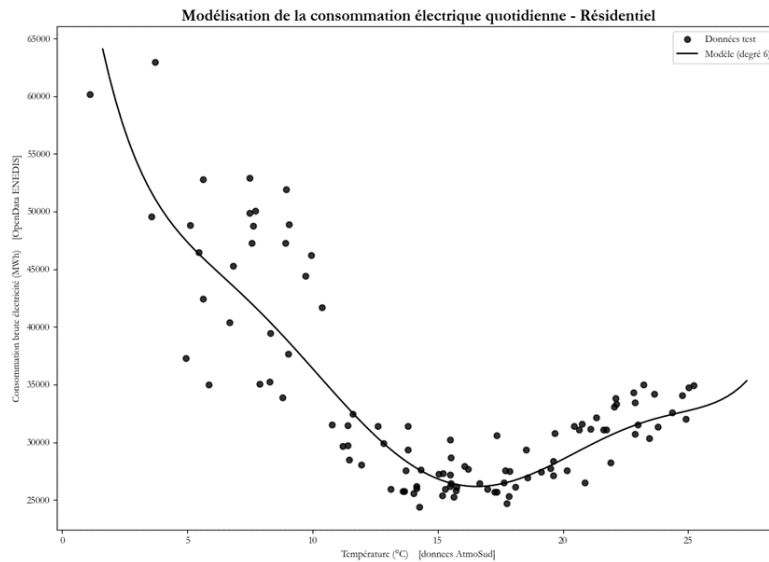
Autre avantage de la régression polynomiale, qui va nous renforcer dans l'idée d'utiliser cette méthode, elle est facilement transposable, c'est-à-dire qu'il est facile de l'exporter pour appliquer le modèle où l'on en a besoin ; il suffit d'extraire la formule polynomiale. Cela aurait été bien plus compliqué, et surtout nécessité plus de ressources, avec une « boîte noire » comme la Random Forest.

Modèle retenu à ce moment du processus :

$$\text{conso} = 87430.11 + (-20095.08277786156 * \text{tempe}^{**1}) + (4219.728678066393 * \text{tempe}^{**2}) + (-482.5300092060183 * \text{tempe}^{**3}) + (28.074054017762844 * \text{tempe}^{**4}) + (-0.7877522906343015 * \text{tempe}^{**5}) + (0.008504717233948255 * \text{tempe}^{**6})$$

### 3.4.3 Gestion des extrêmes (Edge-Case)

On a donc un modèle représenté comme vu précédemment :



Interrogation : que se passe-t-il avec ce polynôme si on lui donne en input une température extrême non rencontrée lors de l'entraînement ? C'est une chose qui peut arriver, d'autant plus avec le changement climatique.

- ➔ Le polynôme est entraîné pour s'ajuster aux données seulement sur l'intervalle considéré lors de l'entraînement. Au-delà, il est susceptible de diverger.

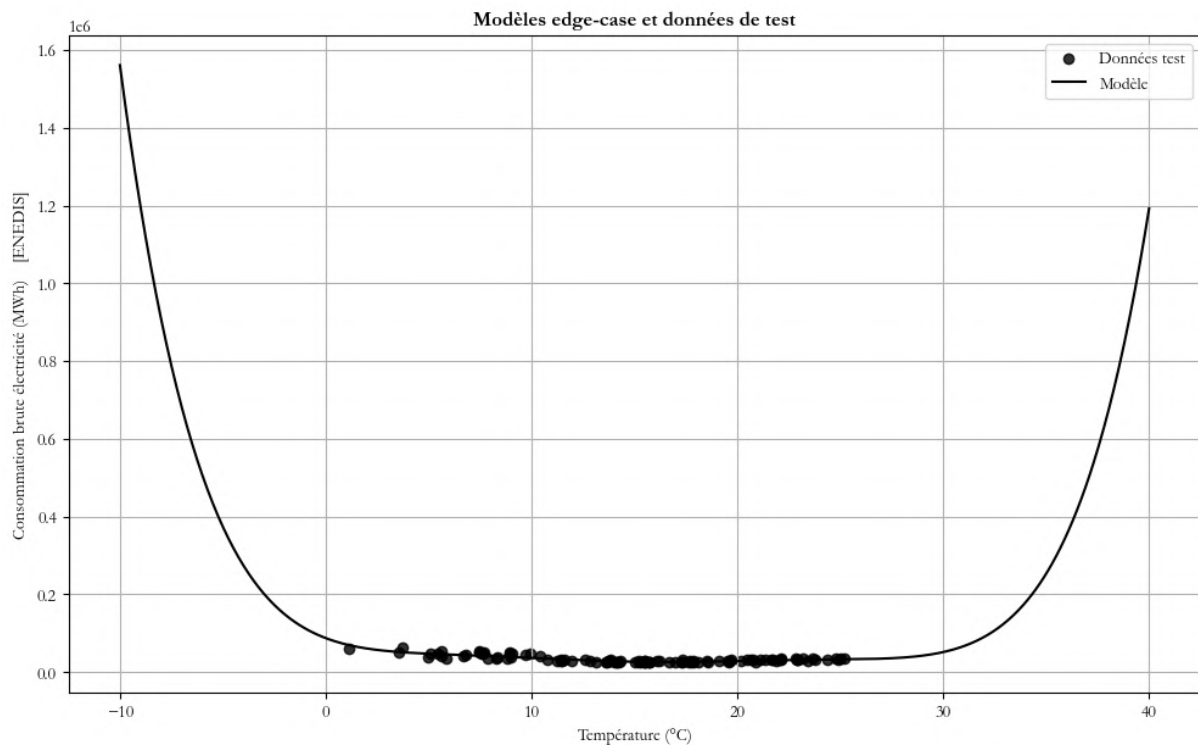


Figure 11 : Régression polynomiale aux valeurs extrêmes

On souhaite alors contrôler ces comportements en affectant une fonction (ou modèle) moins instable aux valeurs de température extrêmes.

Pour les valeurs extrêmes (disons  $< 3^{\circ}\text{C}$  ou  $> 27^{\circ}\text{C}$ ), on va remplacer le polynôme par des régressions linéaires, une pour le froid et une pour le chaud, afin d'avoir un modèle contrôlé comme ceci :



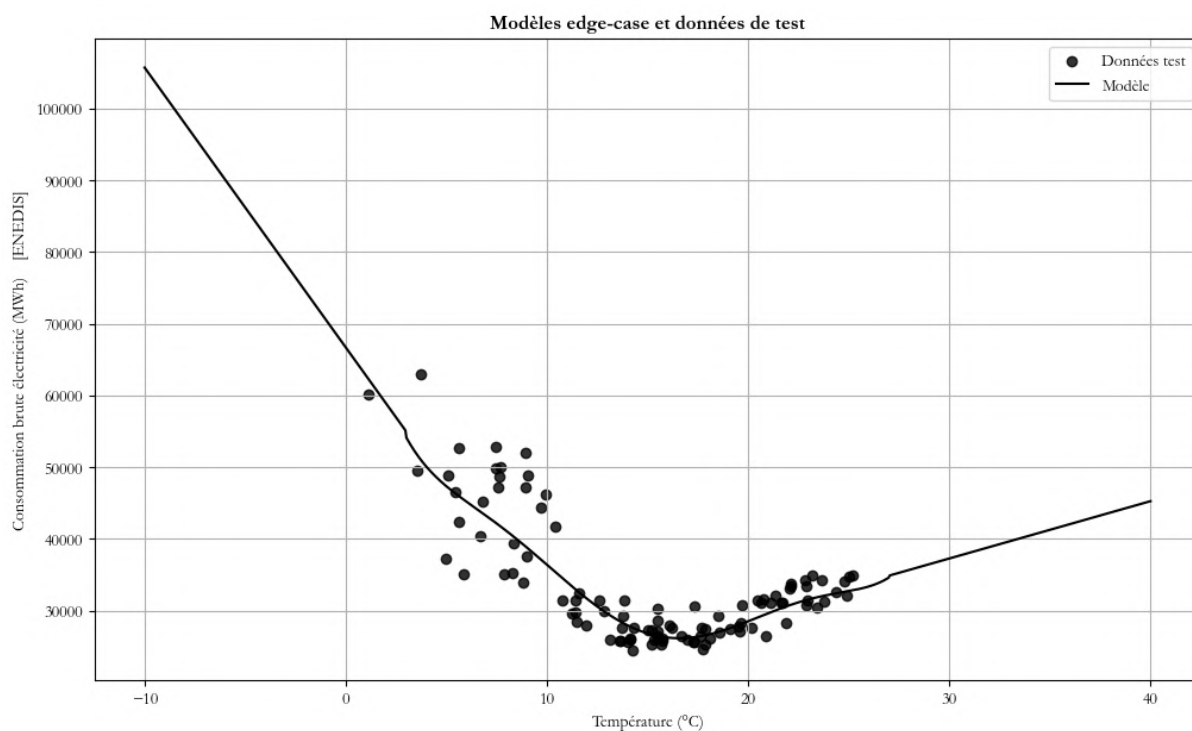


Figure 12 : Représentation du modèle général avec valeurs extrêmes contrôlées

L'entraînement des 2 régressions linéaires se fait en filtrant les données : on choisit un seuil (6°C) en-dessous duquel les données sont considérées pour entraîner la régression du froid extrême, et un seuil (22°C) au-dessus duquel les données sont considérées pour entraîner la régression du chaud extrême. Ces seuils sont déterminés par tâtonnement et visualisation dans le but d'avoir des pentes cohérentes. L'absence d'une méthodologie rigoureuse pour le choix de ces seuils ne pose pas de problème majeur, sachant qu'il s'agit ici de contrôler des cas extrêmes rares.

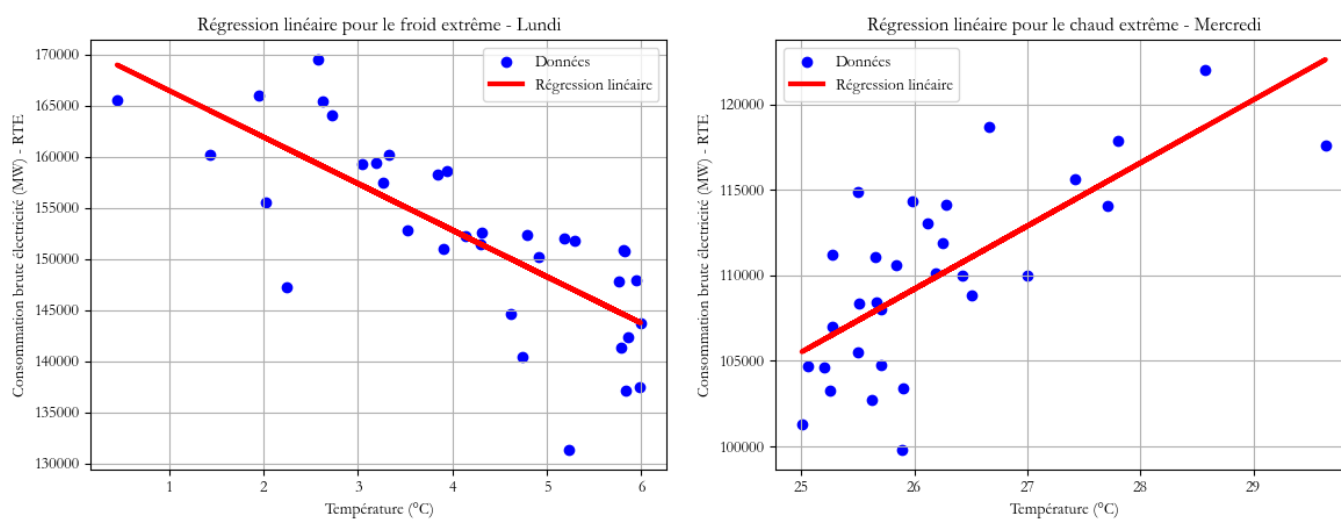


Figure 13 : Régressions linéaires pour les valeurs extrêmes

### 3.5 - Export du modèle pour utilisation

A ce moment du processus, le modèle d'estimation de la consommation électrique en région PACA en fonction de la température est prêt. Il est simple à exporter, avec les formules suivantes :

**Lundi**

- SI température < 1°C :  $ConsoPaca = 66704.24 + (-3902.89 * \text{température})$
- SI  $1^\circ\text{C} \leq \text{température} < 29^\circ\text{C}$  :  $ConsoPaca = 87430.11 + (-20095.0827 * \text{température}) + (4219.7286 * \text{température}^2) + (-482.530009 * \text{température}^3) + (28.07405 * \text{température}^4) + (-0.78775 * \text{température}^5) + (0.0085047 * \text{température}^6)$
- SI température  $\geq 29^\circ\text{C}$  :  $ConsoPaca = 13295.94 + (799.40 * \text{température})$

Pour passer de cette estimation de la consommation dans tout PACA en fonction de la température moyenne en PACA à une estimation de la consommation d'une commune en fonction de la température de cette commune, on a :

$ConsoCommune(\text{tempe}) = ConsoPACA(\text{tempe}) * \text{part\_conso}$

où  $\text{part\_conso} = \frac{ConsoCommune_{2021}}{ConsoPACA_{2021}}$  est la part de consommation de la commune dans la région PACA en 2021, issue de l'inventaire d'AtmoSud.

### 3.6 - Test du modèle

Tout est prêt pour déploiement, on peut donc estimer les consommations au jour J d'une commune dès lors que l'on a accès à une donnée de température.

Visualisons la qualité de prédiction du modèle.

