
Rapport de stage individuel

5^{ème} année

Accompagnement au décret tertiaire

Développement d'un modèle prédictif des économies d'énergies
mobilisables sur le patrimoine ferroviaire

AREP

16 Avenue d'Ivry
75013 PARIS



AREP

Tuteurs entreprise :

Milan BILLOIN, Frédéric RONSEAUX

Tuteur académique :

Mindjid Maizia

Autrice :

Anne-Sophie Girot

IUT

Année scolaire :

2022-2023

Table des matières

Remerciements	3
L'organisme d'accueil	4
L'entreprise	4
Le stage	4
Présentation de la mission	5
Contexte de l'affaire.....	5
Contexte de la mission.....	6
Description et enjeux de la mission.....	6
Résolution de la mission	7
Etat de l'art.....	7
Proposition d'un premier modèle.....	10
Test du modèle sur des données de validation et stratégies d'amélioration.....	13
Test du modèle sur des données terrain.....	15
Pistes pour l'élaboration d'une stratégie à l'échelle du patrimoine.....	19
Analyse critique et pistes d'amélioration	20
Retour d'expérience et perspectives professionnelles.....	23
Retour d'expérience.....	23
Perspectives professionnelles	24
Bibliographie	25
Webographie.....	26

Table des compléments

Equations

Équation 1 : nombre de simulations par action selon le cardinal de chaque critère (A.S. Girot, 2023)	11
Équation 2 : détermination du coefficient multiplicateur optimal (A.S. Girot, 2023).....	13
Équation 3 : Gains selon les besoins de chauffage et rendements des systèmes de chauffage initial et final (A.S. Girot, 2023)	14
Équation 4 : Gains selon consommation initiale et rendements (A.S. Girot, 2023)	14
Équation 5 : Approche multiplicative adoptée pour le calcul des gains de bouquets d'actions, à scénario le plus favorable (A.S. Girot, 2023)	19

Figures

Figure 1 : objectifs de performance énergétique à échéances 2030, 2040 et 2050 (2022)	5
Figure 2 : performances énergétiques initiales et objectifs horizon 2050 (A. Goupil, 2022)	5
Figure 3 : Organisation de la phase 3 de la mission décret tertiaire (W.Teyssier, 2022)	6
Figure 4 : interface du logiciel ClimAudit dans le cadre des calculs d'audits (BBS logiciels, s.d.).....	7
Figure 5 : Formule finale et poids statistiques utilisés pour le calcul des besoins de chauffage (A.Rogeu,2020).....	8
Figure 6 : valeurs forfaitaires en kWhcumac pour l'action d'isolation des murs extérieurs par l'intérieur (Ministère Ecologie Energie Territoires, 2023)	8
Figure 7 : Répartition du nombre de gares par classe de surfaces (m²) (. Goupil, 2022)	10
Figure 8 : Régressions selon DJU chauffage pour l'action pose de vitrages performants (AS. Girot, 2023) ..	11
Figure 9 : Moyennes des gains par sous-groupe d'orientation, action pose de vitrages performants (AS. Girot, 2023)	11
Figure 10 : Coefficients et incertitudes normalisées, régression pour l'action d'isolation des planchers hauts, sous groupe local chauffé donnant sur combles (A.S. Girot, 2023)	12
Figure 11 : Premier modèle avec coefficients pour l'action d'isolation des planchers hauts (A.S. Girot, 2023)	12
Figure 12 : répartition des erreurs relatives (A.S. Girot, 2023)	13
Figure 13 : évolution de l'erreur relative selon corrections apportées (A.S. Girot, 2023)	13
Figure 14: Analyse comparés des gains obtenus par action, selon méthode utilisée, pour deux gares auditées (A.S. Girot, 2023)	14
Figure 15 : Heatmap des gains rapportés à la consommation, par gare et par action testées (A.S. Girot, 2023)	16
Figure 16 : Heatmap des gains en MWh _{eff} /an, par gare et par action testées (A.S. Girot, 2023)	18
Figure 17 : Diagramme de Sankey, avec consommations initiales, à horizon 2050 et gains totaux, regroupés par typologie des gares étudiées (A.S. Girot, 2023)	19
Figure 18 : Répartition des surfaces et consommations par classe de surfaces pour les gares IDF et DOT (A. Goupil, 2022).....	20

Tableaux

Tableau 1: Tableau récapitulatif des diverses méthodes de calcul étudiées (A.S Girot, 2023)	9
Tableau 2 :critères et valeurs des critères pour l'action d'isolation des planchers hauts (AS. Girot, 2023)...	10

Remerciements

Je remercie l'ensemble de l'équipe pédagogique du Département Aménagement et Environnement de Polytech Tours, pour leur accompagnement et leur bienveillance. Je remercie particulièrement mon tuteur académique, Professeur Mindjid Maizia, pour sa rigueur intellectuelle sans faille et sa pédagogie poussée.

Je remercie mes tuteurs professionnels, Frédéric Ronseaux et Milan Billoin, qui m'ont chacun accompagné d'une manière bienveillante tout au long de mon stage. Je tiens particulièrement à remercier Milan pour sa pédagogie, son écoute et sa sympathie. Je tiens ensuite à remercier les personnes constituant l'équipe rapprochée décret tertiaire, qui ont été les mentors officiels de mon stage. Je remercie ainsi William Teyssier pour la confiance qu'il a su m'accorder au fur et à mesure de mon stage. Je remercie Jack Suddaby pour son soutien indéfectible, son humilité précieuse et son esprit de synthèse hors pair. Je remercie Antoine Hubert pour son accessibilité, sa patience et l'ensemble de ses conseils avisés. Je remercie mes deux comparses alternants, Valentin Khamly et Antonin Desplanques, pour leurs nombreux coups de mains salvateurs, ainsi que la richesse et la drôlerie de tous nos échanges. Je remercie également mes deux collègues stagiaires Yasmina Nina et Noé Godailler, qui ont faits l'impossible, grâce à leur détermination sans faille et leur organisation minutieuse. Je tiens finalement à remercier l'ensemble de mes collaborateurs AREP, pour la richesse de nos échanges et la bonne entente générale. Je remercie particulièrement l'équipe CVC de m'avoir accueilli aussi chaleureusement.

L'organisme d'accueil

L'entreprise

L'agence d'architecture pluridisciplinaire AREP (Architecture Recherche Engagement Post-Carbone) filiale de SNCF Gares & Connexions (SNCF G&C), se consacre à l'invention d'un avenir post-carbone. Première agence d'architecture de France, elle est aujourd'hui représentée à l'international par ses plus de mille collaborateurs de trente nationalités différentes. Sous la direction de Raphaël Ménard depuis 2018, AREP s'engage activement dans la transition écologique du monde ferroviaire et architectural, notamment grâce à son référentiel *Energie Carbone Climat Biodiversité* (EMC2B). L'innovation et la recherche sont mises au profit de solutions « élégantes et sobres, efficaces et pérennes » (AREP, s.d.).

L'agence AREP possède huit domaines d'expertise : Architecture, Environnement et numérique, Conseil et Programmation, Territoires, Management de projet AMO, Design, Ingénierie, et Flux et Mobilité (AREP, 2023). Le domaine de l'ingénierie regroupe plusieurs sujets, tels que la structure, le confort climatique (CVC), l'électricité, la recherche, l'économie, l'acoustique, l'accessibilité, la sécurité incendie et la coordination de projet. Chaque sujet constitue un pôle.

Le stage

Rattachée aux équipes du pôle CVC, j'ai pu bénéficier de leur expertise sur la dépense énergétique du patrimoine de SNCF G&C, notamment au travers des retours de terrain ou d'audits énergétique. J'ai également collaboré avec d'autres équipes. L'équipe du pôle recherche, aussi appelé Hypercube, a partagé son expérience dans la création de modèles numériques complexes et la simulation numérique. J'ai également été amenée à échanger avec les équipes d'Environnement et Numérique. Ces échanges m'ont notamment aidé à retranscrire et pérenniser mes travaux de recherches sous Python.

Présentation de la mission

Contexte de l'affaire

Contexte national

Le 1^{er} septembre 2019 entre en vigueur l'obligation Eco Energie Tertiaire (EET). Cette obligation, élaborée par le gouvernement français, vise à réduire, comparativement à une année de référence, la consommation en énergie finale du parc immobilier tertiaire français (Ministère de la Transition énergétique, 2023). Les bâtiments tertiaires cumulant une surface plancher supérieure à 1000 m² sont concernés par EET. Les objectifs de réduction, illustrés figure 1, sont déterminés selon les performances énergétiques initiales des bâtiments. Le suivi de réduction des consommations s'opère par un reporting annuel et obligatoire, effectué sur une plateforme dédiée.

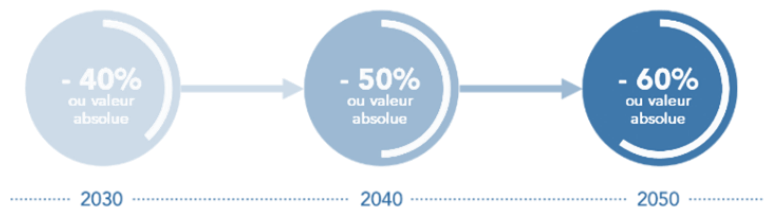


Figure 1 : objectifs de performance énergétique à échéances 2030, 2040 et 2050 (2022)

Contexte entreprise

Le patrimoine de SNCF G&C est concerné par l'obligation EET. Ainsi, sur les 3000 gares recensées en France métropolitaine, 340 sont concernées pour tout ou partie par l'obligation.

Une étude sur le patrimoine concerné a été faite en amont, afin de déterminer quelle stratégie de réduction adopter (voir fig. 2). Au vu des caractéristiques constructives et énergétiques des bâtiments, la stratégie de réduction relative des consommations est retenue.

