# Résumé

# Tables des matières

[Résumé 1](#_Toc179732692)

[Tables des matières 2](#_Toc179732693)

[Listes des tableaux 4](#_Toc179732694)

[Listes des figures 5](#_Toc179732695)

[1 Introduction 6](#_Toc179732696)

[1.1 Contexte général du marché du Gazole Non Routier en France 6](#_Toc179732697)

[1.2 Énoncé de la problématique et objectifs de recherche 6](#_Toc179732698)

[1.3 Aperçu de l’approche méthodologique et de la structure du rapport 7](#_Toc179732699)

[2 Revue de littérature 9](#_Toc179732700)

[2.1 Le marché du GNR en France : contexte et facteurs d’influence 9](#_Toc179732701)

[2.1.1 Évolution historique et réglementaire du GNR depuis 2011 9](#_Toc179732702)

[2.1.2 Importance économique et secteurs clés 10](#_Toc179732703)

[2.1.2.1 Part du GNR dans la consommation énergétique française 10](#_Toc179732704)

[2.1.2.2 Analyse sectorielle de l’utilisation du GNR 10](#_Toc179732705)

[2.1.3 Conclusion sur les facteurs économiques et les secteurs clés du GNR 11](#_Toc179732706)

[2.2 Facteurs influençant les prix du GNR 12](#_Toc179732707)

[2.2.1 Facteurs macroéconomiques 12](#_Toc179732708)

[2.2.2 Facteurs géographiques et logistiques 13](#_Toc179732709)

[2.2.3 Facteurs démographiques et économiques régionaux 14](#_Toc179732710)

[2.2.4 Facteurs saisonniers et météorologiques 15](#_Toc179732711)

[2.2.5 Conclusion 16](#_Toc179732712)

[2.3 Méthodes d'analyse des disparités régionales des prix énergétiques 17](#_Toc179732713)

[2.3.1 Analyse statistique descriptive 17](#_Toc179732714)

[2.3.2 Analyse spatiale 18](#_Toc179732715)

[2.3.3 Analyse temporelle 18](#_Toc179732716)

[2.4 Collecte et structuration des données 18](#_Toc179732717)

[2.4.1 Approche méthodologique de la collecte de données 18](#_Toc179732718)

[2.4.1.2. Critères de sélection des sources de données 19](#_Toc179732719)

[2.4.2 Techniques de web scraping utilisées 20](#_Toc179732720)

[3.2.2 Architecture du système de scraping 20](#_Toc179732721)

[3.2.3 Défis et solutions dans le processus de scraping 21](#_Toc179732722)

[3.2.4 Gestion des défis techniques 21](#_Toc179732723)

[3.2.5 Considérations éthiques et légales du web scraping 21](#_Toc179732724)

[3 Collecte et structuration des données de l’étude 22](#_Toc179732725)

[3.3 Sources des différentes variables 22](#_Toc179732726)

[3.3.1 Sources pour les données de prix du GNR 22](#_Toc179732727)

[3.3.2 Sources pour les facteurs macroéconomiques 22](#_Toc179732728)

[3.3.3 Sources pour les données géographiques et démographiques 22](#_Toc179732729)

[3.4 Couverture géographique des données collectées 22](#_Toc179732730)

[3.4.1 Présentation de la carte des codes postaux scrapés 22](#_Toc179732731)

[3.4.2 Analyse de la répartition géographique des données 22](#_Toc179732732)

[3.4.3 Discussion sur la représentativité de l'échantillon 22](#_Toc179732733)

[3.5 Fréquence et période de collecte des données 22](#_Toc179732734)

[3.6 Structuration de la base de données 22](#_Toc179732735)

[3.7 Organisation des données collectées 22](#_Toc179732736)

[3.8 Description finale de la base de données 22](#_Toc179732737)

[4 Méthodologie d’analyse 23](#_Toc179732738)

[5 Analyse des disparités régionales des prix du GNR 24](#_Toc179732739)

[5.1 Cartographie des prix à l'échelle nationale 24](#_Toc179732740)

[5.2 Identification des facteurs influents par région 24](#_Toc179732741)