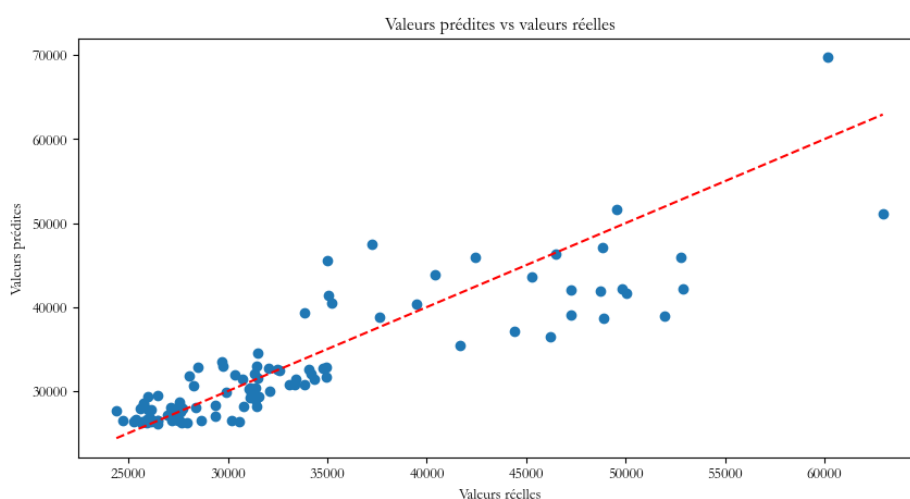
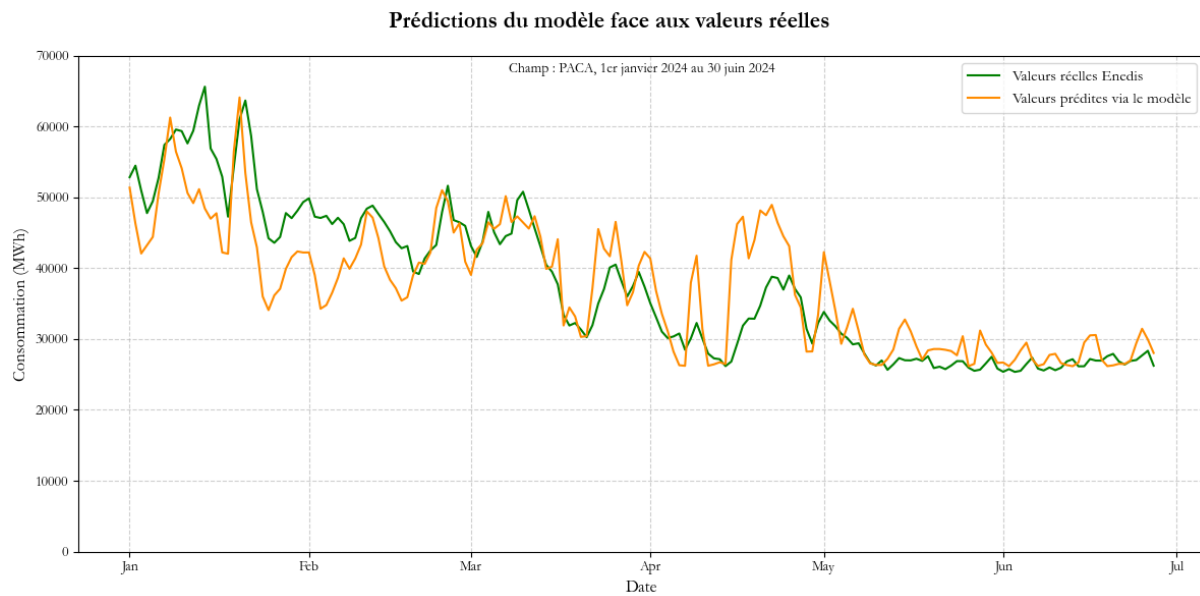


Figure 14 : Valeurs prédites VS Valeurs réelles (données test)

Ensuite, comparons l'évolution 2024 des consommations quotidiennes estimées avec celle des données réelles d'Enedis, sur lesquelles a été, entre autres, entraîné le modèle. Ceci permet de voir si on suit correctement les tendances.



*Figure 15 : Valeurs estimées VS valeurs réelles en 2024*

Enfin, en chiffres, on retrouve sur ces données test

- **une erreur RMSE** (Root Mean Squared Error) de 4103 MWh, soit **12.3%** du Root Mean Square des consommations.
- **Une erreur MAE** (Mean Absolute Error) de 2825 MWh, soit **8.5%** de la consommation moyenne.

➔ Bien qu'évidemment limitées, ces performances peuvent être considérées **satisfaisantes** étant donné :

- La simplicité du modèle (transposable, explicable, rapide à entraîner)
- La difficulté à trouver des données de consommation journalières du résidentiel

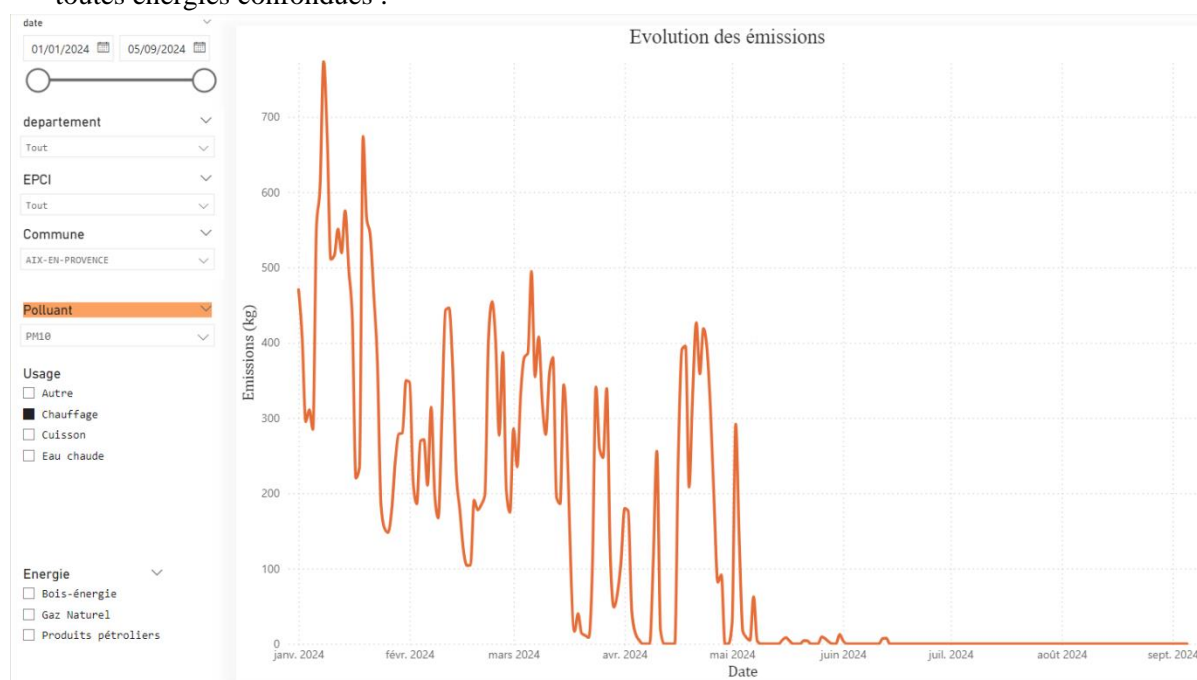
## IV - Méthode de calcul des émissions en temps réel du secteur Résidentiel

### 4.1 - Objectif

On vise à estimer les émissions issues du Résidentiel des différents polluants suivis

- « en temps réel », c'est-à-dire chaque jour
- Dans chaque commune
- Par usage (Cuisson, Lavage, Chauffage, ECS etc..)
- Par Energie (Electricité, Bois, Gaz etc..)

Exemple avec les émissions de PM10 à Aix-en-Provence (13001) issues de l'usage Chauffage, toutes énergies confondues :

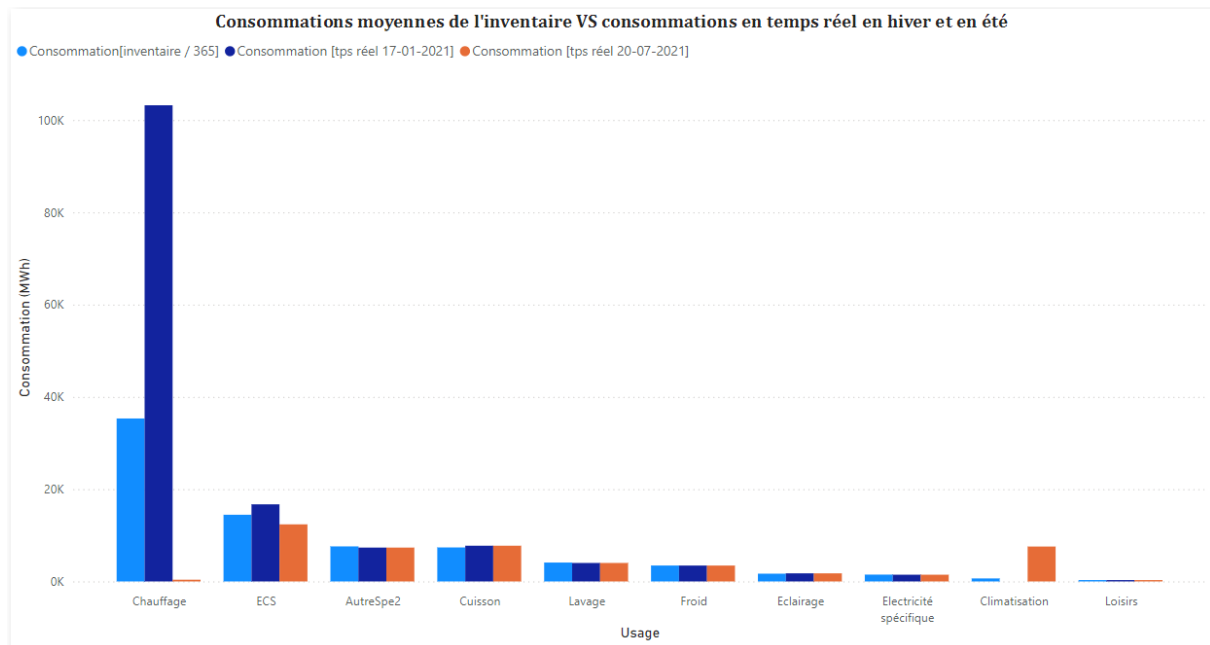


*Illustration 6 : Exemple de résultat final*

Pour cela, comme les **émissions** du secteur Résidentiel sont en lien direct avec les **consommations**, l'on s'attache d'abord à identifier ces consommations du jour dans chaque commune, pour chaque usage et chaque énergie. Il ne restera alors plus qu'à appliquer les facteurs d'émissions de polluants correspondants.

### 4.2 - Problématique

Tout le problème de cette notion de « temps réel » est de savoir comment prendre en compte les spécificités journalières. Ne pas les prendre en compte reviendrait à prendre directement les émissions annuelles de l'inventaire des émissions et de les diviser par 365 pour avoir les consommations d'un jour « moyen », c'est ce qui est en bleu clair ci-dessous.



*Figure 16 : Consommations moyenne, hiver, et été*

L'idée est donc au contraire de prendre en compte ces spécificités, de façon à ce qu'on retrouve plus de Chauffage un jour d'hiver (bleu foncé) et plus de Climatisation un jour d'été (orange), par exemple grâce à la température du jour.

La méthode d'adaptation des consommations par usage et énergie selon le jour (donc la température) a ici été revue par rapport à la partie II, il s'agit ici d'une 2<sup>e</sup> méthode, que l'on trouve moins contraignante, moins complexe, et fournissant de meilleurs résultats en validation.

## 4.3 - Présentation générale de la méthode

### CALCUL

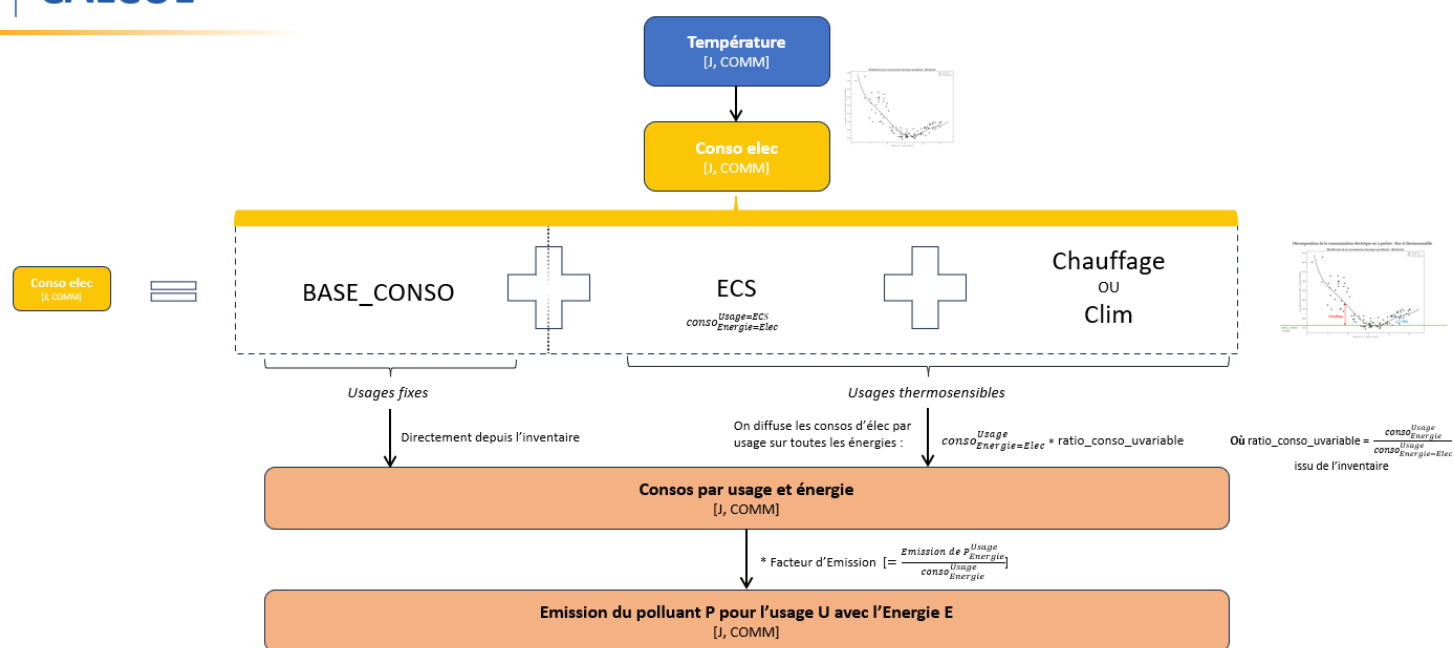
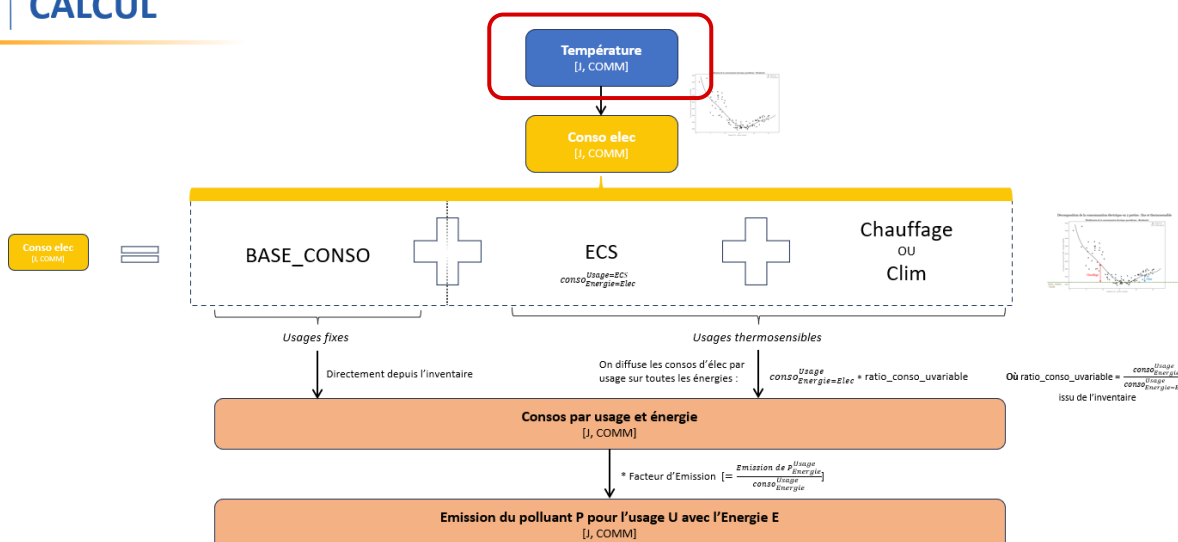


Illustration 7 : Méthode de calcul

## 4.4 - Extraction des températures communales journalières

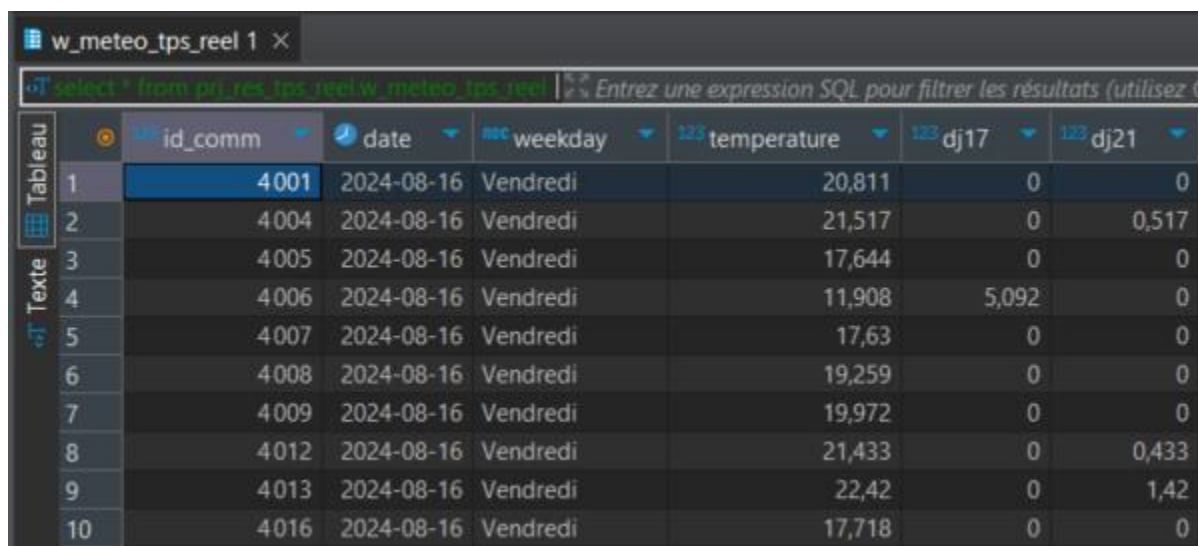
### CALCUL



La température du jour, pour chaque commune, est le fondement de la méthode utilisée ici, puisqu'elle nous permet à la fois d'estimer la consommation totale et d'identifier les consommations liées aux usages thermosensibles comme le Chauffage ou la Climatisation.