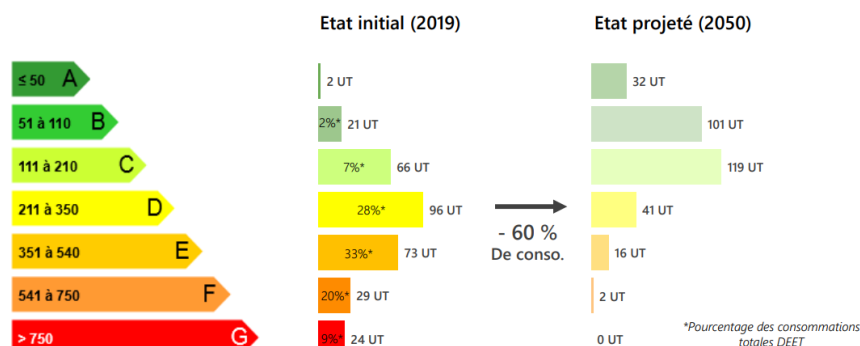


Figure 2 : performances énergétiques initiales et objectifs horizon 2050 (A. Goupil, 2022)

Contexte de la mission

Les bâtiments EET identifiés, il s'agit ensuite de déterminer l'ensemble des actions à mettre en place afin de respecter les objectifs de réduction des consommations.

37 actions individuelles ont été recensées par les équipes AREP comme potentiels leviers d'économies d'énergie. Afin de déterminer les actions à mettre prioritairement en place au sein des gares, il s'agit de déterminer les économies d'énergies potentielles permises par chaque action et combinaison d'actions. Certaines actions possèdent déjà une méthode de calcul propre, mais la plupart des actions n'ont pas de méthode de calcul déterminé.

La mission intervient à un stade déjà avancé de l'offre, et se chevauche avec d'autres phases qui dépendent des résultats obtenus, comme démontré par la figure 3. Le temps de résolution de la mission est donc fortement restreint, et les différentes étapes de résolution doivent concorder avec les besoins des autres phases.

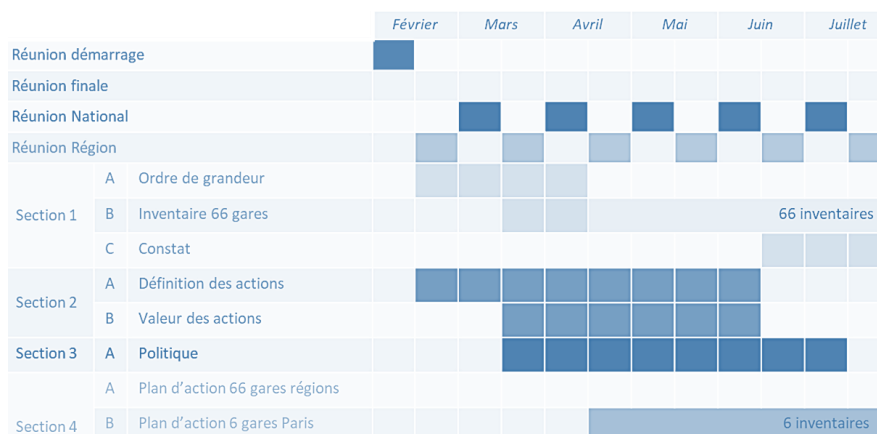


Figure 3 : Organisation de la phase 3 de la mission décret tertiaire (W.Teyssier, 2022)

Description et enjeux de la mission

La mission consiste à fournir, pour chaque action puis combinaison d'actions, des ordres de grandeur des potentielles économies d'énergie. Les ordres de grandeurs obtenus doivent être adaptés aux spécificités des gares étudiées.

L'enjeu de la mission est donc de mettre en place un modèle prédictif d'économies d'énergies, adapté au patrimoine de SNCF G&C, qui soit conjointement fiable, rapide et simple à mettre en œuvre.

Résolution de la mission

Etat de l'art

Diverses approches visent à identifier les économies d'énergies potentielles d'un ou plusieurs systèmes énergétiques.

Approche réglementaire

L'approche réglementaire consiste en la modélisation énergétique fine d'un bâtiment, au travers de logiciels dédiés. Le logiciel réglementaire utilisé par l'équipe CVC est ClimAudit. Ce logiciel présente différents modes de calculs, selon les besoins utilisateur et l'année de construction des bâtiments (BBS logiciels, 2022).

Dans le cadre des audits énergétiques réalisés sur le patrimoine ferroviaire, l'équipe CVC utilise principalement le module RT- Existant, basé sur le calcul réglementaire Th-C-E-Ex. L'interface dédiée à ce module est présentée figure 4. Une fois le bâtiment décrit énergétiquement par l'utilisateur, un premier calcul est lancé afin de déterminer la répartition des consommations par poste. Il est ensuite possible de construire des *variantes* du système énergétique décrit. Les actions ou bouquet d'actions mises en place dans les *variantes* sont renseignées manuellement par l'utilisateur. Un calcul des nouvelles répartitions des consommations est lancé sur les *variantes*.

Cette approche permet une description précise et complète des effets des actions et bouquets d'actions sur les consommations d'énergie. Elle requiert cependant une description très fine des systèmes analysés. La quantité d'informations requises et la vitesse d'exécution des calculs rendent cependant la démarche fastidieuse à des échelles plus complexes tels qu'un groupement de bâtiment ou un parc (Rogea, 2020). Dans le cadre de notre étude, l'usage de cette méthode semble donc compromis.

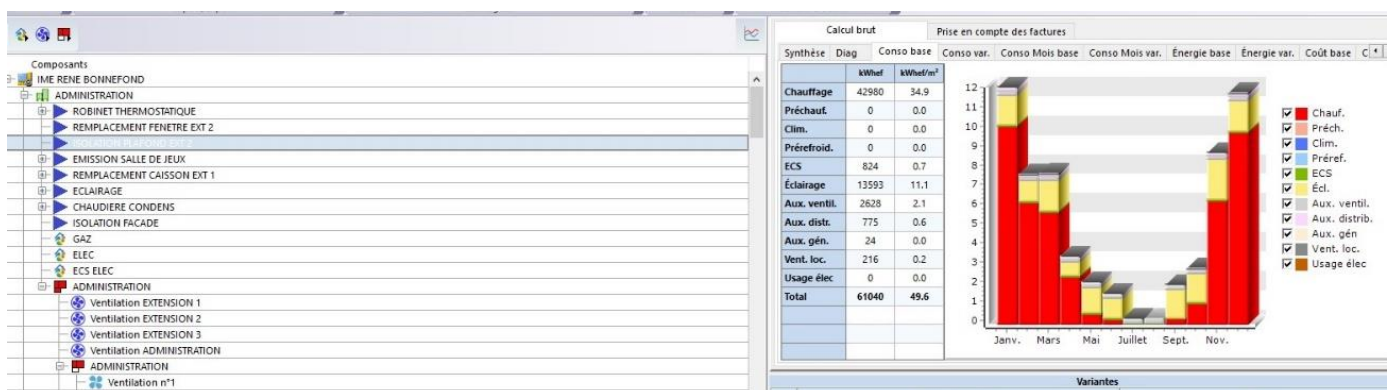


Figure 4 : interface du logiciel ClimAudit dans le cadre des calculs d'audits (BBS logiciels, s.d.)

Approche statistique

Rogean (2020) et Berthou (2013) mentionnent dans leurs thèses respectives l'approche statistique de type régression multilinéaire comme méthode de modélisation de la demande énergétique de bâtiments. Cette approche est préférée aux approches réglementaires pour sa simplicité d'élaboration et sa compréhensivité (Rogean, 2020). Des modèles de régressions linéaires de type Multiple Input Single Output (MISO) sont élaborées par Berthou (2013) dans le cadre de la prédiction des puissances appelées et des températures effectives, à la fois pour le chauffage et la climatisation. L'auteur conclut que ces modèles conviennent notamment « en chauffage et sur les bâtiments anciens », avec des erreurs de prédiction ramenée en énergie « relativement faibles » (Berthou, 2013). Ces modèles se révèlent moins performants concernant la prédiction des consommations de refroidissement, avec des erreurs qui restent cependant « acceptables ». Rogean (2020) souligne par ailleurs la linéarité des « caractéristiques impactées par la rénovation thermique » lorsque des modèles thermiques simplifiés sont utilisés pour quantifier la demande énergétique des bâtiments. L'auteur propose finalement un « métamodèle » semi-multilinéaire présenté figure 5, intégrant des poids explicatifs déterminés par régression, afin de déterminer les besoins annuels de chauffage d'un bâtiment (Rogean, 2020). L'application de la méthode statistique au calcul d'économies d'énergie sans passer par la reconstruction des consommations semble cependant peu explorée à ce jour.

$$Q_b = \left(\sum_{p \in P} \beta_p DHU_p u_{b,p} A_{b,p} + \beta_{vent} DHU \frac{c_{air}}{3.6} \tau_b^{vent} V_b \right) - \beta_{gain} IHU_{b,vitres} g_{b,vitres}$$

β_{murs}	β_{toit}	β_{vitres}	β_{sol}	β_{vent}	β_{gain}
0.779	0.873	1.527	0.360	0.641	0.606

Figure 5 : Formule finale et poids statistiques utilisés pour le calcul des besoins de chauffage (A.Rogean,2020)

Approche forfaitaire

Depuis 2014, le gouvernement incite l'ensemble des secteurs à améliorer leurs performances énergétiques, notamment au travers du dispositif *Certificats d'Économies d'Énergie* (CEE) (Ministère Écologie Énergie Territoires, 2023). Ces certificats présentent, pour chaque secteur, les montants forfaitaires d'économie d'énergies associées à la mise en œuvre d'une action. Les économies d'énergie sont présentées en kWhcumac. Cette unité comprend l'ensemble des kWh économisés sur la durée de vie des installations relatives à chaque action. Les forfaits de chaque action dépendent de plusieurs critères. Pour les fiches relatives aux bâtiments tertiaires, on observe notamment la redondance des critères du secteur d'activité, et de la rigueur du climat comme illustré figure 6.

5. Montant de certificats en kWh cumac

Montant en kWh cumac par m² d'isolant		
Zone climatique	Énergie de chauffage	
	Électricité	Combustible
H1	3 000	4 800
H2	2 500	3 900
H3	1 600	2 600

Secteur d'activité	Facteur correctif
Bureaux, Enseignement, Commerces	0,6
Hôtellerie - Restauration	0,7
Santé	1,3
Autres secteurs	0,6

Surface d'isolant en m²	
S	

Figure 6 : valeurs forfaitaires en kWhcumac pour l'action d'isolation des murs extérieurs par l'intérieur (Ministère Écologie Énergie Territoires, 2023)

Approche terrain

Certains acteurs de la construction évoquent la nécessité d'une approche agile face aux enjeux de baisse des consommations des bâtiments tertiaire. La société Incub' (Incub, s.d.) promeut ainsi des solutions basées uniquement sur une visite poussée des bâtiments et des échanges réguliers avec les usagers.