[5.3 Patterns géographiques et clusters de prix 24](#_Toc179732742)

[5.4 Cas d'étude : focus sur des régions spécifiques 24](#_Toc179732743)

[Annexes 26](#_Toc179732744)

[Références bibliographiques 27](#_Toc179732745)

# Listes des tableaux

# Listes des figures

# 1 Introduction

## 1.1 Contexte général du marché du Gazole Non Routier en France

Le Gazole Non Routier (GNR) occupe une place centrale dans le paysage énergétique français, jouant un rôle crucial dans des secteurs économiques clés tels que l'agriculture, la sylviculture et les travaux publics. Introduit en 2011 dans le cadre des efforts européens visant à réduire les émissions polluantes, le GNR est rapidement devenu un élément indispensable pour de nombreuses activités économiques (TotalEnergies, 2024).

La dynamique du marché du GNR en France est caractérisée par une complexité croissante, influencée par une multitude de facteurs allant des fluctuations du marché pétrolier mondial aux politiques énergétiques nationales et européennes. L'annulation récente, en janvier 2024, de la hausse progressive de la taxe sur le GNR prévue jusqu'en 2030 illustre la sensibilité de ce marché aux décisions politiques et aux pressions économiques (Ministère de la Transition Écologique, 2022).

Les disparités régionales dans les prix du GNR soulèvent des questions importantes sur l'équité d'accès à cette ressource énergétique et sur ses implications pour la compétitivité des entreprises dans différentes régions de France. Ces variations de prix, parfois significatives, peuvent avoir des répercussions importantes sur les coûts opérationnels des entreprises et, par extension, sur l'économie locale et nationale.

## 1.2 Énoncé de la problématique et objectifs de recherche

Dans ce contexte, notre étude vise à répondre à la question centrale suivante : "Quels sont les facteurs déterminants des disparités régionales des prix du GNR en France, et comment peut-on prédire ces variations de prix de manière fiable ?"

Pour aborder cette problématique, nous nous fixons les objectifs spécifiques suivants :

1. Cartographier et analyser en profondeur les disparités régionales des prix du GNR en France.
2. Identifier et hiérarchiser les facteurs influençant ces variations de prix, qu'ils soient économiques, géographiques, ou politiques.
3. Développer des modèles prédictifs robustes pour anticiper les évolutions futures des prix du GNR à l'échelle régionale.
4. Évaluer l'impact potentiel de ces disparités de prix sur la compétitivité des entreprises et l'économie régionale.
5. Formuler des recommandations basées sur nos résultats pour les acteurs du marché et les décideurs politiques.

## 1.3 Aperçu de l’approche méthodologique et de la structure du rapport

Pour atteindre ces objectifs, notre étude s'appuie sur une approche multidimensionnelle combinant :

* Une analyse approfondie du contexte historique et réglementaire du marché du GNR en France.
* Une collecte de données utilisant des techniques de web scraping pour obtenir des informations détaillées sur les prix du GNR dans plus de 2000 codes postaux.
* L'utilisation de techniques avancées d'analyse spatiale et temporelle pour identifier les patterns et les tendances dans les variations de prix.
* L'application de modèles statistiques et de machine learning pour analyser les facteurs d'influence et développer des capacités prédictives.

Le rapport est structuré comme suit :

* La section 2 présente une analyse détaillée du marché du GNR en France, incluant son évolution historique et les facteurs influençant ses prix.
* La section 3 décrit notre méthodologie de collecte et de structuration des données, mettant en lumière notre approche innovante de web scraping.
* La section 4 expose les méthodologies d'analyse employées dans notre étude.
* Les sections 5 et 6 présentent respectivement nos analyses des disparités régionales et de la dynamique temporelle des prix du GNR.
* La section 7 discute des implications de nos résultats et propose des recommandations.
* Enfin, la conclusion synthétise nos principales découvertes et ouvre des perspectives pour de futures recherches.

À travers cette étude, nous visons à apporter un éclairage nouveau sur les dynamiques complexes du marché du GNR en France, offrant ainsi des insights précieux pour les acteurs économiques et les décideurs politiques dans ce domaine crucial de l'économie française.