La 1<sup>re</sup> étape est donc d'extraire proprement des données de température en temps réel à l'échelle communale. Ces données sont disponibles au pas de temps horaire, avec une prévision à J+3, sur le réseau interne d'AtmoSud.

Il ne reste plus qu'à les regrouper par jour, en faisant la moyenne des 24 températures horaires dans chaque commune, ainsi qu'à prendre la valeur de la colonne 'echeance' la plus petite parmi les lignes ayant la même combinaison [commune, date] afin d'avoir la prévision la plus récente, pour obtenir notre table :

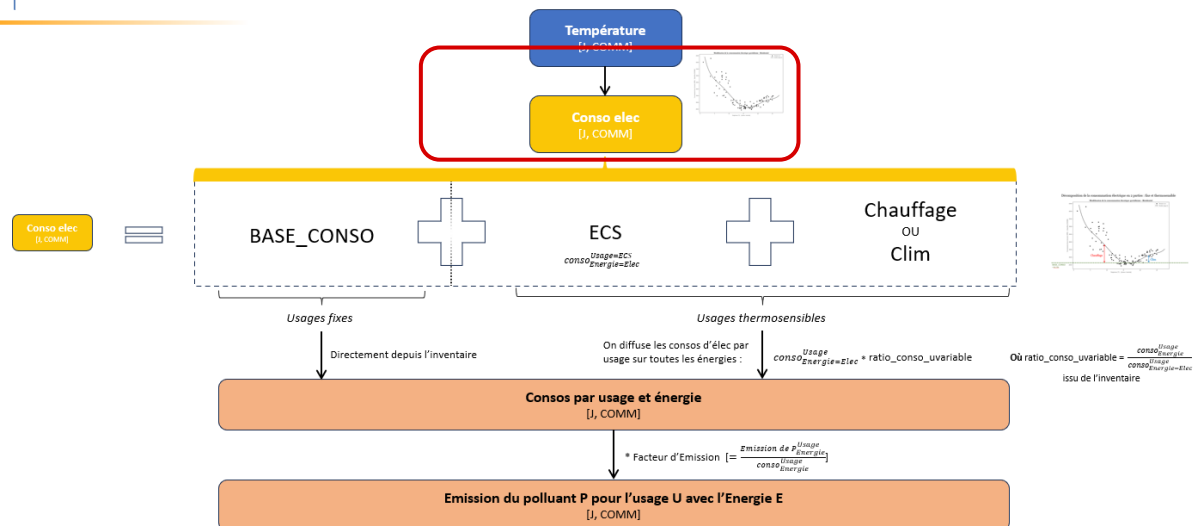


	id_comm	date	weekday	temperature	dj17	dj21
1	4001	2024-08-16	Vendredi	20,811	0	0
2	4004	2024-08-16	Vendredi	21,517	0	0,517
3	4005	2024-08-16	Vendredi	17,644	0	0
4	4006	2024-08-16	Vendredi	11,908	5,092	0
5	4007	2024-08-16	Vendredi	17,63	0	0
6	4008	2024-08-16	Vendredi	19,259	0	0
7	4009	2024-08-16	Vendredi	19,972	0	0
8	4012	2024-08-16	Vendredi	21,433	0	0,433
9	4013	2024-08-16	Vendredi	22,42	0	1,42
10	4016	2024-08-16	Vendredi	17,718	0	0

Tableau 10 : Relevés météo (head)

## 4.5 - Consommation électrique à partir de la température

### CALCUL



Le modèle peut être représenté face à des données test non rencontrées lors de l'entraînement ainsi :

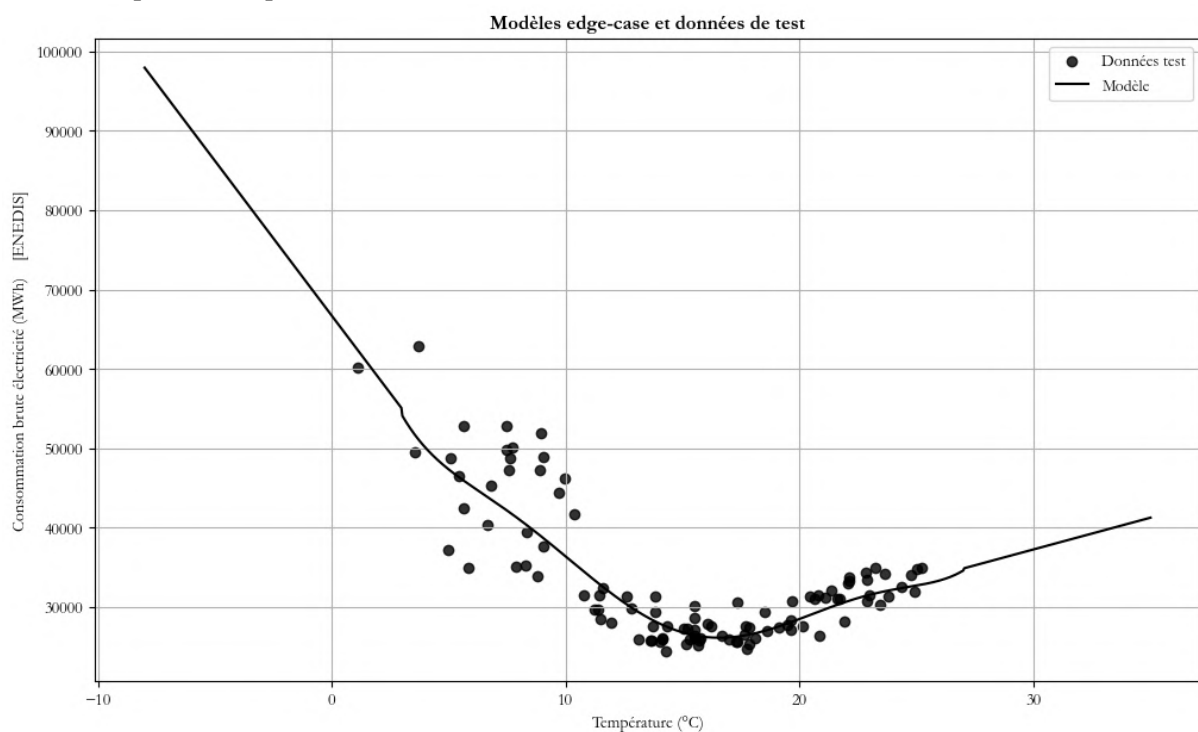


Figure 17 : Modélisation de la consommation (PACA) en fonction de la température

Tout le détail de cette modélisation se trouve dans la partie III.

On obtient ainsi une consommation électrique du Résidentiel chaque jour dans chaque commune.

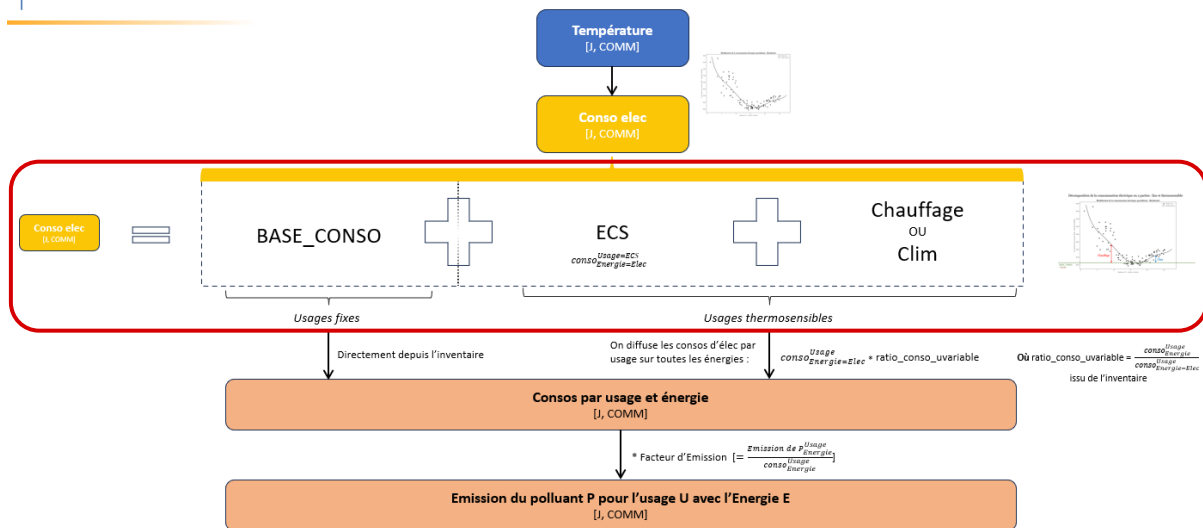


w_conso_res_estim_tr 1						
select * from pti_res_tps_real w_conso_res_estim_tr   %s %s Entrez une expression SQL pour filtrer les résultats (utilisez Ctrl+Espace)						
	date	123 id_comm	123 temperature	123 part_conso_comm	123 consommation_est_mwh	
1	2024-08-16	4001	20,811	0,0002633246	7,7519257461	
2	2024-08-16	4004	21,517	0,0001550417	4,6840346192	
3	2024-08-16	4005	17,644	0,0000264817	0,7002470717	
4	2024-08-16	4006	11,908	0,0005304745	16,8916895835	
5	2024-08-16	4007	17,63	0,0000142378	0,3763851935	
6	2024-08-16	4008	19,259	0,0002770984	7,6830816328	
7	2024-08-16	4009	19,972	0,0000038778	0,1104848262	
8	2024-08-16	4012	21,433	0,0000250426	0,7543451112	
9	2024-08-16	4013	22,42	0,0001270327	3,9497155455	
10	2024-08-16	4016	17,718	0,0000115899	0,3069204502	
11	2024-08-16	4017	17,159	0,0000186407	0,4892562462	
12	2024-08-16	4018	21,745	0,0002297752	6,9961401836	
13	2024-08-16	4019	16,56	0,0007687195	20,1088010299	

Tableau 11 : Estimations de consommation électrique (head)

## 4.6 - Décomposition de la consommation électrique

### CALCUL



Une fois que l'on a une consommation électrique totale, on souhaite la décomposer pour identifier les parts de consommation liées aux usages thermosensibles. On considèrera 3 usages comme thermosensibles : Chauffage, Climatisation, ECS (Eau Chaude Sanitaire). Pour cette décomposition, on réutilise le modèle d'estimation vu précédemment.

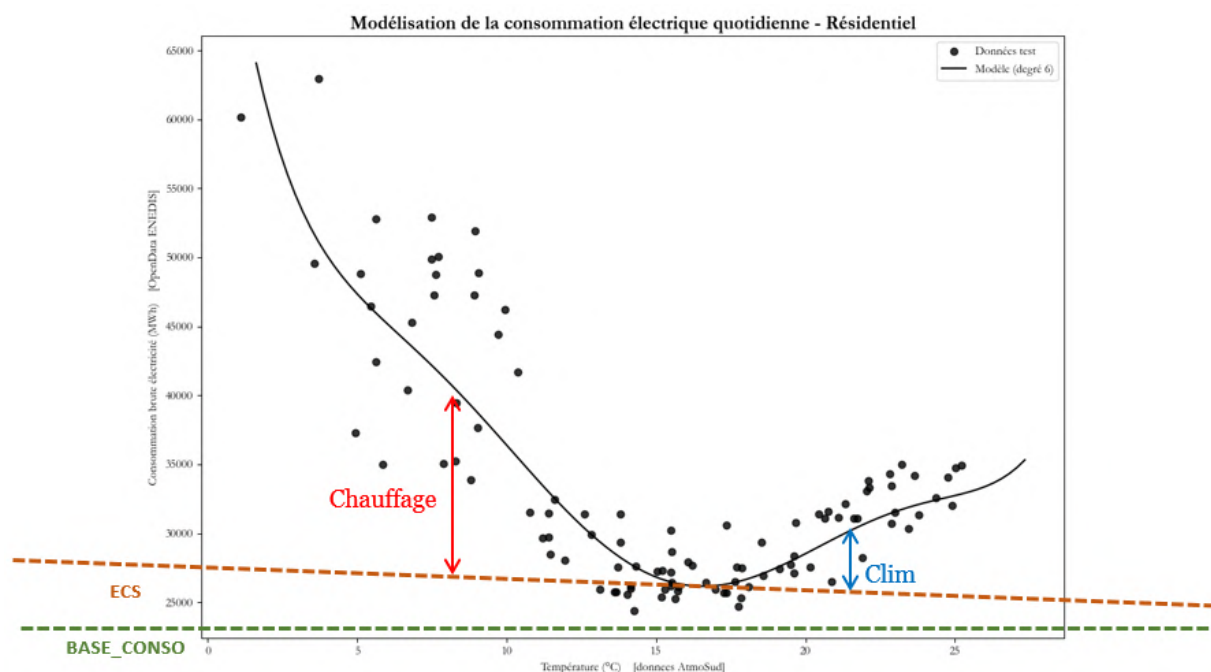


Illustration 8 : Décomposition de la consommation électrique

### 4.6.1 Détermination de BASE\_CONSO

« BASE\_CONSO » représente la consommation électrique journalière non liée aux usages thermosensibles (Chauffage, Clim, ECS). Elle sera donc considérée fixe et unique pour une commune, restant la même tout au long de l'année.