Conclusion

Les différentes solutions abordées dans cette partie sont résumées dans le tableau 1.

Au vu des enjeux de la mission, l'approches statistique semble la plus adaptée. Cette approche requière des bases de données d'apprentissage conséquentes afin de pouvoir générer un modèle prédictif fiable.

APPROCHE	CALCUL DE LA CONSOMMATION	CALCUL DES ECONOMIES D'ENERGIE	AVANTAGES	LIMITES
Réglementaire	RT existant (Audit)	Construction d'une <i>variante</i> puis recalcul	-Modélisation fine	-Quantité conséquente de données d'entrées -Complexité d'application
Statistique	Régressions		-Simplicité d'application -Rapidité des calculs	-Encore peu développée -Modèles parfois décrits comme <i>boîtes noires</i>
Forfaitaire		Ratios selon critères	- Simplicité des calculs	-Peu précis à l'échelle bâtiment -Combinaison d'actions impossible
Terrain			-Très peu coûteux -Mise en œuvre rapide	-Peu généralisable à plusieurs bâtiments -Nécessite une visite et un suivi poussés

Tableau 1 : Tableau récapitulatif des diverses méthodes de calcul étudiées (A.S Girot, 2023)

Proposition d'un premier modèle

Construction des données d'apprentissage

Afin de générer un modèle prédictif fiable, il est nécessaire d'avoir des données d'apprentissage conséquentes qui puissent représenter le patrimoine dans son ensemble. L'enjeu de notre modèle étant de prédire les gains permis par des actions spécifiques ou bouquet d'actions, la base de données doit comprendre des bâtiments aux caractéristiques variées, ainsi que leurs consommations avant et après mise en place des actions ou bouquets d'actions. Pour cette raison, il n'est pas choisi de partir de bases de données existantes, mais de générer des données d'apprentissage.

Après discussion avec les équipes, il est décidé de générer de la donnée à partir des modèles énergétiques des gares précédemment auditées. Ce choix est fait afin d'avoir des résultats adaptés aux spécificités énergétiques des gares. Les gares étant modélisés énergétiquement majoritairement sous ClimAudit, il m'est donc imposé de travailler avec ce logiciel pour construire les données d'apprentissage. Ce logiciel ne possédant pas d'API, le nombre de simulations à effectuer pour générer de la donnée d'apprentissage est un facteur limitant. Il est donc décidé de générer des données d'apprentissages à partir d'une gare de référence et non plusieurs gares. La gare de référence est ensuite changée incrémentalement en fonction des besoins de l'étude.

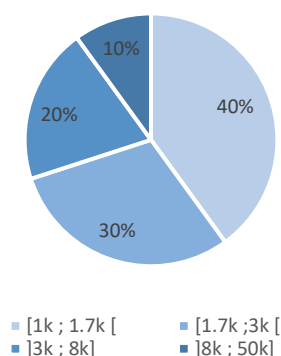


Figure 7 : Répartition du nombre de gares par classe de surfaces (m²) (A. Goupil, 2022)

Parmi les 32 gares auditées, 13 répondent aux critères de surface imposé par EET. Par ailleurs, hors grandes gares parisiennes, 70% des gares EET sont inférieures à 3000 m² (ref. figure 7). Il est donc décidé de choisir une gare appartenant à la classe [1.7k - 3k] afin d'avoir un système fittant au mieux aux petites gares mais aussi aux grandes gares. Cela réduit notre échantillon à 4 gares. La gare de Valence Ville est choisie en sa qualité de majorante de notre classe.

Il s'agit ensuite d'identifier, pour chaque action, les critères à relever par itération qui serviront ensuite au modèle. Le logiciel ne possédant pas d'API, l'ensemble des changements, lancements des simulation et reports des résultats doivent être effectués manuellement. Les critères et valeurs par critères doivent donc être limités afin de minimiser le nombre de simulations à lancer sur ClimAudit. Les critères et valeurs par critère sont définies selon la littérature et les expertises internes AREP. Le tableau 2 présente un exemple pour l'action d'isolation des planchers hauts.

Nom des critères	Gestion température local inférieur	Local supérieur	Epaisseur du plancher (cm)	Zone climatique	Altitude
Valeurs possibles	{Chaud ; Chaud Froid}	{Combles ; Ext}	{10 ; 15 ; 20}	{H1 ; H2 ; H3}	{[0,400[; [400,800] ; 800+}

Tableau 2 : critères et valeurs des critères pour l'action d'isolation des planchers hauts (AS. Girot, 2023)

L'ensemble des critères et valeurs par critère définissent, pour chaque action, un nombre de simulations à effectuer, selon l'équation 1 :

$$n_{simulation} = \prod_{Ei \in \{critères\ action\}} card(Ei)$$

Équation 1 : nombre de simulations par action selon le cardinal de chaque critère (AS. Girot, 2023)

Une fois l'ensemble des valeurs déterminées, les simulations sont lancées au travers de l'outil variante de ClimAudit. Les consommations totales avant et après action sont relevées. L'ensemble est ensuite restitué dans un Excel. Les gains sont ensuite ramenés par unité fonctionnelle afin de garantir la transposabilité des résultats

Régressions

Une fois les simulations effectuées, il s'agit de proposer un premier modèle prédictif, identifiant les critères de sensibilité principaux. Il est choisi d'opter pour un modèle *boîte noire* de type régression multilinéaire. Ce modèle est choisi pour sa simplicité d'application et de compréhension, ainsi que pour sa rapidité. Ce modèle ne peut cependant pas être appliqué sans traitement préalable pour notre étude. En effet, plusieurs de nos actions contiennent des critères qualitatifs, ce qui ne convient pas pour une régression classique. Des analyses préalables sont donc menées.

Des premières sorties graphiques sont faites afin d'avoir une idée des critères qualitatifs déterminants. Les critères visibles à l'œil nu comme non pertinents sont éliminés.

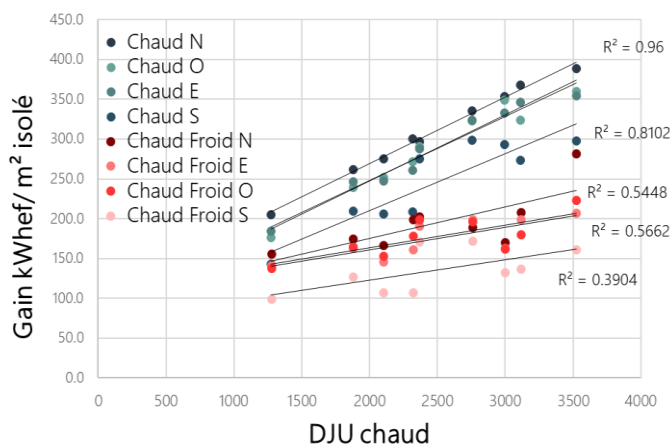


Figure 8 : Régressions selon DJU chauffage pour l'action pose de vitrages performants (AS. Girot, 2023)

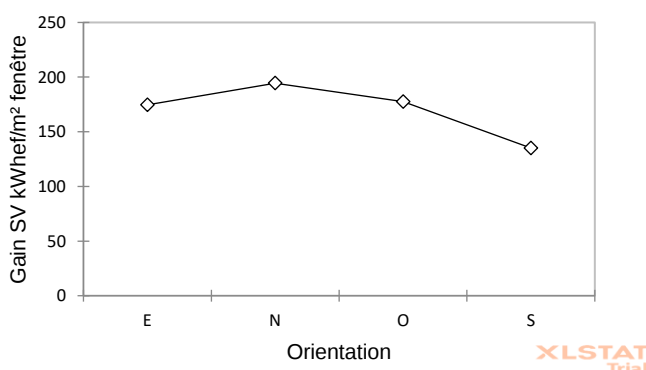


Figure 9 : Moyennes des gains par sous-groupe d'orientation, action pose de vitrages performants (AS. Girot, 2023)

Concernant les critères d'orientation et d'altitude, ils semblent influents pour la majorité des actions. Une tentative de linéarisation à travers la prise en compte du DJU chauffage est menée, telle que sur la figure 8. Cette linéarisation se révèle concluante pour la majorité des simulations. Il est donc décidé de regrouper les critères d'altitude et de zone climatique en prenant en compte le DJU chauffage.

Les critères ou valeurs de critères qualitatifs restants qui sont potentiellement redondants sont finalement identifiés à travers une analyse statistique de type ANCOVA menée sous XLSTAT. Les moyennes par sous-groupes qualitatifs sont analysées, et la pertinence de certains éléments de l'ensemble est validée. Pour l'exemple figure 9, les moyennes des orientations E et O sont identiques. Ces orientations sont donc fusionnées.

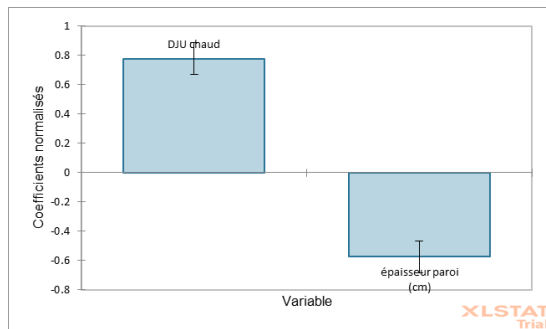


Figure 10 : Coefficients et incertitudes normalisés, régression pour l'action d'isolation des planchers hauts, sous-groupe local chauffé donnant sur combles (A.S. Girot, 2023)

Une fois ces simplifications faites, il s'agit d'effectuer des régressions multilinéaires selon chaque sous-groupe incompressible. Une vérification du coefficient de détermination donne une première idée de la capacité de prédiction du modèle. Les coefficients de la régression sont normalisés et analysés graphiquement par le logiciel (voir fig. 10). Au cas où l'incertitude normalisée est non négligeable par rapport au coefficient normalisé, la ou les variables concernées ne sont pas retenues. Une nouvelle régression est donc lancée sans ses variables. Si l'ensemble des

variables sont jugées incertaines, la régression est abandonnée et un ratio moyen est appliqué. D'autre part, si le R^2 est jugé insatisfaisant et que l'ensemble des variables identifiées sont retenues, une nouvelle régression est lancée en incluant une nouvelle variable.