# 2 Revue de littérature

## 2.1 Le marché du GNR en France : contexte et facteurs d’influence

### 2.1.1 Évolution historique et réglementaire du GNR depuis 2011

L'introduction du Gazole Non Routier (GNR) en France s'inscrit dans le cadre plus large de la politique énergétique européenne. La directive 2009/30/CE du Parlement européen et du Conseil du 23 avril 2009 a posé les bases de cette évolution, visant à réduire l'impact environnemental des carburants utilisés dans les secteurs non routiers (Commission européenne, 2009). Cette directive a marqué un tournant dans la réglementation des carburants en Europe, en imposant des normes plus strictes sur la qualité des carburants et en encourageant l'utilisation de carburants à faible teneur en soufre.

En France, l'implémentation de cette directive s'est traduite par l'introduction du GNR le 1er mai 2011, remplaçant le fioul domestique dans les engins mobiles non routiers, les tracteurs agricoles et forestiers. Cette transition a été encadrée par l'arrêté du 10 décembre 2010, qui a défini les spécifications techniques du GNR et les conditions de son utilisation (Ministère de la Transition écologique, 2010). Les premières années ont été caractérisées par une phase d'apprentissage et d'ajustement, au cours de laquelle les acteurs du marché ont dû s'adapter à ce nouveau carburant et à ses spécificités.

Entre 2014 et 2017, le marché du GNR a connu une période de stabilisation et de croissance régulière. Cette phase a vu une amélioration des réseaux de distribution et une meilleure compréhension des avantages du GNR par les utilisateurs finaux. Cette relative stabilité a cependant été mise à l'épreuve en 2018, lorsque la hausse des cours du pétrole a fortement impacté les prix du GNR, suivie d'une chute marquée qui a souligné la vulnérabilité du marché face aux fluctuations internationales (INSEE, 2018).

La pandémie de COVID-19 en 2020 a introduit un nouveau niveau de complexité, entraînant une volatilité accrue des prix et des perturbations dans les chaînes d'approvisionnement. Cette période a mis en lumière la nécessité d'une plus grande résilience dans le secteur énergétique (Terre-net, 2020).

Les années 2021 à 2023 ont été marquées par un renforcement des normes environnementales et des débats intenses sur l'avenir fiscal du GNR (Assemblée nationale, 2021). Ces discussions reflétaient la tension croissante entre les objectifs de transition écologique et les préoccupations économiques des secteurs dépendant du GNR.

L'année 2024 a débuté avec une décision politique majeure : l'annulation de la hausse progressive de la taxe sur le GNR prévue jusqu'en 2030. Cette décision illustre la complexité des arbitrages entre considérations environnementales, économiques et sociales dans la gestion de ce marché (Inspection générale des finances, 2024).

## 2.1.2 Importance économique et secteurs clés

### 2.1.2.1 Part du GNR dans la consommation énergétique française

Le Gazole Non Routier (GNR) est un élément important du mix énergétique en France, surtout dans des secteurs tels que l’agriculture, les travaux publics et la sylviculture. Selon le Bilan énergétique de la France publié par le Service de la Donnée et des Études Statistiques (SDED) du Ministère de la transition Écologique, le pétrole représente environ 39 % de la consommation finale d'énergie en 2022, dont une part significative est allouée aux carburants pour les secteurs industriels et agricoles (Service de la Donnée et des Études Statistiques, 2023). Cela inclut le GNR, essentiel pour les machines lourdes et les engins non routiers dans ces secteurs.

### 2.1.2.2 Analyse sectorielle de l’utilisation du GNR

Le Gazole Non Routier (GNR) joue un rôle crucial dans plusieurs secteurs clés de l'économie française. Cette section analyse l'utilisation de celui-ci dans les principaux secteurs consommateurs.