Pour l'identifier s'assurer de sa pertinence, on procède ainsi :

$$BASE_{CONSO} = \min(Conso_{totale\ estimee}) - ECS_{fixe}$$

- Où  $\min(Conso_{totale\ estimee})$  est le minimum du modèle d'estimation de consommation de la commune, soit le point le plus bas de la courbe noire
- Où  $ECS_{fixe}$  est la consommation électrique moyenne journalière pour l'ECS dans la commune, issue de l'inventaire AtmoSud. Autrement dit c'est la consommation d'ECS avant correction par la température (voir section 4.6.2)

Tout l'intérêt de cette valeur réside dans l'identification des consommations de Chauffage ou Climatisation :

$$Conso_{Chauffage\ ou\ Clim} = Conso_{totale\ estimee} - ECS - BASE\_CONSO$$

D'autres méthodes pour déterminer BASE\_CONSO ont été testées, par exemple en utilisant l'inventaire AtmoSud pour avoir des consommations électriques moyenne hors usages thermosensibles. Cette méthode a fourni des résultats médiocres en validation car les estimations de consommation étaient trop basses par rapport à cette BASE\_CONSO, impliquant l'omission de beaucoup de chauffage et climatisation.

### 4.6.2 Correction de l'ECS

Avant correction, on a  $ECS_{fixe}$  ; la consommation électrique moyenne journalière pour l'ECS dans la commune, issue de l'inventaire AtmoSud. La correction de cette consommation électrique liée à l'ECS se fait en considérant que + il fait chaud, moins on consomme d'eau chaude sanitaire, et ce de façon linéaire.

$$ECS = ECS_{fixe} * (1 + \frac{temperature_{moy} - temperature}{\max(|temperature - temperature_{moy}|)} * 0.28)$$

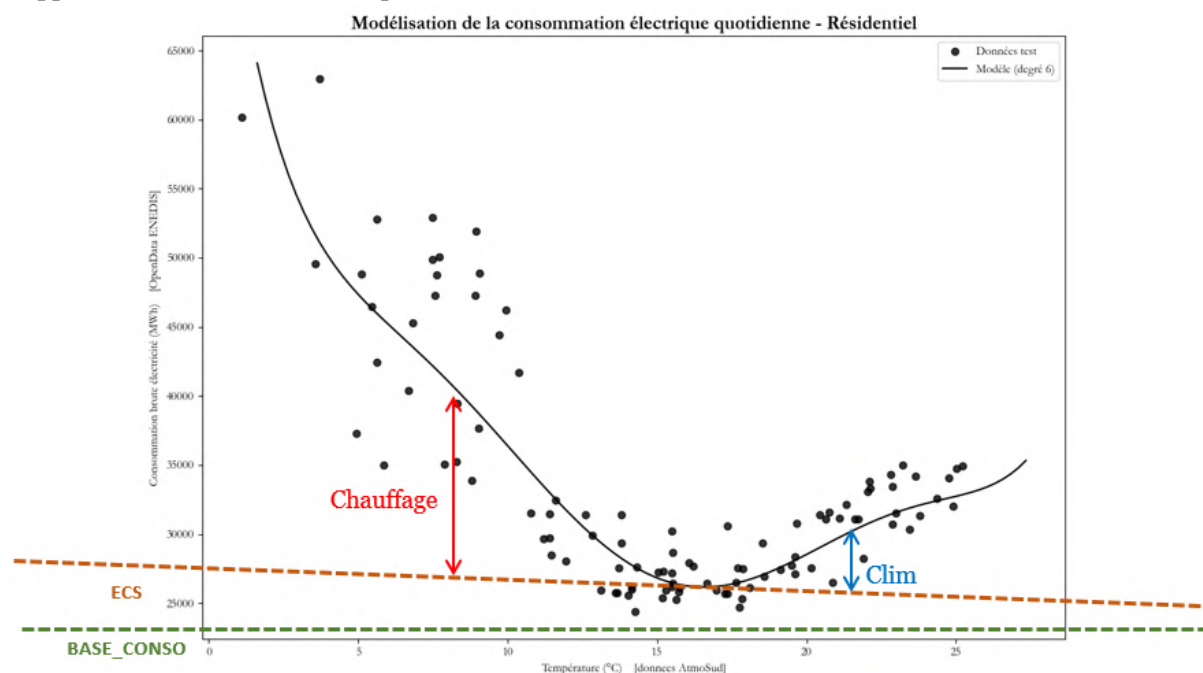
- Où  $temperature$  est la température du jour dans la commune
- $temperature_{moy}$  est la température moyenne de la commune de 2021 à 2023 inclus
- $\max(|temperature - temperature_{moy}|)$  est l'écart max de température observé par rapport à la moyenne. Ce jour-là, l'ECS sera augmenté, ou diminué, de 28%. D'où \* 0.28

Ainsi, on encadre la correction de l'usage ECS en s'assurant qu'il ne sera jamais corrigé de plus que de 28%, le terme dans la parenthèse étant toujours compris entre 0.72 et 1.28.

Cette limite de 28% provient d'une étude sur la consommation des chauffe-eau. (source en références)

### 4.6.3 Identification du chauffage et de la climatisation

Rappel de la méthode de décomposition :



Une fois que l'on a fixé BASE\_CONSO et estimé une consommation électrique corrigée pour l'ECS, il ne reste plus qu'à identifier ce qu'il manque à notre consommation d'électricité totale et à déterminer si le solde est du Chauffage ou de la Climatisation. Deux cas de figure :

- Temperature  $< 17^{\circ}\text{C}$   $\rightarrow$  Chauffage
- Temperature  $> 21^{\circ}\text{C}$   $\rightarrow$  Climatisation
- Température entre 17 et 21  $\rightarrow$  Ni chauffage ni climatisation

Ces seuils proviennent du Service des données et études statistiques (SDES).

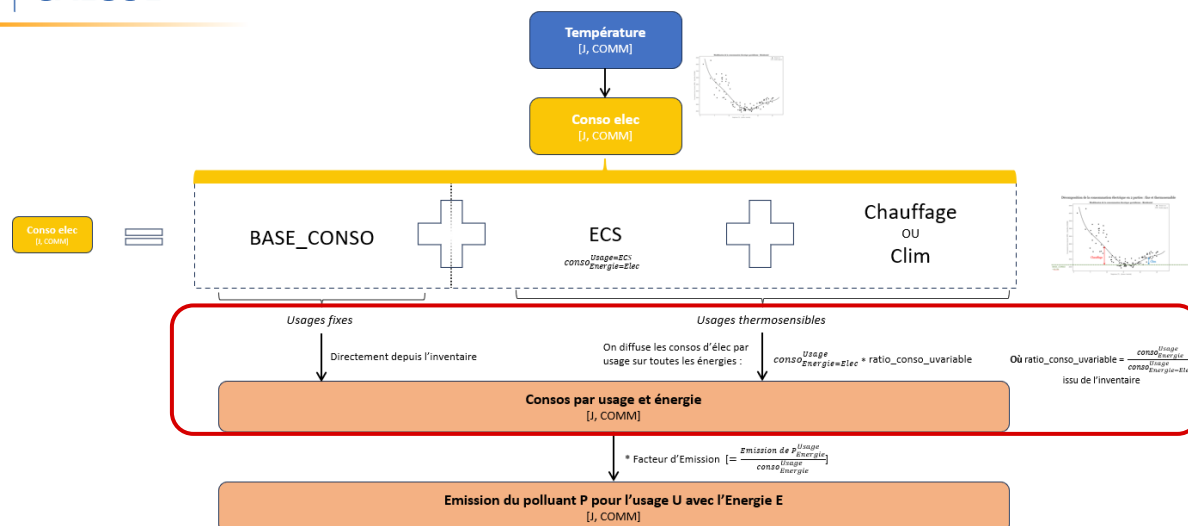
Nous voilà enfin avec une consommation électrique bien décomposée, et surtout avec des consommations électriques adaptées à la température du jour pour nos usages thermosensibles.

	date	id_comm	base_consoelec_mwh	temperature	consommation_est_mwh	conso_elec_ecs_fixe_mwh	conso_elec_ecs_mwh	conso_elec_chauffage_mwh	conso_elec_clim_mwh
1	2024-08-16	4001	4.7515478665	20.811	7.7519257461	2.1367140978	1.7756652612	0	0
2	2024-08-16	4004	2.5981027434	21.517	4.6840346192	1.4576038142	1.2129486713	0	0.8729832045
3	2024-08-16	4005	0.4524729527	17.644	0.7002470717	0.2404189805	0.19830063	0	0
4	2024-08-16	4006	8.9214991084	11.908	16.8916895835	4.9558813877	4.1451473526	3.8250431225	0
5	2024-08-16	4007	0.2750333866	17.63	0.3763851935	0.097441331	0.0804694125	0	0
6	2024-08-16	4008	4.7675697804	19.259	7.6830816328	2.4824077428	2.0381584482	0	0
7	2024-08-16	4009	0.0676895978	19.972	0.1104848262	0.0337498903	0.0274983567	0	0
8	2024-08-16	4012	0.4247882171	21.433	0.7543451112	0.2302977141	0.1888788582	0	0.1406780359
9	2024-08-16	4013	2.2040202017	22.42	3.9497155455	1.1190067215	0.9123517385	0	0.8333436052
10	2024-08-16	4016	0.2111521979	17.718	0.3069204502	0.0920374828	0.0744044563	0	0
11	2024-08-16	4017	0.3051726094	17.159	0.69961401836	0.1824507021	0.1494636965	0	0
12	2024-08-16	4018	3.8787117526	21.745	6.9961401836	2.1319755785	1.7392762323	0	1.3781521987
13	2024-08-16	4019	13.1468338946	16.56	20.1088010299	6.0619671353	5.7275800014	1.2343691339	0
14	2024-08-16	4020	0.4231507683	16.766	0.6260144952	0.2026447225	0.1659187311	0.0369449959	0
15	2024-08-16	4021	0.4114765314	20.331	0.7041988485	0.2222654351	0.1836578911	0	0

Tableau 12 : Consommations électriques séparées par usages thermosensibles (head)

## 4.7 - Consommations par usage et énergie

### CALCUL



Désormais, on cherche à retrouver des consommations par usage et par énergie, là où jusqu'à maintenant on ne considérait que l'électricité en termes d'énergie.

### 4.7.1 Pour les usages fixes

Nul besoin d'utiliser BASE\_CONSO, on prend directement dans l'inventaire AtmoSud les consommations moyennes journalières par usage et par énergie. Ces moyennes sont calculées sur la base des inventaires AtmoSud de 2018 à 2021 inclus.

### 4.7.2 Pour les usages thermosensibles

On diffuse les consommations électriques par usage sur toutes les autres énergies en appliquant un ratio de consommation, calculé à partir des données d'inventaire annuel.

$$conso_{Energie}^{Usage} = conso_{Energie=Elec}^{Usage} * ratio\_conso\_variable$$

$$Où \text{ ratio\_conso\_variable} = \frac{conso_{Energie}^{Usage}[\text{inventaire 2021}]}{conso_{Energie=Elec}^{Usage}[\text{inventaire 2021}]}$$

w2_ratios_conso_variables 1 X									
select * from rat_conso_variables w2_ratios_conso_variables 1 X Entrez une expression SQL pour filtrer les résultats (utilisez Ctrl+Espace)									
	id_comm	id_usage	usage_old	code_cat_energie	energie_old	consommation	id_unite	consommation_elec	ratio_conso_variable
1	4001	0	Chauffage	1	Gaz Naturel	557,9381582026	-30007	1282.3296365715	0,4350972966
2	4001	0	Chauffage	2	Produits pétroliers	1443,2091297379	-30007	1282.3296365715	1,1254587655
3	4001	0	Chauffage	4	Bois-énergie	2395,7802942806	-30007	1282.3296365715	1,8683029901
4	4001	0	Chauffage	8	Electricité	1282.3296365715	-30007	1282.3296365715	1
5	4001	1	ECS	1	Gaz Naturel	166,4293055854	-30007	778.3085425112	0,2138346125
6	4001	1	ECS	2	Produits pétroliers	22,342652534	-30007	778.3085425112	0,0287066778
7	4001	1	ECS	8	Electricité	778.3085425112	-30007	778.3085425112	1
8	4001	2	Cuisson	1	Gaz Naturel	106,2767361021	-30007	402.7626460051	0,2638693959
9	4001	2	Cuisson	2	Produits pétroliers	2,4603516183	-30007	402.7626460051	0,0061086887
10	4001	2	Cuisson	8	Electricité	402.7626460051	-30007	402.7626460051	1
11	4001	3	Electricité spécifique	8	Electricité	23,7506009036	-30007	23.7506009036	1
12	4001	11	Froid	8	Electricité	302,7330923412	-30007	302.7330923412	1
13	4001	12	Lavage	8	Electricité	363,957406833	-30007	363,957406833	1
14	4001	33	Eclairage	8	Electricité	145,5018707497	-30007	145,5018707497	1
15	4001	34	AutreSpe2	8	Electricité	691,1918078501	-30007	691,1918078501	1
16	4001	35	Loisirs	2	Produits pétroliers	81,7685915165	-30007	0	[NULL]
17	4001	35	Loisirs	8	Electricité	0	-30007	0	[NULL]

Tableau 13 : ratios de consommation à utiliser pour les usages thermosensibles (head)

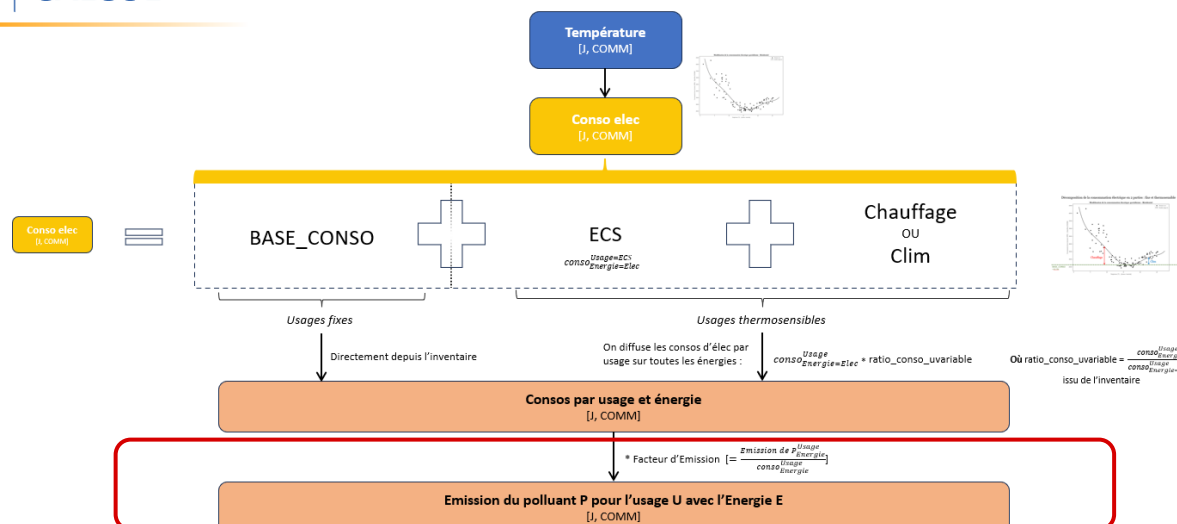
Et nous voilà avec des consommations par usage et par énergie adaptées à la température du jour, pour chaque commune.

w2_consos_u_e_estim_tr 1 X											
Entrez une expression SQL pour filtrer les résultats (utilisez Ctrl+Espace)											
	date	id_comm	base_consoelec_mwh	temperature	conso_elec_chauffage_mwh	conso_elec_clim_mwh	id_usage	usage_old	code_cat_energie	ratio_conso_variable	conso_u_e_mwh
1	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	0	Chauffage	8	1	0
2	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	0	Chauffage	2	0,8099532089	0
3	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	0	Chauffage	1	1,7364664138	0
4	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	0	Chauffage	4	1,6607292399	0
5	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	0	Chauffage	6	0,4514460815	0
6	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	1	ECS	1	0,4844191397	100,0937129346
7	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	1	ECS	2	0,0160050361	3,30700606707
8	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	1	ECS	8	1	206,6262555027
9	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	2	Cuisson	1	0,7897215446	103,4637921195
10	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	2	Cuisson	2	0,0038560665	0,4489136433
11	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	2	Cuisson	8	1	107,2383788368
12	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	3	Electricité spécifique	8	1	11,9414303309
13	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	11	Froid	8	1	87,8684541943
14	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	12	Lavage	8	1	101,1181904442
15	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	21	Climatisation	8	1	231,6370529905
16	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	33	Eclairage	8	1	45,3188837615
17	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	34	AutreSpe2	8	1	193,8243681543
18	2024-08-16	13001	377,9288597349	26,11	0	231,6370529905	35	Loisirs	2	[NULL]	1,9376854841

Tableau 14 : Consommations par usage et énergie (head)

## 4.8 - Emissions de polluants par usage et par énergie

### CALCUL



Une fois que l'on a nos consommations par usage et par énergie adaptées à la température du jour dans la commune concernée, il ne reste plus qu'à appliquer les facteurs d'émission de **chaque polluant** à ces consommations pour avoir les émissions correspondantes. Le lien direct entre consommations et émissions s'exprime ici.