La figure 11 illustre l'ensemble des étapes citées précédemment ainsi que la formule et les coefficients retenus pour l'action isolation des planchers hauts. Ainsi, pour certains sous-groupes, l'erreur sur le coefficient normalisé du DJU chauffage s'est révélé trop importante. Cette variable a donc été abandonnée, et l'inclusion du DJU refroidissement s'est révélé concluant.

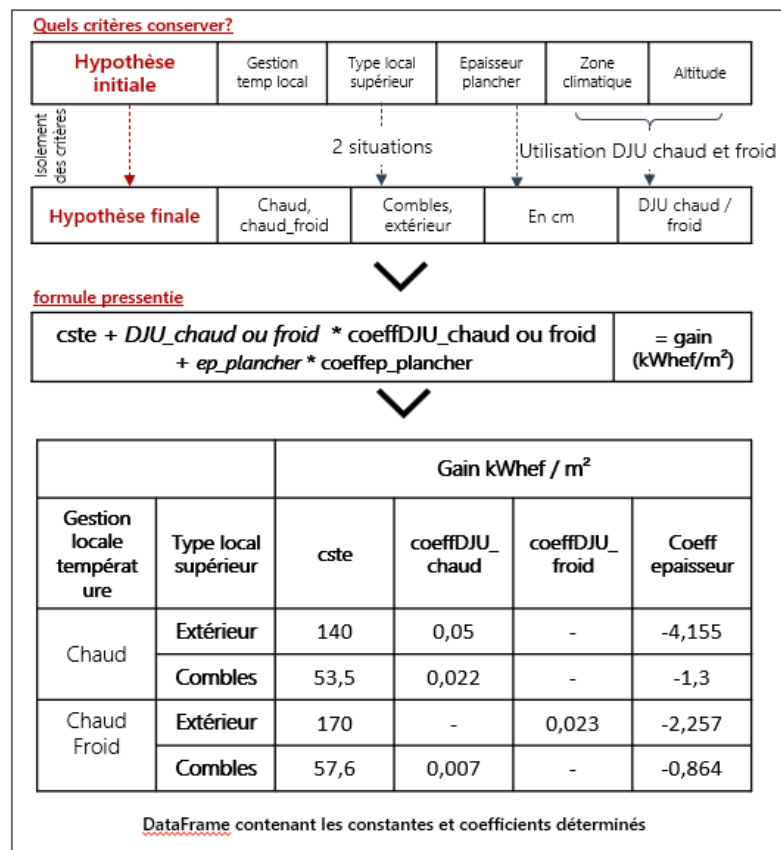


Figure 11 : Premier modèle avec coefficients pour l'action d'isolation des planchers hauts (A.S. Girot, 2023)

Test du modèle sur des données de validation et stratégies d'amélioration

Par unité fonctionnelle

Une fois notre premier modèle mis en place, il s'agit de le tester sur des données de validation. Cette étape est notamment incontournable pour assurer la cohérence des résultats obtenus lors de l'application du modèle sur l'ensemble du patrimoine, et ainsi éviter l'*overfitting* (Babyak, 2004). D'autres gares auditées, les plus différentes de la gare d'apprentissage, sont ainsi choisies. Le protocole de construction des données est réappliqué sur ces gares. Les données obtenues sont ensuite comparées aux résultats calculés par notre modèle (voir fig. 12).

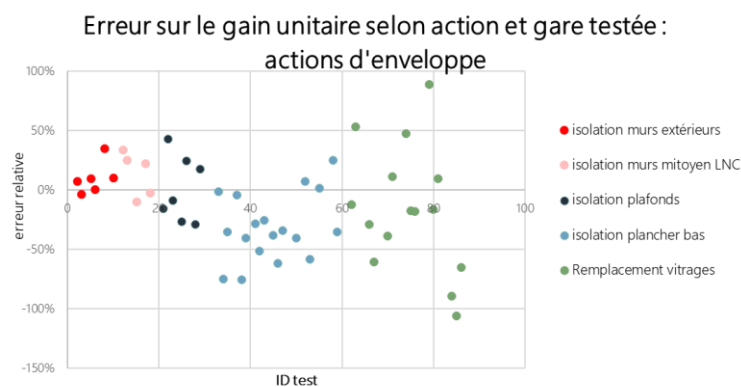


Figure 12 : répartition des erreurs relatives (A.S. Girot, 2023)

La répartition des erreurs est ensuite analysée action par action. En cas de dispersion forte des erreurs, il est envisageable que le modèle passe à côté de variables prédictives exerçant une forte influence sur les résultats. Des variables sont donc ajoutées, en suivant des lois de thermique simples. Les formules obtenues se rapprochent ainsi plus de formules de type *boîte grise* (Berthou, 2013). En cas d'erreurs tendancielle de surestimation ou sous-estimation, une simple recalibrage des coefficients utilisés est envisagée. Le recalibrage se fait par l'ajout d'un coefficient multiplicateur m tel que défini en équation 2 :

$$\frac{1}{n} * \sum_{i \in \text{test}} 1 - m * \frac{\text{gain}_{\text{calc},i}}{\text{gain}_{\text{mes},i}} \xrightarrow{m_{\text{opti}}} 0, m \in]0, +\infty[$$

Équation 2 : détermination du coefficient multiplicateur optimal (A.S. Girot, 2023)

Cette démarche est choisie car elle ne nécessite pas de recalcul via XLSTAT, et est donc plus simple et rapide à mettre en œuvre. Dans l'idéal, les calculs de régressions doivent être relancés jusqu'à convergence des erreurs vers 0.

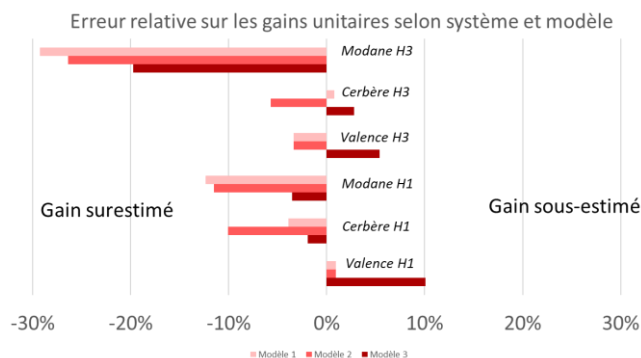


Figure 13 : évolution de l'erreur relative selon corrections apportées (A.S. Girot, 2023)

Les stratégies d'enrichissement et de recalibrage peuvent être employées pour une même action. C'est le cas pour l'action d'isolation des planchers hauts, illustrée figure 13. Ainsi, l'inclusion de nouvelles variables a rendu l'erreur tendancielle (Modèle 2). Une révision des coefficients est donc effectuée (Modèle 3).

En cas d'échec des stratégies d'enrichissement et de recalibrage, l'approche statistique est totalement abandonnée. Le gain est recalculé en se basant uniquement sur des formules thermiques et physiques évoquées dans la littérature. Cette stratégie est adoptée pour l'ensemble des actions concernant le changement de génération. Ainsi, en supposant les besoins de chauffage Bch constants, on peut définir le gain G associé à ces changements de génération tel qu'exposé dans l'équation 3 :

$$G = \frac{Bch}{\rho_{init}} - \frac{Bch}{\rho_f}$$

Équation 3 : Gains selon les besoins de chauffage et rendements des systèmes de chauffage initial et final (A.S. Girot, 2023)

Etant donné que les besoins de chauffage ne sont pas estimés par notre modèle, ils sont donc inconnus. Ils doivent alors être exprimés à travers des variables qui soient connues par la visite de terrain. Nous choisissons ainsi d'exprimer Bch selon la consommation initiale de chauffage C_{init} , disponible par facture pour les chaudières gaz et fioul, et le rendement ρ_{init} de ces chaudières, suivant l'équation 4:

$$G = \frac{C_{init} * \rho_{init}}{\rho_{init}} - \frac{C_{init} * \rho_{init}}{\rho_f} = C_{init} * (1 - \frac{\rho_{init}}{\rho_f})$$

Équation 4 : Gains selon consommation initiale et rendements (A.S. Girot, 2023)

Cette approche par le calcul ne permet cependant pas d'exprimer le gain de remplacement pour des systèmes de chauffage de type électrique, car cela supposerait d'avoir le détail des postes de consommation dans les données de consommation électrique. Cette précision n'étant pas disponible à ce jour, la formule, ne se limite donc qu'aux changements des générations usant du gaz ou du fioul. Cela suppose également que ces consommations ne sont allouées qu'au poste de chauffage.

Ajout des quantités

Les unités fonctionnelles sont couplées à des quantités afin d'obtenir des gains s'exprimant en MWh d'énergie finale (MWh_{ef}). Afin d'avoir une idée de la précision des résultats de notre modèle par rapport aux résultats des audits, mais aussi par rapport à d'autres approches, nous décidons par ailleurs d'adapter la méthode CEE de l'ADEME afin de calculer les gains de certaines actions. Les comparaisons pour les gares de Modane et Sélestat sont disponibles figure 14. Cette comparaison révèle un écart significatif du gain espéré par le changement de génération par une

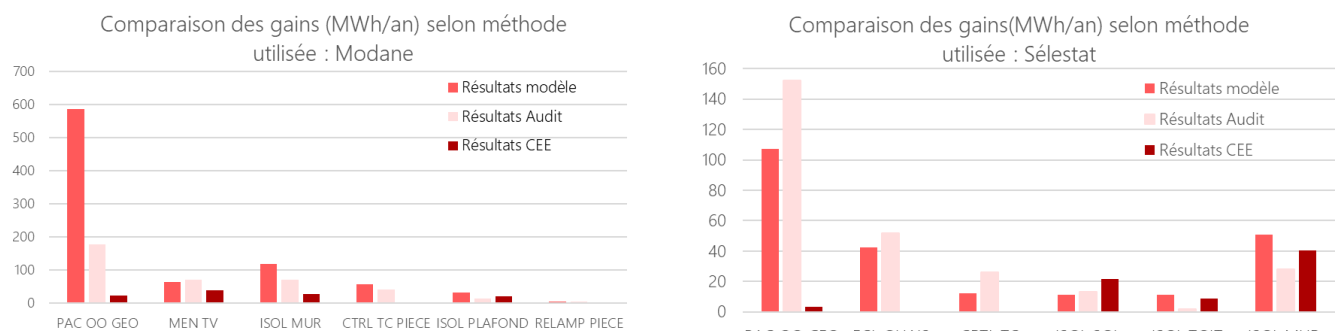


Figure 14: Analyse comparée des gains obtenus par action, selon méthode utilisée, pour deux gares auditées (A.S. Girot, 2023)

pompe à chaleur (PAC) eau-eau. Cela peut notamment s'expliquer par les hypothèses prises sur les rendements initiaux et coefficients de performance finaux des systèmes de chauffage. Dans le cas de notre modèle, l'erreur peut aussi s'expliquer par une erreur sur les consommations initiales. En effet, notre modèle utilise les données renseignées par DEEPI et la SNCF, tandis que les audits reposent sur l'analyse directe des factures. Concernant les autres actions, les analyses ne révèlent pas d'erreurs systémiques. Le modèle est donc validé. Ce modèle décrit l'ensemble des critères nécessaires et suffisants à relever sur le terrain pour les actions concernées.