#### Secteur agricole :

D'après Audrey Trévisiol, Coordinatrice Agriculture et Changement Climatique à l'Agence de l'Environnement et de la Maîtrise de l'Énergie (ADEME), le secteur agricole est le premier consommateur de GNR en France (Trévisiol, 2023). Cette affirmation est confirmée par le Ministère de l’Agriculture et de l'Alimentation, qui a indiqué qu’en 2021, le secteur agricole a utilisé environ 29,4 TWh de Gazole Non Routier, soit 60 % de sa consommation énergétique totale. Ces données soulignent l'importance du GNR pour soutenir les activités agricoles, en en faisant une ressource énergétique clé pour ce secteur​ (SSP, 2024).

#### Secteur des travaux publics et de la construction :

En 2015, le secteur de la construction en France dépend fortement des produits pétroliers, qui représentent 89 % de sa consommation énergétique totale. Le Gazole Non Routier (GNR) en constitue 22 %, étant largement utilisé dans les travaux publics pour les engins lourds et les grands chantiers. Le gazole routier, dominant à 61 %, est plus courant dans le bâtiment pour les petits chantiers et le transport. Dans les travaux publics, 97 % du GNR est dédié aux activités de chantier, tandis que le transport et les locaux administratifs occupent une part secondaire, illustrant ainsi la dépendance majeure de ce secteur aux énergies fossiles (Commissariat général au développement durable, 2018).

#### Secteur forestier :

Selon l'Office National des Forêts (ONF), le secteur forestier, bien que moins consommateur que l'agriculture ou la construction, dépend fortement du GNR pour ses opérations ( Office National des Forêts, 2021). En 2021, la consommation de GNR dans ce secteur était estimée à environ 100 000 mètres cubes par an, représentant une part significative des coûts opérationnels. Une étude menée par l'Institut technologique FCBA (Forêt Cellulose Bois-construction Ameublement) en 2021 a révélé que le GNR constitue en moyenne 12% des coûts totaux d'exploitation forestière, pouvant atteindre 18% pour l'abattage mécanisé et 15% pour le débardage. Ces chiffres soulignent l'importance du GNR dans les activités forestières, particulièrement dans les zones difficiles d'accès où l'électrification des engins n'est pas encore viable (Institut technologique Forêt Cellulose Bois-construction, 2021).

### 2.1.3 Conclusion sur les facteurs économiques et les secteurs clés du GNR

L'analyse de l'importance économique du Gazole Non Routier (GNR) et de son utilisation dans les secteurs clés de l'économie française révèle son rôle crucial dans le paysage énergétique national. Le GNR s'impose comme une ressource énergétique essentielle, particulièrement dans les domaines de l'agriculture, des travaux publics, de la construction et de la sylviculture.

La part significative du GNR dans la consommation énergétique française, notamment dans les secteurs industriels et agricoles, souligne sa place prépondérante dans le mix énergétique national. Cette importance est d'autant plus marquée lorsqu'on examine son utilisation sectorielle.

Le secteur agricole se distingue comme le principal consommateur de GNR, représentant 60% de sa consommation énergétique totale. Cette dépendance illustre le rôle vital du GNR dans le maintien et le développement des activités agricoles françaises.

Dans le domaine des travaux publics et de la construction, le GNR occupe également une place centrale, constituant 22% de la consommation énergétique du secteur. Son utilisation prédominante dans les grands chantiers et pour les engins lourds témoigne de son importance dans la réalisation des infrastructures et des projets de construction à grande échelle.

Bien que moins consommateur en volume, le secteur forestier dépend fortement du GNR pour ses opérations, particulièrement dans les zones difficiles d'accès. La part significative du GNR dans les coûts d'exploitation forestière souligne son impact économique dans ce secteur.

Cette analyse sectorielle met en lumière la dépendance de secteurs clés de l'économie française au GNR, soulignant ainsi l'importance stratégique de ce carburant. Elle révèle également les défis potentiels liés à cette dépendance, notamment en termes de transition énergétique et d'impact environnemental.

La compréhension de ces dynamiques sectorielles est cruciale pour appréhender les enjeux liés aux prix du GNR et leurs répercussions sur l'économie française. Elle fournit un contexte essentiel pour l'analyse des disparités régionales des prix du GNR et souligne l'importance d'une approche nuancée, prenant en compte les spécificités de chaque secteur dans l'élaboration de politiques énergétiques et économiques.