Plus concrètement, qu'est-ce qu'un facteur d'émission ? C'est la quantité d'un polluant émise dans l'air pour une activité de consommation donnée. Par exemple, lorsqu'on consomme 1 MWh de Bois-Energie pour l'usage Chauffage, on émet environ 0.09kg de Benzène.

On utilise donc des facteurs d'émissions induits par polluant, par usage et par énergie, par commune, comme présentées ci-dessous. Ces facteurs d'émission par polluant sont recalculés pour chaque commune par usage et par énergie, à partir des données d'inventaire annuel.

w_facteurs_emiss_res 1									
Entrez une expression SQL pour filtrer les résultats (utilisez Ctrl+Espace)									
	**id_comm	**id_polluant	**nom_court_polluant	**id_usage	**code_cat_energie	**emission_kg	**consommation_mwh	**facteur_emission_kg_by_mwh	
1	4001	2	As	0	1	0,0000239333	557,9381582026	0,0000000429	
2	4001	2	As	0	2	0,0000095884	1443,2091297379	0,0000000066	
3	4001	2	As	0	4	0,0467759517	2395,7802942806	0,0000195243	
4	4001	2	As	1	1	0,0000071392	166,4293055854	0,0000000429	
5	4001	2	As	2	1	0,0000045588	106,2767361021	0,0000000429	
6	4001	2	As	35	2	0,0000020605	81,7685915165	0,0000000252	
7	4001	3	Benzene	0	1	0,3286371485	557,9381582026	0,0005890207	
8	4001	3	Benzene	0	2	0,2234974288	1443,2091297379	0,0001548614	
9	4001	3	Benzene	0	4	223,1819059104	2395,7802942806	0,0931562491	
10	4001	3	Benzene	1	1	0,0980303132	166,4293055854	0,0005890207	
11	4001	3	Benzene	1	2	0,0130299859	22,342652534	0,0005831889	
12	4001	3	Benzene	2	1	0,0625992021	106,2767361021	0,0005890207	
13	4001	3	Benzene	2	2	0,0014348496	2,4603516183	0,0005831889	
14	4001	3	Benzene	35	2	20,666418329	81,7685915165	0,2527427457	
15	4001	4	BenzoAPyren	0	1	0,0000012051	557,9381582026	0,0000000022	
16	4001	4	BenzoAPyren	0	2	0,0000630814	1443,2091297379	0,0000000437	

Tableau 15 : Facteurs d'émission (head)

La multiplication de nos consommations (par jour et commune, par usage et énergie) par les facteurs d'émission correspondants nous donne l'estimation finale de nos émissions par polluant, usage, énergie, dans telle commune pour un jour donné.



emissions_estim_tr 1											
Entrez une expression SQL pour filtrer les résultats (utilisez Ctrl+Espace)											
	date	id_comm	base_consoulelec_mwh	temperature	id_polluant	id_usage	code_cat_energie	ratio_conso_variable	conso_u_e_mwh	facteur_emission_kg_by_mwh	emission_tps_reel_kg
1	2024-08-16	4001	4.7515478665	20.811	2	0	1	0.4350972966	0	0.0000000429	0
2	2024-08-16	4001	4.7515478665	20.811	2	0	2	1.1254587655	0	0.0000000066	0
3	2024-08-16	4001	4.7515478665	20.811	2	0	4	1.8683029901	0	0.0000195343	0
4	2024-08-16	4001	4.7515478665	20.811	2	1	1	0.2138346125	0.379698693	0.0000000429	0.0000000163
5	2024-08-16	4001	4.7515478665	20.811	2	2	1	0.2638693959	0.2922098114	0.0000000429	0.0000000125
6	2024-08-16	4001	4.7515478665	20.811	2	35	2	[NULL]	0.221106073	0.0000000252	0.0000000056
7	2024-08-16	4001	4.7515478665	20.811	3	0	1	0.4350972966	0	0.0005890207	0
8	2024-08-16	4001	4.7515478665	20.811	3	0	2	1.1254587655	0	0.0001548614	0
9	2024-08-16	4001	4.7515478665	20.811	3	0	4	1.8683029901	0	0.0931562491	0
10	2024-08-16	4001	4.7515478665	20.811	3	1	1	0.2138346125	0.379698693	0.0005890207	0.00002236504
11	2024-08-16	4001	4.7515478665	20.811	3	1	2	0.0287066778	0.0509734505	0.0005831889	0.00000297271
12	2024-08-16	4001	4.7515478665	20.811	3	2	1	0.2638693959	0.2922098114	0.0005890207	0.0001721176
13	2024-08-16	4001	4.7515478665	20.811	3	2	2	0.0061086887	0.0072556118	0.0005831889	0.0000042314
14	2024-08-16	4001	4.7515478665	20.811	3	35	2	[NULL]	0.221106073	0.2527427457	0.055882956
15	2024-08-16	4001	4.7515478665	20.811	4	0	1	0.4350972966	0	0.0000000022	0
16	2024-08-16	4001	4.7515478665	20.811	4	0	2	1.1254587655	0	0.0000000437	0
17	2024-08-16	4001	4.7515478665	20.811	4	0	4	1.8683029901	0	0.00000545157	0
18	2024-08-16	4001	4.7515478665	20.811	4	1	1	0.2138346125	0.379698693	0.0000000022	0.0000000008
19	2024-08-16	4001	4.7515478665	20.811	4	1	2	0.0287066778	0.0509734505	0.0000000022	0.0000000001
20	2024-08-16	4001	4.7515478665	20.811	4	2	2	0.2638693959	0.2922098114	0.0000000022	0.0000000006

Tableau 16 : Emissions (head)



## 4.9 - Validation de la méthode

Pourquoi a-t-on choisi cette méthode plutôt qu'une autre parmi toutes celles testées ? Comment s'assurer que les résultats sont cohérents ? Les réponses à ces deux questions se trouvent dans la phase de validation.

L'idée est de comparer la somme de nos consommations journalières par usage et énergie sur tout 2021 avec l'inventaire AtmoSud 2021, qui donne lui des consommations annuelles par usage et par énergie. De plus, on quantifie les différences observées ainsi grâce à deux métriques : RMSE (Root Mean Squared Error) et MAE (Mean Absolute Error). Ce sont deux métriques couramment utilisées en Machine Learning pour évaluer l'erreur de prédiction d'un modèle par exemple. La 1<sup>re</sup>, utilisant des carrés, a tendance à donner plus de poids aux grands écarts ; elle sera alors pertinente à regarder si on veut avant tout limiter les grosses erreurs, quitte à en faire plus de petites. La 2<sup>e</sup> est une simple moyenne des écarts.

On peut faire la comparaison sur l'ensemble de la région :

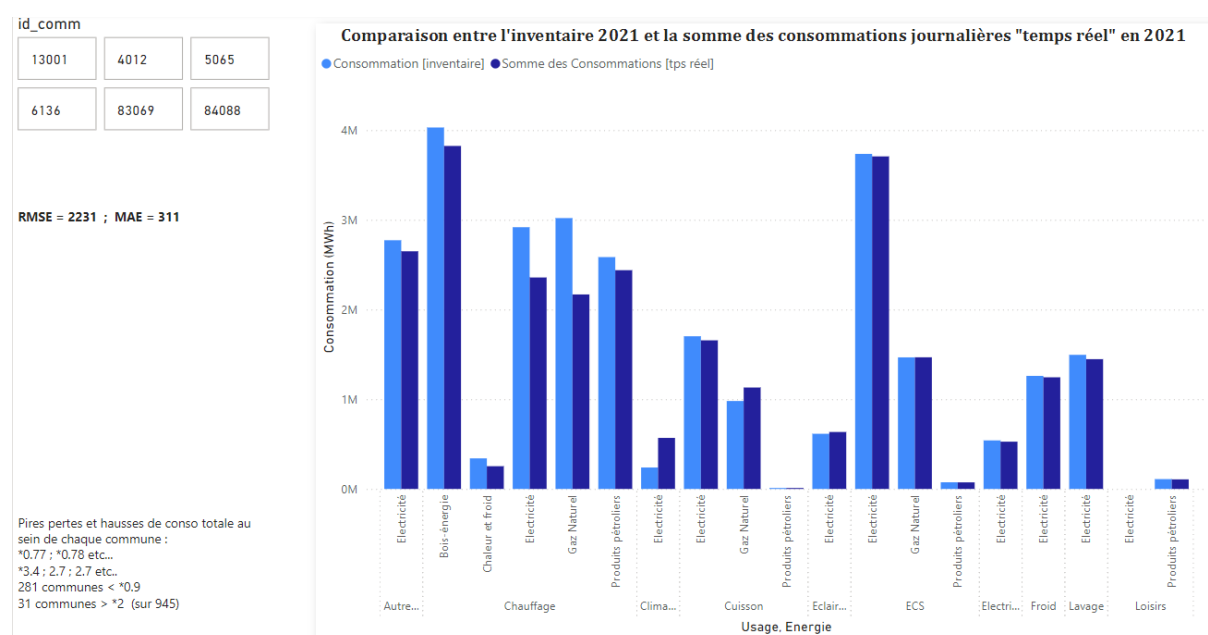
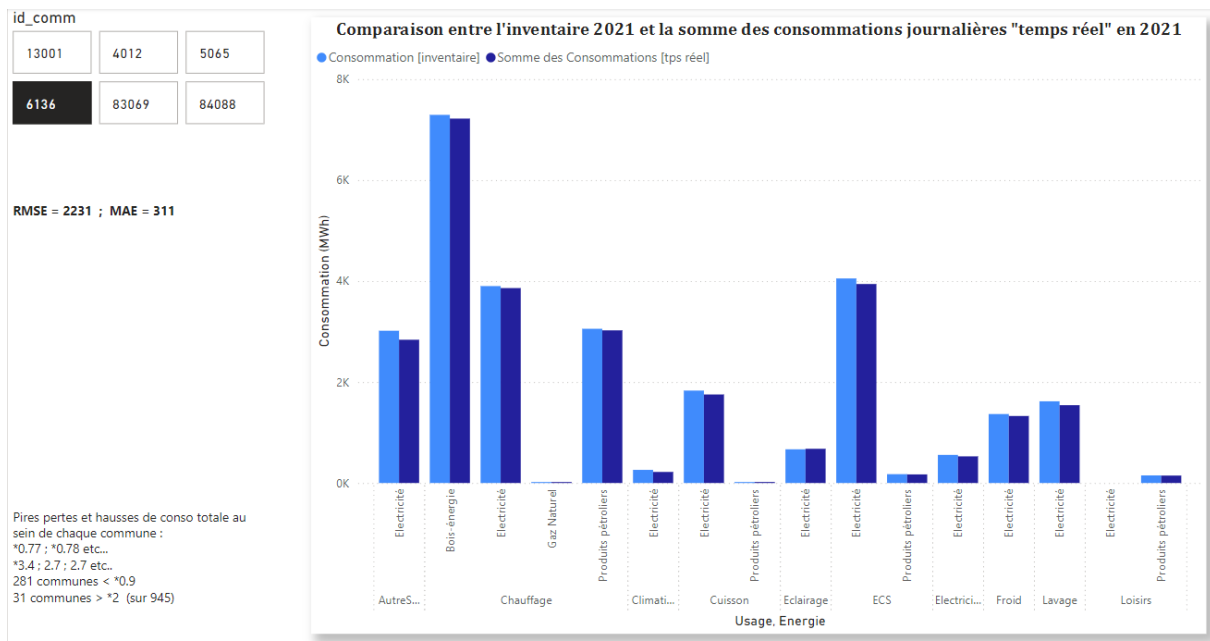
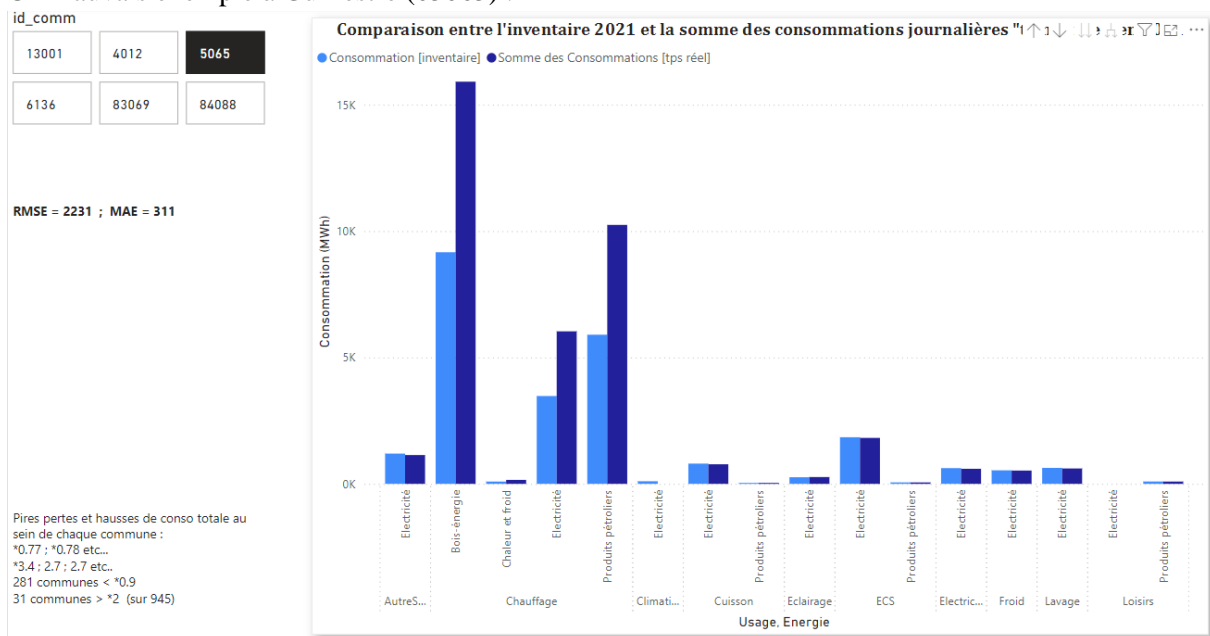