Test du modèle sur des données terrain

Principe

L'ensemble des formules de notre modèle et l'ensemble des formules identifiées précédemment sont retranscrites au sein d'un même code Python. Le codage facilite ainsi l'analyse automatique des résultats. Par ailleurs, un document intermédiaire Excel est rédigé afin de pouvoir stocker rigoureusement l'ensemble des informations relevées lors des visites des gares EET. Ce document Excel est ensuite directement lu par le Python.

Construction de la base de données DJU

Les calculs précédents étant réalisés sur des gares auditées, les DJU utilisés sont ceux renseignés dans la base de données météo de ClimAudit. Cette base de données n'est cependant pas publique. Il est donc nécessaire de passer par une autre base de données. Le site Infoclimat propose un calcul des DJU chauffage, mais la donnée reste souvent lacunaire et l'export automatique semble impossible. Le Ministère de la Transition Ecologique et de la Cohésion des Territoires (s. d.) propose une base de données des DJU chauffage, mais dont la granularité des résultats ne permet pas de tenir compte précisément de l'altitude des gares. Il est donc décidé de créer une base de données à partir des méthodes de calcul des DJU renseignés par MétéoFrance et GRDF (2023). L'hypercube étant fort de précédentes recherches sur le sujet, les méthodes donnent des résultats équivalents (A.Hubert, Comm. Pers., 2023). La méthode de calcul dite « météo » est donc préférée pour sa simplicité. La base de données est ensuite construite sous Python, à partir des coordonnées géographiques des gares et des données météo opensource de Meteostat. Le code utilisé est présenté en annexe 1.

Analyse graphique

La base de données des DJU faite, le code est donc lancé sur plusieurs gares. Les gains absolus sont calculés. Ces gains sont rapportés à la consommation. Les gains absolus et relatifs sont reportés sur une heatmap afin de faciliter la lecture graphique.

Une première analyse est faite sur les gains relatifs, présentés figure 15. Ainsi, l'ordre de priorité des actions individuelles colle avec la littérature, ainsi que les résultats des gares auditées. On retrouve ainsi la pose de PAC et les travaux de rénovation des parois opaques et vitrées comme actions à impact significatif. Les actions de régulation des consignes de chauffage et de climatisation hors hall sont également significatives, tout type de gare confondues. L'action de suppression de chauffage du hall se révèle peu intéressante à l'échelle du patrimoine, au regard

du faible nombre de gares analysées ayant un hall chauffé. Une exception notable subsiste pour Strasbourg, dont un des halls chauffés consiste en une verrière donnant sur l'extérieur. Les actions sur les équipements semblent peu significatives au regard des consommations des gares. On note cependant un gain significatif pour l'arrêt nocturne des écrans pub dans certaines grandes gares de province, telles que Strasbourg et Caen.

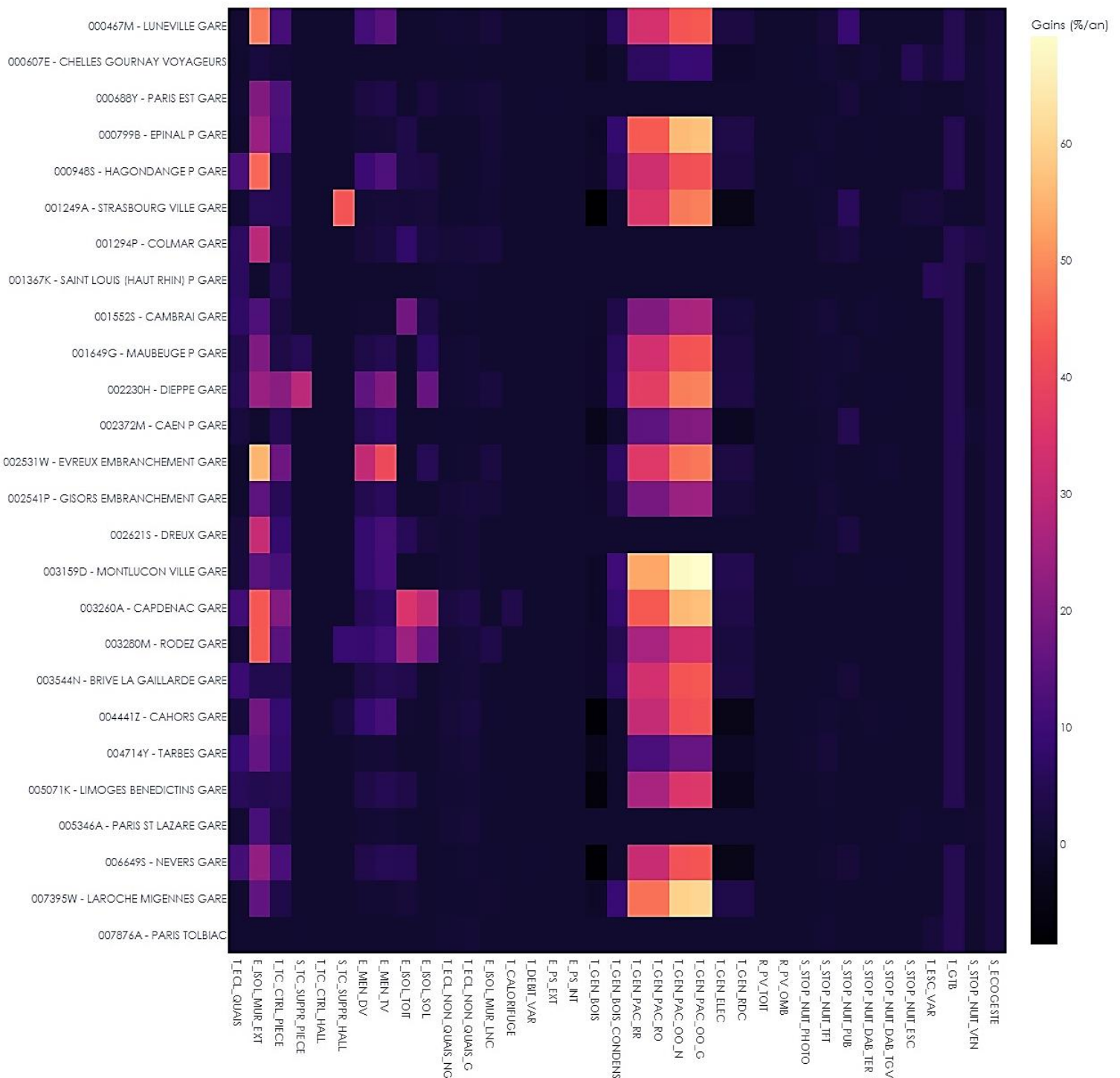


Figure 15 : Heatmap des gains rapportés à la consommation, par gare et par action testées (A.S. Giro, 2023)

Comparativement aux résultats obtenus par les audits précédemment réalisés, certaines différences subsistent. C'est notamment le cas vis-à-vis de l'action de rénovation de l'éclairage des quais. Les gains calculés par notre modèle se révèlent ainsi peu significatifs au regards des résultats précédents sur les gares auditées. Cela peut s'expliquer partiellement par la différence de ratios de surfaces entre les gares auditées, majoritairement non EET, et les gares EET. On constate ainsi que la proportion de surface de quais rapportée à la surface totale diminue lorsque la gare augmente en taille et/ou en fréquentation. Une autre différence tient en la valeur moyenne des gains observés pour l'isolation des parois opaques. Cela s'explique partiellement par des surfaces à isoler potentiellement surévaluée. Même en excluant ces cas, le gain moyen relatif à l'isolation des murs extérieurs reste deux fois plus élevé que la valeur trouvée dans les audits. Cela traduit une erreur systématique potentiellement due à un mauvais calibrage de notre modèle de régression. Le gain moyen trouvé reste cependant cohérent avec d'autres valeurs évoquées dans la littérature. Enfin, on observe des augmentations de la consommation pour la pose de chaudières biomasses classique, ainsi que pour le raccordement à un réseau de chaleur ou le passage au tout électrique. Cela s'explique par les rendements moyens pris pour ces solutions, qui est désavantageux comparativement aux rendements des chaudières relevées sur le terrain.

Une analyse rapide est ensuite menée sur les gains totaux, toujours à l'aide d'une heatmap présentée figure 16. Les résultats se révèlent largement disparates selon les gares concernées. On constate ainsi que les Grandes Gares Parisiennes (GGP) et les grandes gares de province se retrouvent largement mises en avant. La encore, les actions de pose de PAC et isolation des parois se révèlent les plus intéressantes, avec une priorité pour les grandes gares, dans et hors Paris. Au vu de la taille des gares concernées, ces actions semblent cependant complexes à mettre en œuvre, ou bien sont déjà mobilisées. Ainsi, les deux GGP analysées se sont déjà raccordées à des réseaux de chaleurs. Quant à Strasbourg, le système de chauffage vient d'être remplacé en 2023 par une nouvelle chaudière gaz à condensation (M. Billoin, Comm. Pers., 2023) ce qui rends la pose de PAC non mobilisable avant au moins 10 à 20 ans. Il est probable que des situations similaires soient observées sur les autres grandes gares de province. Les actions de sobriété ressortent cependant beaucoup plus avec cette analyse, en particulier sur les grandes gares. Ainsi, le contrôle de la température sur la gare de Paris Est permet d'économiser plus d'énergie que la pose d'une PAC en gare de Limoges. De la même manière, l'arrêt nocturne des écrans pubs sur les gares de Stasbourg et Paris Est permettent d'économiser autant d'énergie que la pose d'une PAC en gare de Caen.