Illustration 9 : Validation avec l'inventaire 2021, sur tout PACA

Mais aussi par commune. Il y a bien sûr des communes où la méthode fonctionne mieux que d'autres, notamment selon la qualité de nos estimations de la consommation électrique de la commune. Un bon exemple à Sospel (06136) :



### Un mauvais exemple à Guillestre (05065) :



C'est ce mode de validation qui nous a mené à choisir cette méthode de calcul des émissions. Les résultats sont effectivement les plus cohérents parmi toutes les méthodes testées. Par exemple, en prenant BASE\_CONSO dans l'inventaire AtmoSud et pas comme une fonction constante de l'estimation de consommation électrique, on obtient ceci :

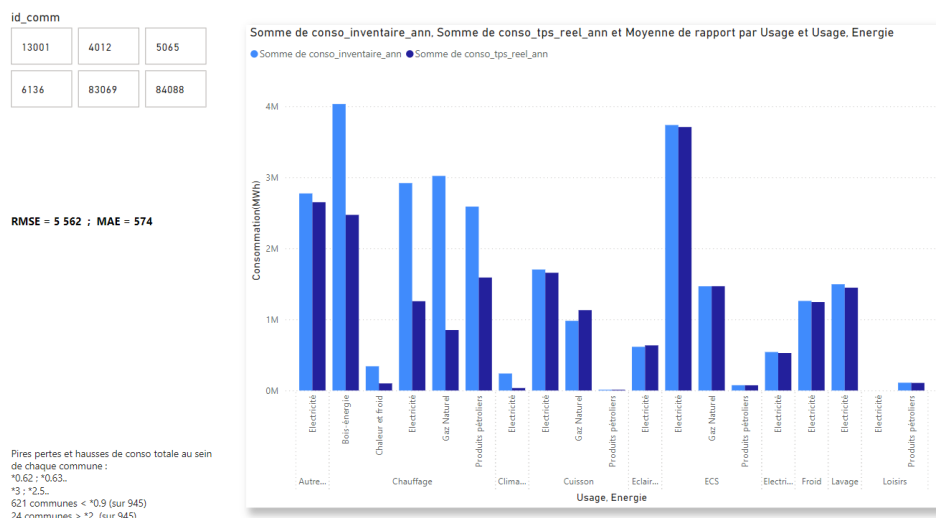


Illustration 12 : Validation avec un autre choix de BASE\_CONSO

## 4.10 - Visualisation

Pour conclure, un tableau de bord Grafana permet de visualiser les résultats en choisissant les filtres souhaités.

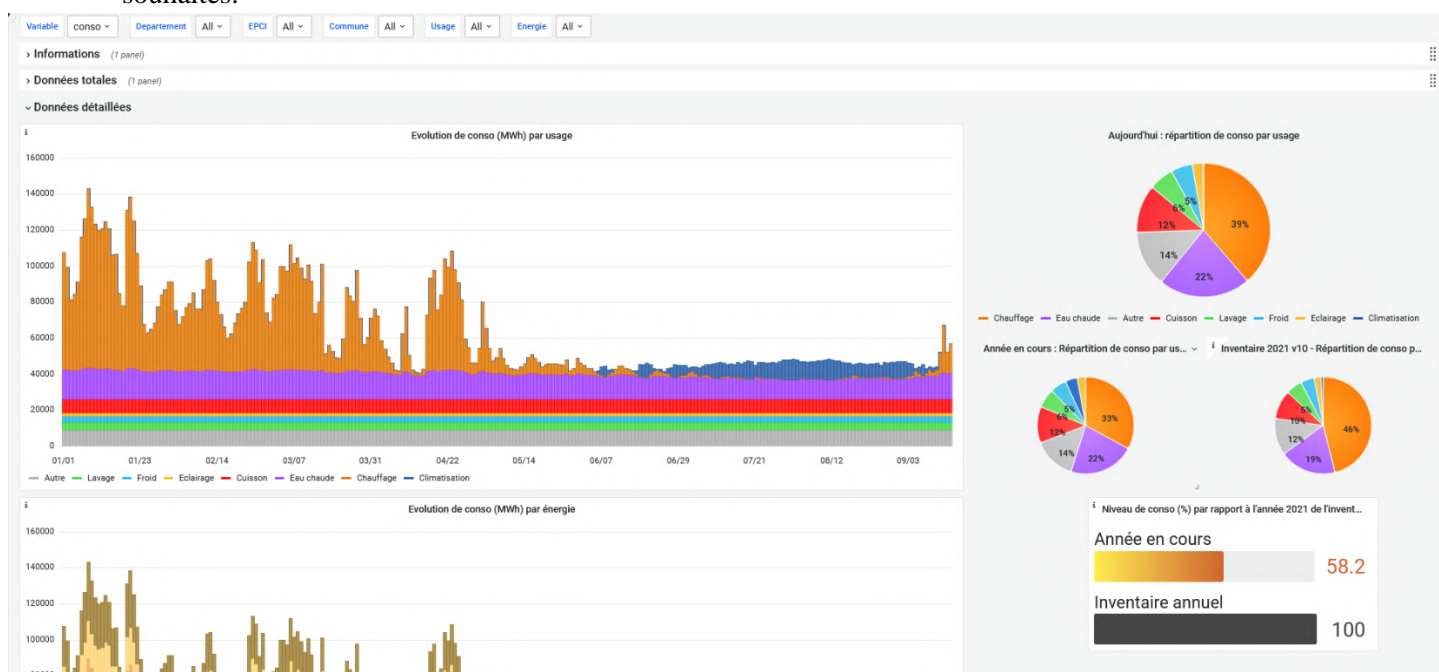


Illustration 13 : Tableau de bord final

Ce tableau de bord est connecté directement aux bases de données mises à jour chaque jour lors du calcul. De plus, ce calcul sous forme de script SQL est effectué automatiquement tous les jours à 13h01, à l'aide d'une machine virtuelle et de l'outil cron. Tout ceci permet d'assurer que les données et visualisations sont toujours à jour, sans nécessiter l'intervention d'un utilisateur.

## 4.11 – Limites et Perspectives

D'abord, la principale limite de cette méthode est l'estimation de la consommation électrique, par nature imparfaite et qui, pourtant, détermine tout le reste. Le modèle est imparfait déjà par sa nature - la température n'explique pas toute la consommation électrique du Résidentiel – mais aussi par son application ; faute de données par communes, nous avons dû entraîner le modèle sur une échelle régionale puis ventiler à la commune à travers une proportion fixe, cela crée une erreur supplémentaire. Tout le calcul découle donc d'une seule donnée « temps réel », la température.

Dans un premier temps, l'on peut se dire que cela reste très utile : on a une idée des variations de consommation et d'émissions attendues selon la météo. En présence d'autres indicateurs, comme par exemple les mesures de concentration effectuées par AtmoSud, on pourrait alors comparer ces indicateurs avec notre indicateur orienté météo, pour identifier ce qui dépend de la météo ou non. Par exemple, si demain j'observe un pic de pollution aux particules fines quelque part grâce aux stations de mesure de concentration, cette méthode saura nous dire plus ou moins précisément dans quelle mesure ce pic est dû à la température ou bien à d'autres phénomènes.

Mais dans un second temps, l'idée reste d'obtenir des données « temps réel » de consommation électrique sans avoir besoin de les estimer avec la température, permettant ainsi de capter toutes les vraies variations de consommation électrique avant de lancer notre calcul. Dans cette optique, ce projet peut faire avancer les futures démarches auprès d'énergéticiens, car il démontre ce qu'AtmoSud peut faire et comment leur données pourront être valorisées.

Ensuite, retraçons quelques hypothèses effectuées lors de ce calcul :

- Les consommations par usage et énergie sont reliées par un ratio de consommation fixe (tous les jours le même, mais différent pour chaque commune). Par exemple, on dit que à Marseille, lorsqu'une unité de consommation électrique (énergie) pour le Chauffage (usage) est consommée, x unités de consommation de Bois (énergie) pour le Chauffage (usage) seront consommées.
- Les usages thermosensibles sont le Chauffage, la Climatisation, l'Eau chaude. Ainsi la Cuisson, l'Eclairage, le Lavage, et le Froid (réfrigérateur et congélateur) sont considérés comme ne dépendant pas de la température, et donc ont un niveau de consommation identique chaque jour

Enfin, avec la meilleure méthode retenue en phase de validation, l'on constate tout de même pour l'usage Chauffage une perte de consommation de 14% au total sur l'année 2021 par rapport à l'inventaire. Néanmoins pour les autres usages, nous n'avons pas d'écart important.

# Conclusion

## Compétences développées

Ce stage de 6 mois m'a permis de grandement renforcer mes compétences et mon expérience sur plusieurs technologies.

D'abord, si j'avais bien eu la chance de suivre des cours de bases de données relationnelles et de SQL à 2 reprises durant mes études, je n'avais jamais eu l'occasion d'exercer ces connaissances en milieu professionnel. Cette fois ça y est ; lors de ce projet, l'intégralité des interactions avec les données stockées sur nos serveurs PostgreSQL s'effectue en SQL. Le script principal du projet qui calcule en temps réel les émissions est donc par exemple intégralement du SQL. Dans le même temps, j'ai pu me familiariser avec la gestion de bases de données PostgreSQL, via l'outil DBeaver. Création et maintenance de tables ainsi que de schémas, connexions pour faire appel à d'autres bases de données au sein d'un même script, autorisations d'accès aux données... Je pense maîtriser l'essentiel.

Dans le même temps, j'ai pu continuer à appliquer mes connaissances déjà correctes en langage Python, après un précédent stage fait exclusivement en Python et de nombreux TD et projets lors de mon M2. Ici, je me suis servi de Python à la fois pour tout ce qui relevait du machine learning, et pour tout ce qui concernait l'extraction de données via une API. J'ai bien progressé dans cette 2<sup>e</sup> application.

Vient ensuite la Data Visualisation ; en + des packages Python comme matplotlib, seaborn ou pyplot, j'ai pu découvrir en profondeur deux outils de visualisation particulièrement intéressants, que sont Power BI et Grafana, en développant des tableaux de bords poussés pour rendre compte de mes calculs en temps réel. C'est une partie de mon stage que j'ai vraiment adorée et qui à coup sûr me servira pour la suite.

D'un point de vue moins centré sur les outils, j'ai pu renforcer quelque peu mes compétences en Machine Learning en appliquant un modèle simple à un problème réel. J'ai eu l'occasion de découvrir un certain nombre de notions sur la qualité de l'air, l'environnement et les émissions, et mon intérêt pour ce domaine n'a fait que grandir au fil des mois passés chez AtmoSud.

## Ressenti personnel

J'ai le sentiment d'avoir vécu une très bonne expérience professionnelle lors de ce stage. Grâce à une bonne ambiance de travail et à des missions en adéquation avec mon expérience, je suis toujours venu au bureau avec le sourire, et n'ai pas compté les semaines. De plus, j'ai l'impression d'avoir bien réalisé ma mission et d'avoir réellement contribué aux projets d'AtmoSud, d'avoir pu apporter. Et cela renforce ma confiance pour la suite de ma carrière, en plus de me conforter dans le fait que j'apprécie ce genre de métier. Il n'y a donc que du positif, et je pourrai désormais me servir de cette expérience réussie pour trouver un 1<sup>er</sup> emploi.

# Références

## Intro

Contexte

-<https://www.ecologie.gouv.fr/politiques-publiques/polluants-lair-situation-impacts-encadrement>

## IV

4.6.2 Correction de l'ECS

- "ANDREAU N., LITVINE D., SPILEMON J., 2022. *Le projet ECSpect - Pratiques et consommation d'eau chaude sanitaire : enseignements d'une expérimentation sociotechnique dans le secteur domestique*, 173 pages, pour le compte de l'ADEME"

# Annexes

## Annexe 1 : Bulle d'air diffusée auprès des adhérents d'AtmoSud le 09 août 2024

### **Inventaire des émissions : de l'annuel au temps réel !**

Dans la continuité des travaux menés ces dernières années dans la production de données d'émission, AtmoSud travaille à l'amélioration des sources de données et des méthodologies. Le processus de réalisation des bilans annuels est long (en moyenne 2 ans) puisqu'il nécessite que l'ensemble des données soient pourvues par les fournisseurs, qu'elles soient validées et consolidées par les instances nationales puis mises à disposition sur des plateformes d'accès. S'en suit, le processus de centralisation puis, selon l'évolution des connaissances et donc des hypothèses, de mise en place de calculs.

Ces données, disponibles sur notre plateforme [CIGALE](#), sont donc ce qu'il y a de plus fiable en termes d'émission de polluants et d'identification des sources à l'échelle régionale.

Afin de répondre au mieux aux besoins des acteurs, AtmoSud engage des travaux pour fournir des données à un pas de temps de plus en plus fin.

Cela nécessite de développer de nouvelles méthodes permettant d'obtenir des données journalières conditionnées aux connaissances d'évolution des consommations d'énergies en fonction des périodes, des saisons et de paramètres environnementaux tels que la température.

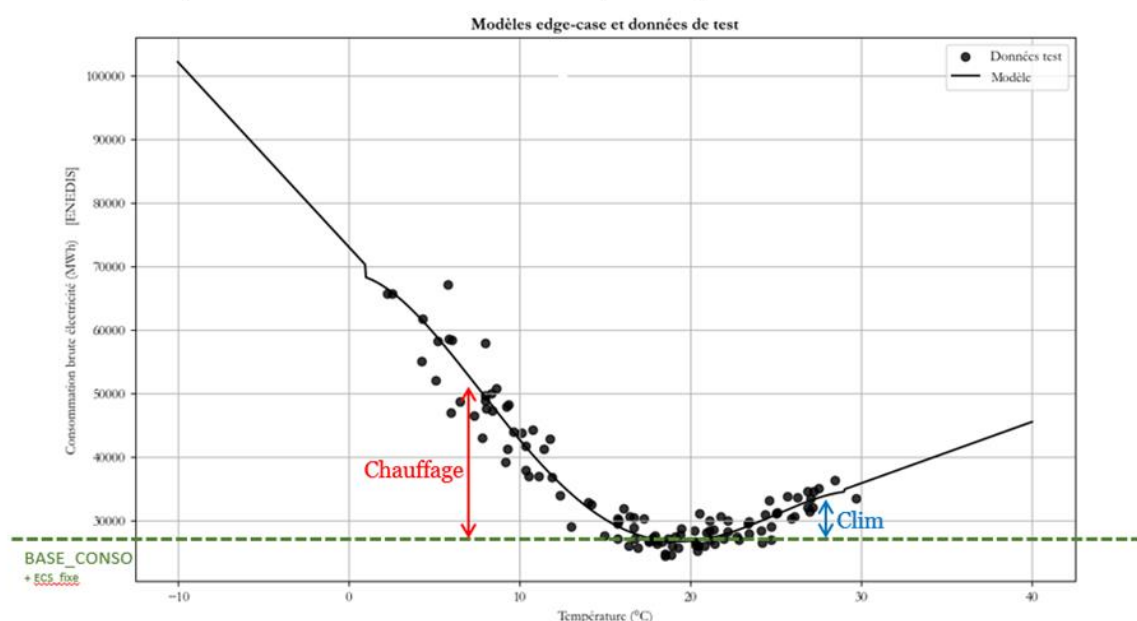
Pour le secteur résidentiel il existe un lien direct entre consommations et émissions. L'objectif est donc tout d'abord d'évaluer les **consommations par usage et par énergie**. Les émissions sont ensuite calculées à partir des consommations en fonction des facteurs d'émission correspondants.