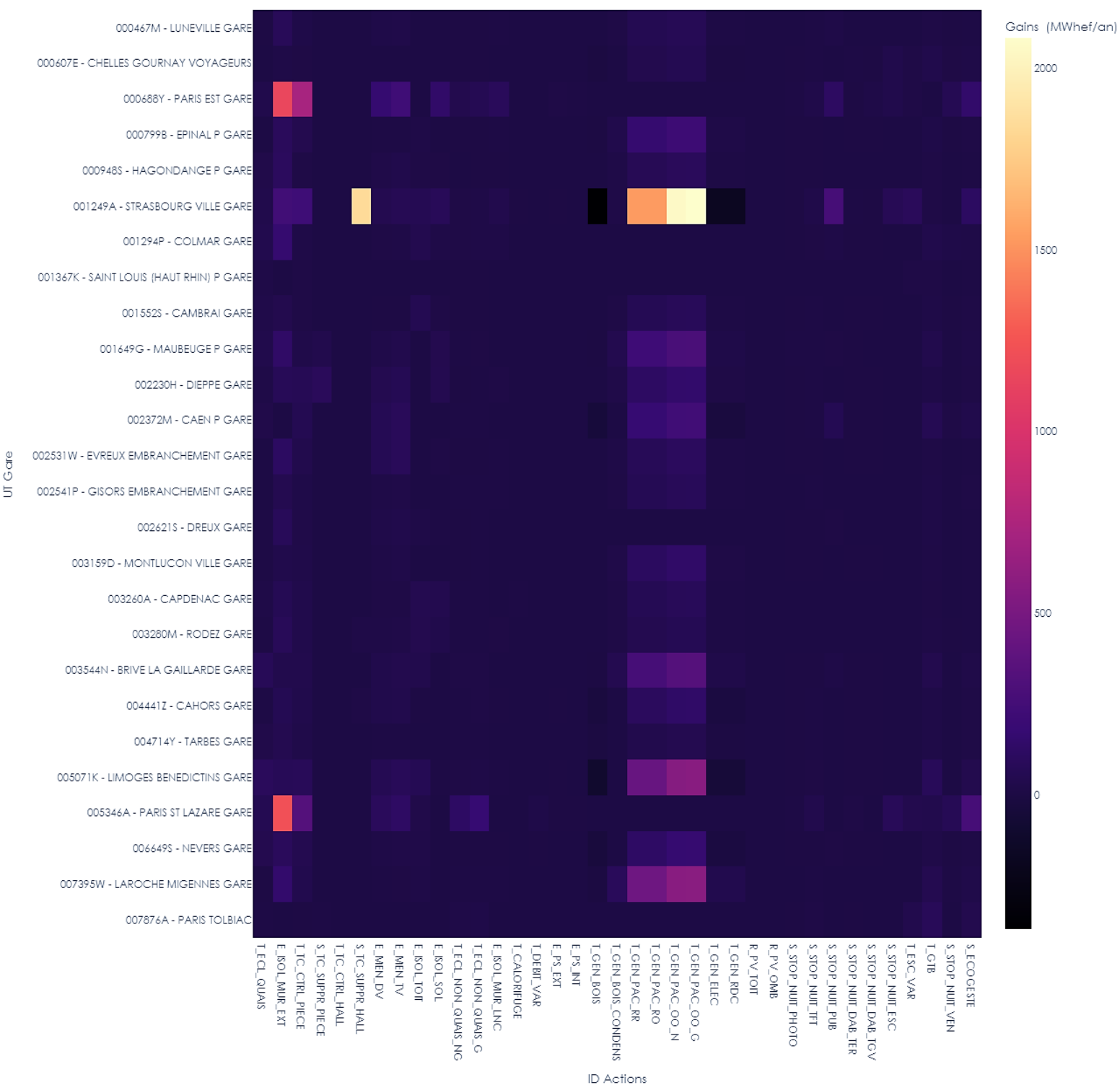


Figure 16 : Heatmap des gains en MWh/ea, par gare et par action testées (A.S. Girot, 2023)

Pistes pour l'élaboration d'une stratégie à l'échelle du patrimoine

Une première tentative de combinaison des actions est menée (voir fig. 5). L'approche choisie est multiplicative, telle que , pour chaque gare j , et pour chaque gain rapporté à la consommation $G\%_{i,j}$:

$$C_{totf,j} = C_{totinit,j} * \prod_{i \in action} (1 - \frac{G\%_{i,j}}{100})$$

Équation 5 : Approche multiplicative adoptée pour le calcul des gains de bouquets d'actions, à scénario le plus favorable (A.S. Girot, 2023)

Dans le cas où plusieurs gains seraient calculés pour un même type d'action, seul les gains majorants sont choisis. Les poses de triple vitrage et d'une PAC eau-eau géothermie sont donc systématiquement retenues pour ce calcul, au détriment des doubles vitrages et autres types de générations.

Les consommations finales à horizon 2050, ainsi que les gains totaux accumulés entre l'année de référence et 2050, sont ensuite regroupés selon le type des gares analysées. On distingue ainsi trois groupes : les GGP, les gares Ile de France hors GGP (IDF), et les gares de province (DOT). L'ensemble est compilé dans un diagramme de Sankey présenté figure 17. Parmi l'ensemble des

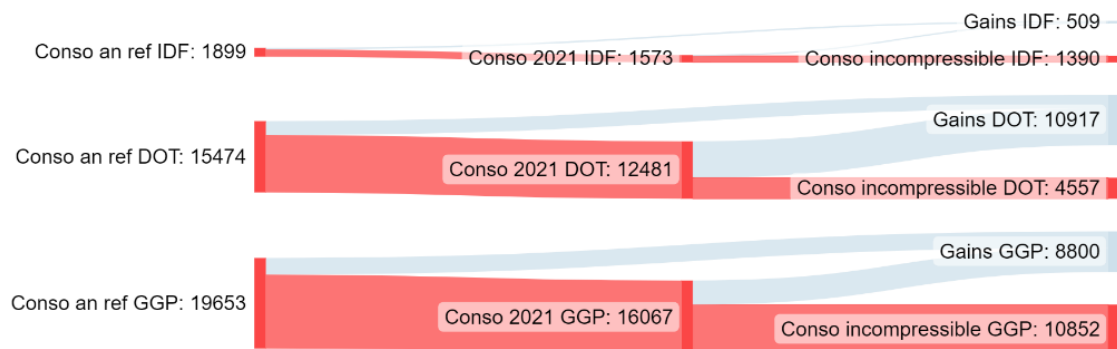


Figure 17 : Diagramme de Sankey, avec consommations initiales, à horizon 2050 et gains totaux, regroupés par typologie des gares étudiées (A.S. Girot, 2023)

gares analysées, notre modèle semble ainsi identifier la majorité des actions à effectuer sur les gares DOT : dans l'hypothèse où l'ensemble des leviers candidats sont activés, la baisse de consommation à horizon 2050 dépasse ainsi les -60% attendus pour ce groupe. Le modèle se révèle moins concluant pour les autres groupes. Concernant les GGP, le modèle peine à identifier les leviers d'économie d'énergie. Cela s'explique en partie par le fait que les gares GGP analysées soient déjà raccordées à un réseau de chaleur, et que ce raccordement ne puisse potentiellement pas être remplacé par la pose de PAC. Une autre explication est que ces gares puissent avoir des sources de consommation spécifiques, notamment du fait d'équipements ferroviaires, bureautiques ou commerciaux non pris en compte par le modèle. Concernant les gares IDF, le diagramme ne tient compte que d'une gare souterraine à ce jour, ce qui biaise potentiellement les résultats présentés.

Ces résultats sont à mettre en regard avec la répartition des consommations totales à l'échelle du patrimoine, selon les groupes et selon la taille des gares. Ainsi, en 2019, Les GGP constituent 32% de la consommation totale du patrimoine EET (A. Goupil, Comm. Pers. ,2022) . Les gares DOT et IDF concernées par EET constituent les 68% restant. En analysant conjointement la consommation et la surface des gares DOT et IDF, 30% des plus grandes gares constituent en 2019 plus de trois quart des consommations de ces groupes (voir figure 18). Ces résultats confirment l'analyse menée à l'aide de la heatmap des gains totaux.

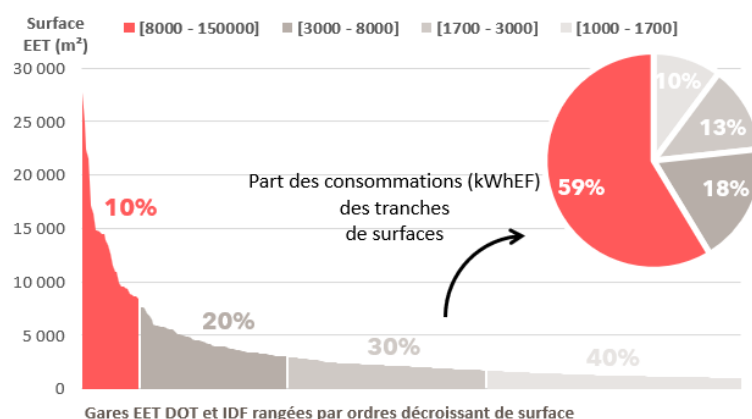


Figure 18 : Répartition des surfaces et consommations par classe de surfaces pour les gares IDF et DOT (A. Goupil, 2022)

Ainsi, les grandes gares, dans et hors Paris, ont un poids non négligeable sur les consommations totales du patrimoine. Les leviers de sobriété et de rénovation appliqués à ces gares doivent donc constituer le cœur de la stratégie énergétique de SNCF G&C. Sans l'activation de ces leviers, l'atteinte des objectifs EET de réduction des consommations à l'échelle du patrimoine est potentiellement compromise.