### **Principe de la méthode consommations "Temps réel résidentiel" retenue :**

- Pour les usages "fixes" dont on considère qu'ils n'ont pas de composante saisonnière, on utilise l'inventaire annuel pour déterminer une consommation journalière par usage et énergie,
- Pour les usages dits "thermosensibles", les consommations journalières sont calculées selon la température.

La décomposition de la consommation électrique est présentée ci-après :

## Décomposition de la consommation électrique en 2 parties : fixe et thermosensible



**BASE\_CONSO** = Moyenne journalière de consommation électrique pour les usages non thermosensibles (Cuisson, Lavage etc..)

Pour chaque jour de l'année, il est comptabilisé une consommation d'électricité fixe (BASE\_CONSO, représentée en vert). A celle-ci peut s'ajouter, en fonction de la température, une surconsommation liée à l'utilisation de moyens de chauffage ou de climatisation. Il en résulte une consommation électrique totale conditionnée aux températures.

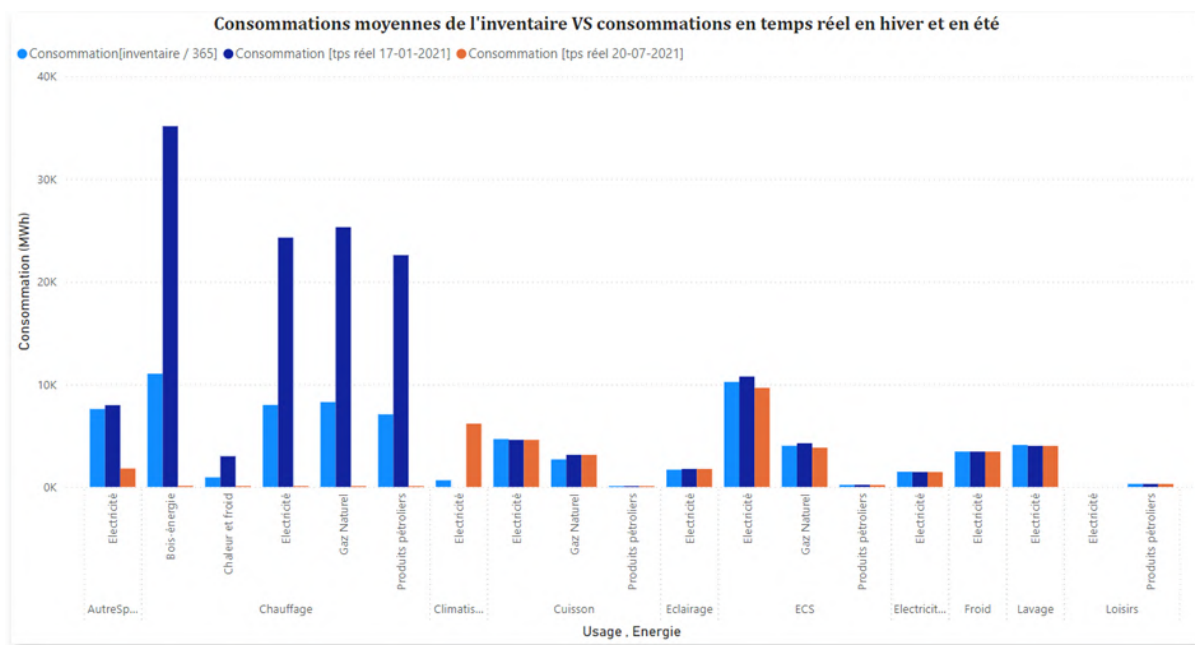
L'inventaire annuel est utilisé pour propager ces consommations électriques sur les autres types d'énergie. Exemple : "Si on consomme 100 MWh d'électricité pour le chauffage, l'inventaire 2021 indique une consommation de 180 MWh de Bois-Energie pour ce même usage Chauffage". On obtient ainsi un niveau de consommation pour chaque usage et chaque énergie.

### Intérêt de la méthode

Une consommation moyenne journalière (consommations annuelles divisée par 365) ne prendrait pas en compte les spécificités de la saison et des conditions climatiques. La méthode développée, qui module les usages selon la température du jour, permet d'adapter les consommations à l'utilisation de chauffage et climatisation chaque jour de l'année pour, *in fine*, être au plus proche de la réalité des émissions du jour.

Regardons un exemple de comparaison des consommations d'un jour d'été (sans chauffage mais avec climatisation), d'un jour d'hiver (chauffage) et de la consommation moyenne journalière :

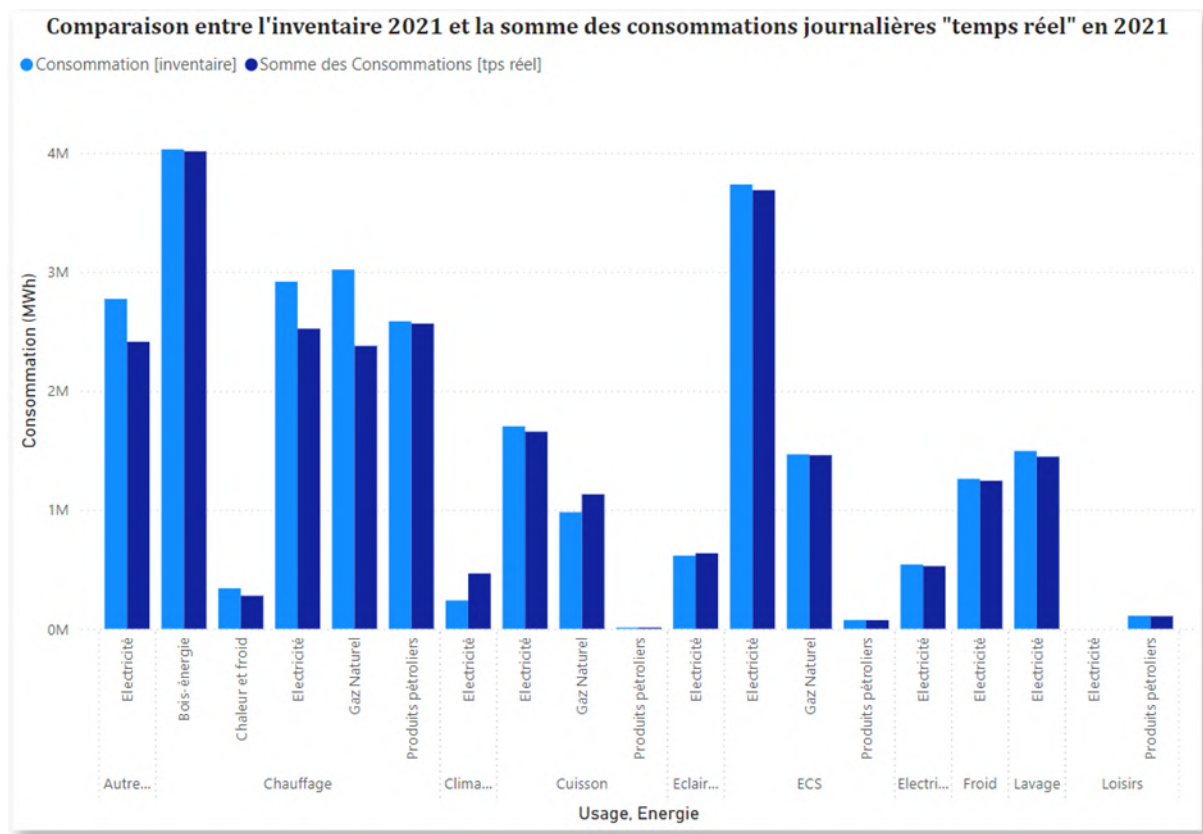




On observe bien que les consommations de bois énergie, d'électricité, de gaz naturel et de produits pétroliers en lien avec le chauffage sont beaucoup plus importants pour un jour d'hiver qu'un jour d'été mais également plus importants que la moyenne des consommations journalière. A l'inverse en été, on observe une augmentation des consommations électriques liées à la climatisation.

### Validation de la méthode

La validation de la méthode demande de vérifier que le cumul annuel des consommations journalières calculées en fonction des températures correspond bien au bilan des consommations annuelles. Le test a été effectué sur l'année 2021 et les résultats sont bien cohérents.



Il ne reste alors plus qu'à appliquer les facteurs d'émissions aux consommations d'énergie pour obtenir les émissions du secteur résidentiel !

## Annexe 2 : Script SQL principal

```
-- Emissions en temps reel du Residentiel. Execution : ~6 minutes
-- METHODE NUMERO 2 - tps réel
--> on considère que toute variation de conso > [conso "de base" sans
Chauff ni clim] est dûe au chauffage ou à la clim.
-- permet de fixer toutes les consos autres, et on fait varier
quotidiennement chauffage et clim en fonction de la température
-- Les tables fixes
/*
-- TEMPERATURE MOY HISTORIQUE DE CHAQUE COMMUNE - utile pour
correction ECS. nécessite src_meteo_hist. Existe dans schema
prj_res_tps_reel. Code dans Hist_res_calcul2.sql si besoin.
drop table if exists prj_res_tps_reel.w2_temperatures_moy;
create table prj_res_tps_reel.w2_temperatures_moy as
SELECT id_station AS id_comm, AVG(Temperature) AS temperature_moy
FROM prj_res_tps_reel.src_meteo_hist
WHERE id_station / 1000 IN (4, 5, 6, 13, 83, 84)
      and EXTRACT(year from date_tu) BETWEEN EXTRACT(year from
CURRENT_DATE) - 3 AND EXTRACT(year from CURRENT_DATE) - 1      --
moy sur 2021-2023
GROUP BY id_station;
comment on table prj_res_tps_reel.w2_temperatures_moy is
'Températures moyennes de chaque commune en Région Sud, prise sur la
période 2018-2021 (4 ans)';
select * from prj_res_tps_reel.w2_temperatures_moy; */
/*
-- PART DE CONSO DE CHAQUE COMMUNE EN PACA (RESIDENTIEL)
drop table if exists prj_res_tps_reel.w2_parts_conso_res_comm;
create table prj_res_tps_reel.w2_parts_conso_res_comm as
select *, Conso_annee / Conso_totale_annee as Part_Conso_comm from
(select id_comm, sum(val) as Conso_Annee
from total.bilan_comm_v10_diffusion bcvd natural join
commun.tpk_polluants tp natural join total.tpk_cat_energie_color tcec
where tp.nom_court_polluant = 'conso'
      and tcec.nom_court_cat_energie = 'Electricité'
      and id_secteur_detail in (3) -- Résidentiel
      and an between (select max(an)- 3 from
total.bilan_comm_v10_diffusion) and (select max(an) from
total.bilan_comm_v10_diffusion) -- 2018-2021 inclus
      and id_comm/1000 in (4, 5, 6, 13, 83, 84)
      and format_detaill_scope_2=1
group by id_comm) as consos
cross join (
select sum(val) as Conso_Totale_Annee
from total.bilan_comm_v10_diffusion bcvd natural join
commun.tpk_polluants tp natural join total.tpk_cat_energie_color tcec
where tp.nom_court_polluant = 'conso'
```

```

        and tcec.nom_court_cat_energie = 'Electricité'
        and id_secteur_detail in (3)      -- Résidentiel
        and      an      between      (select      max(an)-      3      from
total.bilan_comm_v10_diffusion)      and      (select      max(an)      from
total.bilan_comm_v10_diffusion)      -- 2018-2021 inclus
        and id_comm/1000 in (4, 5, 6, 13, 83, 84)
        and format_detaillé_scope_2=1) as conso
;
comment on table prj_res_tps_reel.w2_parts_conso_res_comm is 'Part de
consommation de chaque commune en région PACA, prise sur la période
2018-2021 (4 ans)';
select * from prj_res_tps_reel.w2_parts_conso_res_comm; */
/*
-- CONSO USAGE ENERGIE POUR LES USAGES FIXES : moyenne journalière
des 4 dernières années
drop table if exists prj_res_tps_reel.w2_consos_fixes_usage_energie;
create table prj_res_tps_reel.w2_consos_fixes_usage_energie as (
select id_comm, bcvd.id_usage, usages.nom_usage as usage_old,
bcvd.code_cat_energie, energ.nom_court_cat_energie as energie_old,
sum(val) / (4 * 365.25) as Consommation, id_unite
from total.bilan_comm_v10_diffusion bcvd
natural join total.tpk_secteur_emi_detail sect
natural join commun.tpk_polluants p
natural join total.tpk_cat_energie_color energ
natural join commun.tpk_usages usages
where p.nom_court_polluant = 'conso' -- conso = COnsommation finale à
climat réel
        and lib_secteur_detail in ('Résidentiel')
        and      an      between      (select      max(an)      -      3      from
total.bilan_comm_v10_diffusion)      and      (select      max(an)      from
total.bilan_comm_v10_diffusion)      --2018-2021 inclus -> moyenne
        and format_detaillé_scope_2=1
        and id_comm/1000 in (4,5,6,13,83,84)
group      by      id_comm,      bcvd.id_usage,      usages.nom_usage,
bcvd.code_cat_energie, energ.nom_court_cat_energie, id_unite
order by id_comm, usage_old, energie_old
);
comment on table prj_res_tps_reel.w2_consos_fixes_usage_energie is
'Consommations fixes moyennes journalières par comunne, par usage
(pour les usages non thermosensibles), et énergie, prise sur la
période 2018-2021 (4 ans)';
select * from prj_res_tps_reel.w2_consos_fixes_usage_energie; */

-- TEMPERATURE -> Estimer conso Chauff/Clim
-- connexion à cadastre pour import meteo
/*
create extension if not exists postgres_fdw;