Analyse critique et pistes d'amélioration

Stratégie d'élaboration du modèle

Le calcul direct des gains, en passant par un modèle statistique semi-physique, sans reconstruction de la consommation, semble à ce jour une méthode inédite. Il a été plusieurs fois signalé de ma part qu'il aurait été plus juste et judicieux de passer par la reconstruction des consommations en premier lieu, avant estimation des gains et montage des stratégies énergétiques. En effet, la non-connaissance de la répartition des consommations par poste représente un risque de surestimer des gains, en entamant un poste de consommation normalement non impacté par l'action menée. Cela complique d'autre part le calcul des gains obtenus par l'activation de plusieurs leviers. Malgré l'expression de mon inquiétude, cette piste a été écartée, principalement pour des raisons de manques de moyens économiques et temporels.

D'autre part, le choix des variables à inclure s'est fait sans véritable connaissance des informations pouvant être relevées sur le terrain, ni recherche approfondie sur les caractéristiques

constructives du patrimoine de SNCF G&C. Cela m'a notamment conduit à estimer les résistances thermiques des parois opaques en tenant uniquement compte de l'épaisseur. Ce choix a été fait car cette variable est plus simple à relever que la résistance thermique des matériaux, qui nécessite la connaissance des matériaux utilisés et la mesure ou l'estimation de leurs conductivités thermiques. La construction de la base de données d'apprentissage s'est donc faite sur la base de conductivités thermiques constantes, correspondant à l'usage des matériaux renseignés par défaut dans ClimAudit. Les épaisseurs à tester pour les parois opaques ont par ailleurs été estimées à partir des données moyennes relevées sur les gares auditées, ou bien à dire d'experts. Les maillages testés s'en sont retrouvés restreints. Cela augmente finalement les risques d'hallucination du modèle lors de l'apparition de valeurs en dehors du maillage. Le modèle peut ainsi présenter des risques d'hallucination pour des bâtis aux caractéristiques constructives très éloignées de la gare de référence choisie pour la construction du modèle. Par exemple, un bâtiment en préfabriqué, avec des parois fines mais des conductivités thermiques adaptées, sera estimé par le modèle comme très peu performant, bien que dans la réalité ses performances soient équivalentes à celles d'un bâtiment possédant d'épais murs en pierre. Cette faiblesse du modèle pourrait finalement être corrigée en tenant compte des matériaux utilisés pour les parois opaques.

Un raisonnement similaire a été appliqué pour les parois vitrées, où les U_w ont été estimés par la typologie de la fenêtre (Simple ou Double vitrage + épaisseur du double vitrage), en supposant un cadre acier, et un rapport châssis/vitrage constant. Après discussion avec les équipes, il s'avère que les hypothèses de U_w finales étaient trop favorables au vu des caractéristiques constructives des châssis de fenêtre observées sur le patrimoine de SNCF G&C. Le modèle a donc été corrigé en conséquence. Cependant, la correction ne tient pas compte de la variabilité constructive des fenêtres. De la même manière que pour les parois opaques, des critères en tenant compte pourraient être inclus afin d'assurer la fiabilité des résultats.

Fiabilité du modèle

Les incertitudes sur la première version du modèle ont pu être maîtrisées par l'analyse du coefficient de détermination et le choix des variables étudiées. Le test du modèle sur les données de validation et les corrections apportées ont cependant inclus des incertitudes complexes à quantifier. Par ailleurs, L'usage du modèle sur des gares EET très éloignées des gares utilisées pour l'apprentissage et la validation du modèle favorise potentiellement les hallucinations. L'ensemble de ces difficultés pose la question de la représentativité des données d'apprentissage. Une solution possible et évidente aurait été d'entraîner notre modèle initial à l'aide d'une base de données plus fournie (Babiyak, 2004), comprenant des données sur des gares de toute taille, ou à défaut des bâtiments tertiaires présentant de fortes similitudes avec les gares. Il m'a cependant été imposé de travailler à partir des modèles ClimAudit de l'équipe CVC. Le logiciel ne possédant pas d'API, la construction de la base de données à partir d'une seule gare a pris un temps conséquent, et la temporalité du projet ne m'a pas permis d'initier le même travail sur d'autres modèles.

Par ailleurs, le couplage du modèle à des quantités relevées sur le terrain et des données de consommation potentiellement lacunaires ou reconstruites induit des sources supplémentaires

d'erreurs. L'application du modèle sur des gares non précédemment auditées rends finalement plus complexe l'identification des potentielles sources d'erreurs. Ce problème peut être minimisé par la double vérification des données de terrain et des données de consommation avant lancement du calcul.

Enfin, les gains calculés par notre modèle sont difficilement comparables au vu des spécificités du patrimoine étudié. Même en restreignant la comparaison aux résultats des précédents audits de l'équipe CVC, cette comparaison reste critiquable. Cela tient du fait de la différence constitutive entre les gares auditées, et les gares EET visitées et analysées par notre modèle. L'appréciation de la fiabilité de notre modèle lors de son application aux gares EET a donc pour l'instant été faite à dire d'experts. Une idée de fiabilisation du modèle pourrait être à minima de borner les gains à l'aide des résultats des audits précédents, et des résultats obtenus par calcul CEE.

Combinaison d'actions

L'approche multiplicative est choisie pour sa simplicité et sa capacité à représenter la dégression progressive des gains lors de la mise en place de plusieurs actions. En effet, certaines actions peuvent avoir des effets de couplage critique, telles que le changement de génération de chauffage couplé à l'isolation des parois opaques et vitrées. Cette approche ne permet cependant pas de rendre compte de l'additivité de certaines actions entre elles.

Une piste d'amélioration du modèle combinatoire est d'allier des approches additive et itérative. L'approche additive serait ainsi utilisée pour combiner les actions indépendantes entre elles. L'approche itérative serait utilisée pour les actions interdépendantes. L'approche itérative consisterait à effectuer des recalculs en actualisant à chaque itération les données de consommation et de rendement. Cette approche est notamment pertinente pour le poste de chauffage car énormément d'actions, et notamment les actions de rénovation d'enveloppe, dépendent dans notre modèle du rendement global du système de chauffage. De la même manière, l'action de dépose des chaudières fossiles dépend des besoins de chauffage, eux même influencés par le niveau d'isolation de l'enveloppe. L'approche itérative suppose cependant de connaître l'ordre de réalisation des actions, car cette approche combinatoire n'est pas commutative contrairement aux approches additive et multiplicative.

Prise en compte des temporalités

Notre modèle fige les gains dans une certaine temporalité. L'évolution du climat, l'apparition de nouveaux postes de consommations ou bien la dégradation des matériaux et systèmes techniques ne sont pas prises en compte. L'évolution du climat pourrait être prise en compte dans notre modèle par la construction de bases de données de DJU tenant compte de l'évolution des températures en fonction des divers scénarios d'émissions de CO₂ décrits par le GIEC (Masson-Delmotte et al., 2021). L'accroissement ou l'apparition de consommations liées au poste de refroidissement est également à prévoir (ADEME, CODA STRATEGIES, 2021). Cet effet pourrait être pris en compte par l'inclusion d'un facteur de risque d'installation. Ce facteur pourrait ainsi rajouter des consommations. Enfin, la dégradation des matériaux et systèmes techniques pourrait être prise en compte, d'une manière similaire à la méthode employée dans les fiches CEE.

Retour d'expérience et perspectives professionnelles

Retour d'expérience

Mon stage au sein du pôle ingénierie chez AREP m'a permis de développer un ensemble varié et solide de compétences techniques. J'ai ainsi renforcé ma capacité à mettre en œuvre avec succès un modèle de machine learning, transformant des concepts abstraits en solutions concrètes et applicables. Cette aptitude a renforcé ma confiance dans l'application pratique des connaissances théoriques et m'a préparé à relever des défis complexes dans le domaine de l'intelligence artificielle. La familiarisation aux méthodes et outils dédiés aux audits énergétiques m'a également fait prendre conscience des enjeux liés à la rénovation énergétique des bâtiments. D'autre part, cette expérience m'a confronté au défi d'innover tout en tenant compte des contraintes financières et temporelles inhérentes au projet. Cet équilibre subtil entre créativité et réalisme m'a fait progresser sur ma capacité à proposer des solutions novatrices tout en garantissant leur faisabilité opérationnelle. Enfin, le management d'équipe a été pour moi une découverte lors de ce stage. Cela m'a permis de développer des compétences en leadership et en coordination, me préparant ainsi à des responsabilités futures dans le monde professionnel.

Ce stage m'a par ailleurs permis de développer un ensemble solide de compétences liées au savoir-être. Être une force de proposition au sein du projet est devenu une seconde nature pour moi. J'ai appris à générer des idées novatrices et à les exprimer de manière convaincante, contribuant ainsi activement à l'évolution des projets. Ma capacité à maîtriser ma frustration a été mise à l'épreuve et a abouti à une meilleure gestion des défis et des périodes de pression. J'ai également développé un sens aigu de l'intégration, en veillant non seulement à mon propre sentiment d'appartenance, mais aussi à celui de mes collègues. Cette qualité a renforcé la cohésion au sein de l'équipe et a favorisé un environnement collaboratif et harmonieux. Mon stage chez AREP m'a ainsi permis de grandir non seulement en tant que professionnelle compétente, mais aussi en tant qu'individu adaptable et axé sur le travail d'équipe.

Mon immersion m'a également éclairée sur certaines spécificités liées au domaine du bâtiment. J'ai pris conscience que ce secteur est déjà fortement encadré par des lois et réglementations telles que la RT (Réglementation Thermique), la RE (Réglementation Environnementale), ainsi que l'EET. La multiplication de ces cadres réglementaires souligne la prise en compte croissante des enjeux liés à ce domaine. J'ai malgré tout constaté lors de mon stage que les réponses à apporter à ces enjeux pouvaient être sources de désaccords entre les divers experts de ce secteur. Ainsi, certaines initiatives de sobriété énergétique restent peu explorées, en raison de limites d'acceptabilité sociale ou de rentabilité financière. Cette prise de conscience m'a amené à réfléchir plus profondément aux compromis nécessaires entre les aspects environnementaux et économiques, et à considérer les moyens de surmonter ces obstacles. Ces observations m'ont incité à considérer la nécessité d'une approche équilibrée et innovante pour concilier les contraintes économiques avec les impératifs de contraction des consommations énergétiques.