```

```

CREATE SERVER fdw_cadastre
FOREIGN DATA WRAPPER postgres_fdw
OPTIONS (host '172.16.13.249', port '5432', dbname 'cadastre',
updatable 'false');
create user mapping for user
    server fdw_cadastre
    options ( user 'postgres', password 'postgres'); -- le mapping
est crée */

-- import table meteo
drop foreign table if exists prj_res_tps_reel.src_meteo_tps_reel;
create foreign table prj_res_tps_reel.src_meteo_tps_reel(
    id_station_meteo INT,
    date_jour timestamp,
    temperature FLOAT,
    echeance INT
)
server fdw_cadastre
options (schema_name 'airesv5', table_name 'meteo_routier_tps_reel');
-- group by JOUR et Calcul DJ17, DJ21.
    LONG (+1min)
drop table if exists prj_res_tps_reel.w_meteo_tps_reel;
create table prj_res_tps_reel.w_meteo_tps_reel as
WITH MinEcheance AS (
    SELECT
        id_station_meteo,
        CAST(date_jour AS DATE) AS date,
        MIN(echeance) AS min_echeance
    from prj_res_tps_reel.src_meteo_tps_reel
    GROUP by id_station_meteo, CAST(date_jour AS DATE)
)
select t.id_station_meteo as id_comm, CAST(t.date_jour AS DATE) AS
date,
    CASE
        WHEN EXTRACT(DOW FROM date) = 0 THEN 'Dimanche'
        WHEN EXTRACT(DOW FROM date) = 1 THEN 'Lundi'
        WHEN EXTRACT(DOW FROM date) = 2 THEN 'Mardi'
        WHEN EXTRACT(DOW FROM date) = 3 THEN 'Mercredi'
        WHEN EXTRACT(DOW FROM date) = 4 THEN 'Jeudi'
        WHEN EXTRACT(DOW FROM date) = 5 THEN 'Vendredi'
        WHEN EXTRACT(DOW FROM date) = 6 THEN 'Samedi'
    end as Weekday,
    AVG(t.temperature) AS temperature,
    case
        when AVG(t.temperature) < 17 then 17 - AVG(t.temperature)
        else 0
    end as DJ17,
    case

```

```

        when AVG(t.temperature) > 21 then AVG(t.temperature) - 21
        else 0
    end as DJ21
from prj_res_tps_reel.src_meteo_tps_reel t INNER join MinEcheance me
ON
    t.id_station_meteo = me.id_station_meteo
    AND CAST(t.date_jour AS DATE) = me.date
    AND t.echeance = me.min_echeance
GROUP by t.id_station_meteo, CAST(t.date_jour AS DATE), EXTRACT(DOW
FROM date)
ORDER by date desc, t.id_station_meteo;
;
comment on table prj_res_tps_reel.w_meteo_tps_reel is 'Température,
DJ17 et DJ21 quotidiens pour chaque commune en Région Sud jusqu'à
aujourd'hui';

```

```

--CONSO ELEC ESTIMEE AVEC MODELE
-- conso modele : entraînés sur des données de températures PACA
(moyenne brute)
drop table if exists prj_res_tps_reel.w_conso_res_estim_tr;
create table prj_res_tps_reel.w_conso_res_estim_tr as
select date, id_comm, temperature, part_conso_comm,
    CASE WHEN Temperature < 3 THEN (66704.23588245558 + (-
3902.887827704622 * Temperature)) * part_conso_comm
    WHEN Temperature BETWEEN 3 and 27 THEN (87430.11 + (-
20095.08277786156 * POW(Temperature, 1))+ (4219.728678066393 *
POW(Temperature, 2))+ (-482.5300092060183 * POW(Temperature, 3))+
(28.074054017762844 * POW(Temperature, 4))+ (-0.7877522906343015 *
POW(Temperature, 5))+ (0.008504717233948255 * POW(Temperature, 6))) *
part_conso_comm
    ELSE (13295.935215642927 + (799.4012802720838 *
Temperature)) * part_conso_comm
    END AS Consommation_est_mwh
from prj_res_tps_reel.w_meteo_tps_reel meteo natural join
prj_res_tps_reel.w2_parts_conso_res_comm parts
order by date desc, id_comm;
comment on table prj_res_tps_reel.w_conso_res_estim_tr is 'Estimation
de conso électrique res en temps réel (jour) pour chaque commune,
selon la température';

```

```

-- SEPARATION DE LA CONSO
-- sans utiliser l'inventaire : BASE_CONSO[commune] = Min(Estimation)
- ecs_fixe
-- CONSO ELEC = BASE_CONSO + CONSO_CHAUFF + CONSO_CLIM. Identifions
conso chauffage et conso clim, en regardant la diff entre la conso
estimee totale et base_conso.

```

```

drop table if exists prj_res_tps_reel.w2_consos_sep_tr;
create table prj_res_tps_reel.w2_consos_sep_tr as
WITH min_avg_conso AS (
    SELECT
        id_comm,
        min_consommation_est_mwh,
        conso_elec_ecs_fixe
    FROM (SELECT id_comm, MIN(consommation_est_mwh) AS
min_consommation_est_mwh
FROM prj_res_tps_reel.w_conso_res_estim_hist GROUP BY id_comm) as min_est
    natural join (select id_comm, consommation as
conso_elec_ecs_fixe
from prj_res_tps_reel.w2_consos_fixes_usage_energie where id_usage = 1 and
code_cat_energie = 8) as ecs -- conso journalière d'elec pour ECS
    par commune
),
max_tempe_diff as (
    select id_comm, max(abs(tmoy.temperature_moy - temperature) /
0.28) as X -- Pour modif à max +/- 28% l'ECS
    from prj_res_tps_reel.w2_temperatures_moy tmoy natural join
prj_res_tps_reel.w_conso_res_estim_tr
    group by id_comm)
SELECT
    date,
    w.id_comm,
    mac.min_consommation_est_mwh - mac.conso_elec_ecs_fixe as
base_consoelec_mwh,
    temperature,
    consommation_est_mwh,
    mac.conso_elec_ecs_fixe as conso_elec_ecs_fixe_mwh,
    mac.conso_elec_ecs_fixe * (1 + (tmoy.temperature_moy -
temperature) / max_tempe_diff.X) as conso_elec_ecs_mwh, --
Correction ecs selon meteo
    CASE -- Chauffage = Conso_est - base_conso - conso_ECS (corrigée)
        WHEN temperature < 17 THEN consommation_est_mwh -
(mac.min_consommation_est_mwh - mac.conso_elec_ecs_fixe) -
(mac.conso_elec_ecs_fixe * (1 + (tmoy.temperature_moy - temperature)
/ max_tempe_diff.X))
        ELSE 0 -- bornée à 0 en bas
    END AS conso_elec_chauffage_mwh,
    CASE
        WHEN temperature > 21 THEN consommation_est_mwh -
(mac.min_consommation_est_mwh - mac.conso_elec_ecs_fixe) -
(mac.conso_elec_ecs_fixe * (1 + (tmoy.temperature_moy - temperature)
/ max_tempe_diff.X))
        ELSE 0
    END AS conso_elec_clim_mwh
FROM

```

```

prj_res_tps_reel.w_conso_res_estim_tr w
JOIN min_avg_conso mac using (id_comm)
join prj_res_tps_reel.w2_temperatures_moy tmoy using (id_comm)
join max_tempe_diff using (id_comm)
order by date desc;
comment on table prj_res_tps_reel.w2_consos_sep_tr is 'Décomposition
de la conso électrique (BASE_CONSO + ECS + CHAUFF/CLIM) par jour et
commune';

-- RATIOS DE CONSO : voir Calcul_ratios_conso.sql
--select * from prj_res_tps_reel.w2_ratios_conso_uvariables;

-- APPLICATION CONSO * RATIO_CONSO
-- provisoire : clim non incluse dans les lignes car pas de ratio
drop table if exists consos_u_e_estim_tr;
create temp table if not exists consos_u_e_estim_tr AS
select date, id_comm, base_consoelec_mwh, temperature,
conso_elec_chauffage_mwh, conso_elec_clim_mwh, id_usage,
ratios_bis.usage_old, code_cat_energie, ratio_conso_uvariable,
case
when id_usage not in (0, 1, 21) then
consos_ue_inv.consommmation -- usages fixes : moyenne j des consos
de cet usage cette énergie
when id_usage = 0 then conso_elec_chauffage_mwh *
ratio_conso_uvariable --usages variables : conso elec * ratio [ratio
= conso energie usage / conso elec usage]
when id_usage = 1 then conso_elec_ecs_mwh *
ratio_conso_uvariable
when id_usage = 21 then conso_elec_clim_mwh
end as conso_u_e_mwh
from prj_res_tps_reel.w2_consos_sep_tr
inner join prj_res_tps_reel.w2_ratios_conso_uvariables
ratios_bis using (id_comm)
inner join prj_res_tps_reel.w2_consos_fixes_usage_energie
consos_ue_inv using (id_comm, id_usage, code_cat_energie)
order by date desc, id_comm, id_usage, code_cat_energie
;

-- Inclusion lignes CLIM par force
drop table if exists prj_res_tps_reel.w2_consos_u_e_estim_tr;
CREATE table if not exists prj_res_tps_reel.w2_consos_u_e_estim_tr as
WITH climatisation AS (
SELECT distinct
date, id_comm, base_consoelec_mwh, temperature,
conso_elec_chauffage_mwh, conso_elec_clim_mwh, 21 AS
id_usage, 'Climatisation' as usage_old, 8 AS code_cat_energie,

```



```

1 AS ratio_conso_uvariable, conso_elec_clim_mwh AS
conso_u_e_mwh
FROM consos_u_e_estim_tr
order by date desc
)
SELECT date, id_comm, base_consoelec_mwh, temperature,
conso_elec_chauffage_mwh, conso_elec_clim_mwh, id_usage, usage_old,
code_cat_energie,
ratio_conso_uvariable, conso_u_e_mwh
FROM consos_u_e_estim_tr
union ALL
SELECT * FROM climatisation
ORDER BY date DESC, id_comm, id_usage, code_cat_energie
;
-- 7,937% de AutreSpe2 est de la clim -> on les enlève
update prj_res_tps_reel.w2_consos_u_e_estim_tr
set conso_u_e_mwh = conso_u_e_mwh * (1 - 0.07937)
where id_usage = 34 -- AutreSpe2
;
-- Cles étrangères
ALTER TABLE prj_res_tps_reel.w2_consos_u_e_estim_tr
ADD CONSTRAINT fk_id_usage5
FOREIGN KEY (id_usage)
REFERENCES commun.tpk_usages(id_usage),
ADD CONSTRAINT fk_code_cat_energie5
FOREIGN KEY (code_cat_energie)
REFERENCES total.tpk_cat_energie_color(code_cat_energie);
comment on table prj_res_tps_reel.w2_consos_u_e_estim_tr is
'Consommations par usage et énergie en temps réel à partir des
températures (modèle d estimation) et des ratios de consommation. Pour
chaque jour et commune';

-- CALCUL FACTEURS D'EMISSION : Emission_usage_energie /
conso_usage_energie. Par commune, par polluant
-- Voir Calcul_fe.sql pour refaire le calcul.
--SELECT * FROM prj_res_tps_reel.w_facteurs_emiss_res;

-- -- CALCUL EMISSIONS TEMPS REEL : Emission [jour, commune, polluant,
usage, energie] = conso[jour, commune, usage, energie] * fe [commune,
polluant, usage, energie]
drop table if exists prj_res_tps_reel.emissions_estim_tr;
create table prj_res_tps_reel.emissions_estim_tr as

```

```

        select      conso.date,      conso.id_comm,      base_consoelec_mwh,
temperature,      id_polluant,      id_usage,      code_cat_energie,
ratio_conso_uvariable,
facteur_emission_kg_by_mwh,
conso_u_e_mwh      *      facteur_emission_kg_by_mwh      as
Emission_tps_reel_kg
    from prj_res_tps_reel.w2_consos_u_e_estim_tr conso
        natural join prj_res_tps_reel.w_facteurs_emiss_res fe
    order by date desc, id_comm, id_polluant, id_usage,
code_cat_energie;
ALTER TABLE prj_res_tps_reel.emissions_estim_tr -- Cles étrangères
ADD CONSTRAINT fk_id_usage5
FOREIGN KEY (id_usage)
REFERENCES commun.tpk_usages(id_usage),
ADD CONSTRAINT fk_code_cat_energie5
FOREIGN KEY (code_cat_energie)
REFERENCES total.tpk_cat_energie_color(code_cat_energie),
add constraint fk_id_polluant5
foreign key (id_polluant)
references commun.tpk_polluants(id_polluant)
;
COMMENT ON TABLE prj_res_tps_reel.emissions_estim_tr is 'Calcul final
(basé sur des estimations de conso) des émissions temps réel sur les
communes. Emission = ConsoEstimee * FE';

-- Export
drop table if exists prj_res_tps_reel.export_emissions_estim_tr;
create table prj_res_tps_reel.export_emissions_estim_tr as
select distinct date, id_comm, comm.nom_comm, comm.siren_epci,
comm.nom_epci, id_comm/1000 as departement, temperature, id_polluant,
    case
        when nom_court_polluant = 'PRG100_3GES' then 'Total_GES'
        else nom_court_polluant
    end as nom_court_polluant,
    id_usage,
    case -- AutreSpe2, Loisirs et Elec spécifique -> Autre
        when nom_usage in ('AutreSpe2', 'Loisirs', 'Electricité
spécifique') then 'Autre'
        when nom_usage = 'ECS' then 'Eau chaude'
        else nom_usage
    end as nom_usage,
    code_cat_energie, nom_court_cat_energie, conso_u_e_mwh,
emission_tps_reel_kg
from      prj_res_tps_reel.emissions_estim_tr      natural      join
commun.tpk_usages      natural join total.tpk_cat_energie_color      natural
join commun.tpk_polluants

```

```

natural join (select id_comm, nom_comm, siren_epci, nom_epci from
commun.tpk_comm_evol comm1 where
                                an_ref = (select max(an_ref) from
commun.tpk_comm_evol comm2 where comm1.id_comm = comm2.id_comm)) comm
where nom_court_polluant in('CO', 'COVNM', 'NOx', 'PM10', 'PM2.5',
'SOx', 'NH3', 'CH4.EQCO2', 'CO2.bio', 'CO2.nbio', 'N2O.EQCO2',
'PRG100_3GES', 'Benzene', 'BenzoAPyren', '16hap', 'BC', 'conso')
-- filtre date pour limiter la taille ? : and date BETWEEN
CURRENT_DATE - interval '6 months' and CURRENT_DATE + interval '3
days'
order by date desc, id_comm, id_polluant, id_usage, code_cat_energie;
ALTER TABLE prj_res_tps_reel.export_emissions_estim_tr -- Cles
étrangères
ADD CONSTRAINT fk_id_usage6
FOREIGN KEY (id_usage)
REFERENCES commun.tpk_usages(id_usage),
ADD CONSTRAINT fk_code_cat_energie6
FOREIGN KEY (code_cat_energie)
REFERENCES total.tpk_cat_energie_color(code_cat_energie),
add constraint fk_id_polluant6
foreign key (id_polluant)
references commun.tpk_polluants(id_polluant);
COMMENT ON TABLE prj_res_tps_reel.export_emissions_estim_tr is
'Version Visualisation de emissions_estim_tr';
select * from prj_res_tps_reel.export_emissions_estim_tr;

-- DICOS pour optimisation viz
----drop table if exists prj_res_tps_reel.dico_polluants_u_e;
----drop table if exists prj_res_tps_reel.dico_communes;
create table if not exists prj_res_tps_reel.dico_communes as
select distinct departement, siren_epci, nom_epci, id_comm, nom_comm
from prj_res_tps_reel.export_emissions_estim_tr;
create table if not exists prj_res_tps_reel.dico_polluants_u_e as
(select distinct id_polluant,
    case
        when nom_court_polluant = 'PRG100_3GES' then 'Total_GES'
        else nom_court_polluant
    end as nom_court_polluant,
    id_usage, nom_usage, code_cat_energie, nom_court_cat_energie
from prj_res_tps_reel.export_emissions_estim_tr)
;

-- Autorisations extérieures
Grant select on all tables in schema prj_res_tps_reel to readonly;

```