Perspectives professionnelles

Mon expérience au sein d'AREP a affiné ou révélé certaines appétences pour des domaines que j'aimerais plus tard retrouver dans mon parcours professionnel. J'ai ainsi pu affiner mes compétences en data science et cela m'a véritablement motivé à donner le meilleur de moi-même. J'aimerais donc pouvoir utiliser ces compétences ultérieurement, notamment dans le cadre de l'analyse de base de données, ainsi que l'élaboration de visuels pertinents et percutants. Ce stage m'a plus largement initié à l'innovation. J'aimerais poursuivre ce parcours par le biais de développement d'outils, en incluant notamment des démarches de machine learning. Au-delà des aspects techniques, ce stage a également affirmé ma volonté de participer à l'élaboration de stratégies environnementales et énergétiques. J'aimerais notamment initier ou participer au développement d'approches transversales, s'inscrivant dans une démarche de sobriété globale.

Ce stage fut également l'occasion de prendre du recul face à mes expériences professionnelles précédentes. J'ai ainsi pris conscience de ma volonté de m'inscrire dans une démarche de décarbonation du secteur du bâtiment, mais aussi du secteur du transport. Les différentes organisations au sein de chaque entreprise ont finalement affirmé mon envie de travailler au sein d'une entreprise à taille humaine.

Bibliographie

- Babiyak, M. A. (2004). What You See May Not Be What You Get : A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models. *Psychosomatic Medicine*, 66(3), 411-421. <https://doi.org/10.1097/01.psy.0000127692.23278.a9>
- Berthou, T. (2013). *Développement de modèles de bâtiment pour la prévision de charge de climatisation et l'élaboration de stratégies d'optimisation énergétique et d'effacement* [Ecole Nationale Supérieure des Mines de Paris]. <https://pastel.hal.science/pastel-00935434>
- Masson-Delmotte, V., Zhai, P., Pirani, A., Péan, C., & Chen, Y. (2021). *Climate Change 2021 : The Physical Science Basis* (6). IPCC. https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_SPM_final.pdf
- Rogean, A. (2020). *Vers une approche intégrée d'aide à la planification énergétique territoriale : Application à la rénovation énergétique des bâtiments* [Université Paris sciences et lettres]. <https://pastel.archives-ouvertes.fr/tel-02969503>

Webographie

ADEME, CODA STRATEGIES. (2021) *La climatisation de confort dans les bâtiments résidentiels et tertiaires, état des Lieux 2020* – ademe.fr. Consulté le 2 Août 2023, à l'adresse : <https://www.actu-environnement.com/media/pdf/news-37991-ademe-etude-climatisation-france-2020.pdf>

AREP. (s. d.). AREP. ADEME-Résilience des territoires. Consulté le 3 juin 2023, à l'adresse <https://wiki.resilience-territoire.ademe.fr/wiki/AREP>

AREP. (2023, 24 juillet). *Nos métiers - AREP*. Consulté le 17 juin 2023, à l'adresse <https://www.arep.fr/nos-metiers/>

BBS Logiciels. (2022, 19 octobre). *Logiciel Audit Énergétique ClimAudit* | BBS Logiciels. Consulté le 20 mai 2023, à l'adresse : <https://www.bbs-logiciels.com/climaudit/>

GRDF. (2023, juin). *Calcul des degrés jours unifiés (DJU)*. Cegibat.fr. Consulté le 9 juillet 2023, à l'adresse <https://cegibat.grdf.fr/simulateur/calcul-dju>

Ministère de la Transition écologique et de la Cohésion des territoires. (s. d.). *Indice de rigueur - degrés-jours unifiés aux niveaux national, régional et départemental*. Données et études statistiques pour le changement climatique, l'énergie, l'environnement, le logement et les transports. Consulté le 15 juin 2023, à l'adresse <https://www.statistiques.developpement-durable.gouv.fr/indice-de-rigueur-degres-jours-unifies-aux-niveaux-national-regional-et-departemental>

Ministère de la Transition énergétique. (2023, 24 janvier). *Éco Énergie Tertiaire (EET)*. Consulté le 12 mai 2023, à l'adresse <https://www.ecologie.gouv.fr/eco-energie-tertiaire-eet>

Ministères Écologie Énergie Territoires. (2023, 26 mai). *Opérations standardisées d'économies d'énergie*. Consulté le 7 juin 2023, à l'adresse https://www.ecologie.gouv.fr/operations-standardisees-deconomies-denergie#scroll-nav__2

Incub'. (s. d.). *Qu'est ce qu'un audit quantique*.

Consulté le 14 juin 2023, à l'adresse <https://www.incub.net/qu-est-ce-qu-un-audit-quantique/>

Annexes

Annexe 1 : code relatif au calcul des DJU chauds et froids

def DJU_ans(ans, sites):

"""

Parameters

ans : LIST

Année(s) prise en compte dans le calcul des DJU

sites : DATAFRAME

sites dont on souhaite calculer les coordonnées.

Pour chaque site, des coordonnées, une altitude et un nom sont renseignés.

Returns

DATAFRAME

DJU chaud et froid, moyennés sur plusieurs années, de chaque site renseigné

/!\ ce code gère les données météo manquantes ou incomplètes,

en allant chercher des données plus complètes dans une station située jusqu'à 80 km du point initial.

certaines DJU peuvent être calculés selon un nombre d'années moindre que celui renseigné en paramètre

La méthode utilisée pour le calcul est calquée sur la méthode 'météo', décrite par le CSTB, exposée ici <https://cegibat.grdf.fr/simulateur/calcul-dju>

les seuils sont pris à 18 pour le DJU chaud, et 19 pour le DJU froid

"""

sites=sites.reset_index(drop = True)

DJU = pd.DataFrame (index = sites['UT_Site'], data= None)

DJU_chaud_moy = len(sites)*[0]

DJU_froid_moy = len(sites)*[0]

for j in tqdm(range(len(sites))) :

#for j in tqdm([1]) :

DJU_froid_ans = []

DJU_chaud_ans = []

for i in (ans):

data = pd.DataFrame(data=None)

#lat = 44.56

lon=4.75

```

# altitude = 79
lat = sites.loc[j,'Lat']
lon = sites.loc[j,'Lon']
altitude = sites.loc[j, 'Altitude']
start = datetime(i,1,1)
end = datetime(i,12,31)
gare=Point(lat,lon, altitude)
data = Daily(gare, start, end) # calcul des données météo pour chaque jour de l'année i
data = data.fetch()
data = data.dropna(subset = ['tavg'])
data = data.reset_index()
echec = False
if len(data)< 329 : #gérer les cas de données incomplètes (10% ou plus manquant)
    s = Stations()
    num = s.nearby(lat, lon, 80000) #chercher les autres stations dans un rayon de 80 km
    num = num.fetch(10)
    num = num.reset_index()
    k = 0
    while len(data)<329 :
        k = k + 1
        if k > (len(num)-1) :
            echec = True
            break #échec de la boucle si aucune station satisfaisante trouvée à moins de 80 km
        else :
            lat2 = num.loc[k,'latitude']
            lon2 = num.loc[k,'longitude']
            start = datetime(i,1,1)
            end = datetime(i,12,31)
            gare=Point(lat2, lon2, altitude)
            data = Daily (gare, start, end)
            data = data.fetch()
            data = data.dropna(subset = ['tavg'])
            data = data.reset_index()
    if echec == True :

```

```
        continue # Pas de valeurs de DJU calculées pour les années incomplètes
    else :
        DJ_chaud = 0
        DJ_froid = 0

        for x in (range(len(data))): # calcul des DJU selon la méthode 'météo', considérant uniquement la
température moyenne jour
            tmin = data.loc[x,'tmin']
            tmax= data.loc[x,'tmax']
            tavg= (tmax+tmin)/2
            if tavg < 18 :
                DJ_chaud += 18 - tavg # somme des degrés d'inconfort froid selon température moyenne jour
            if tavg > 19:
                DJ_froid += tavg - 19# somme des degrés d'inconfort chaud selon température moyenne jour
            DJU_chaud_ans.append(DJ_chaud) # stockage du DJU chaud de l'année i
            DJU_froid_ans.append(DJ_froid) # stockage du DJU froid de l'année i

        DJU_chaud_moy[j] = (np.mean(DJU_chaud_ans)) # moyenne des DJU chaud sur l'ensemble des années
"complètes"
        DJU_froid_moy[j] = (np.mean(DJU_froid_ans)) # moyenne des DJU froid sur l'ensemble des années
"complètes"

        DJU['DJU_chaud']= DJU_chaud_moy
        DJU['DJU_froid']= DJU_froid_moy
    return (DJU)
```

Accompagnement au décret tertiaire

Développement d'un modèle prédictif des économies d'énergies mobilisables sur le patrimoine ferroviaire

Résumé : En 2019 entre en vigueur l'obligation Eco Energie Tertiaire (EET). Cette obligation vise à réduire, comparativement à une année de référence, la consommation en énergie finale du parc immobilier tertiaire français. 340 gares sont ainsi concernées pour tout ou partie par l'obligation. Plusieurs actions individuelles ont été recensées par les équipes AREP comme potentiels leviers d'économies d'énergie. Afin de déterminer les actions à mettre prioritairement en place au sein des gares, il s'agit de déterminer les économies d'énergies potentielles permises par chaque action et combinaison d'actions. Ce stage consiste en l'élaboration d'un modèle prédictif d'économies d'énergies répondant à ces enjeux. L'approche proposée consiste en l'application de plusieurs régressions, puis leur enrichissement par l'introduction de nouvelles variables. Le modèle obtenu fournit des ordres de grandeurs cohérents avec les résultats de la littérature.

Abstract : In 2019, the Eco Energie Tertiaire (EET) obligation came into effect. This obligation aims at reducing the energy demand of the French tertiary real estate sector compared to a reference year. A total of 340 trains stations are thus affected in whole or in part by this obligation. Several individual actions have been identified by AREP teams as potential means of energy savings. To determine the priority actions to be implemented within the stations, the task is to assess the potential energy savings enabled by each action and combination of actions. This internship implies the development of a predictive model for energy savings that addresses these challenges. The proposed approach consists in the use of a multiple hybrid regressions. The resulting model provides estimates that align with the findings of the literature.

Mots Clés : Bâtiment, Tertiaire, Energie, Rénovation, Modèle, Sobriété énergétique, Machine Learning

AREP

16 Avenue d'Ivry

75013 Paris

Tuteurs entreprise :

Milan BILLOIN

Frédéric RONSEAUX

Tuteur académique :

Mindjid MAIZIA