## 2.2 Facteurs influençant les prix du GNR

La formation des prix du GNR en France est le résultat d'une interaction complexe entre divers facteurs, allant des dynamiques macroéconomiques globales aux spécificités locales. Comprendre ces facteurs est essentiel pour appréhender les disparités régionales et anticiper les évolutions futures des prix.

### 2.2.1 Facteurs macroéconomiques

Les facteurs macroéconomiques, qui opèrent à l'échelle mondiale et nationale, jouent un rôle crucial dans la détermination des coûts de production, d'importation et de distribution du GNR. Parmi ces facteurs, deux se distinguent par leur impact significatif et leur nature dynamique sur le marché international : le cours du baril de pétrole et le taux de change USD/EUR.

#### 2.2.1.1 Cours du baril de pétrole

Le cours du baril de pétrole est un des facteurs clés, à la fois visible et volatil, qui influence les prix du Gazole Non Routier (GNR) en France. En tant que composant principal des produits raffinés, une variation du prix du baril de pétrole brut affecte directement les coûts de production du GNR, en plus des coûts de transport et de distribution qui y sont associés.

Les fluctuations du marché pétrolier mondial, qu'elles soient dues à des événements géopolitiques, des décisions de l'OPEP, ou des changements dans l'offre et la demande globales, se répercutent rapidement sur les prix du GNR en France (Fernandois & Medel, 2020). Cette sensibilité aux cours mondiaux du pétrole souligne l'interconnexion du marché français du GNR avec les dynamiques énergétiques globales.

#### 2.2.1.2 Taux de Change USD/EUR

Le taux de change USD/EUR exerce une influence significative sur les prix du GNR. La majeure partie du pétrole brut étant négociée en dollars américains, les variations du taux de change impactent directement les coûts d'importation du pétrole en France. Un euro fort tend à réduire ces coûts, tandis qu'un euro faible les augmente, affectant ainsi les prix finaux du GNR pour les consommateurs français (Beckmann & Czudaj, 2013).

### 2.2.2 Facteurs géographiques et logistiques

Les prix du GNR sont influencés par des facteurs géographiques et logistiques, en particulier la localisation des dépôts de carburant et la qualité de l'infrastructure de transport. Ces éléments jouent un rôle crucial dans la distribution du GNR, impactant directement les coûts de transport et donc les prix finaux dans différentes régions.

#### 2.2.2.1. Localisation des dépôts

La position des dépôts de carburant par rapport aux zones de consommation affecte directement les coûts de distribution. Plus les dépôts sont éloignés des consommateurs finaux (ex. agriculteurs, sites industriels), plus les coûts de transport sont élevés. En effet, les coûts supplémentaires liés aux trajets prolongés et aux besoins en carburant pour les camions-citernes sont répercutés sur les prix du GNR dans les régions isolées.

#### 2.2.2.2. Infrastructure de transport et prix du carburant

Les coûts associés au transport du Gazole Non Routier (GNR) en France reposent sur une combinaison de facteurs logistiques. Tout d’abord, le carburant, représentant une part significative des dépenses, est extrêmement sensible aux variations de prix du pétrole. Ces fluctuations influencent directement le coût des opérations de transport, impactant ainsi les prix finaux pour les consommateurs. Les frais liés à la main-d'œuvre, notamment les salaires et charges sociales des chauffeurs, sont également importants. Toute augmentation dans ces coûts, due par exemple à des ajustements salariaux ou à des pénuries de conducteurs qualifiés, accroît les charges globales du transport​.

Les frais de maintenance et d’entretien des véhicules, qui varient selon leur fréquence d’utilisation et les conditions de transport, ajoutent une autre couche de coûts. Le vieillissement des camions et les normes de sécurité plus strictes contribuent également à ces dépenses. Par ailleurs, l’amortissement des véhicules, calculé en fonction de leur durée de vie et de la fréquence de renouvellement de la flotte, constitue un coût structurel non négligeable. Enfin, les dépenses en assurances et les coûts de gestion administrative associés au transport – incluant la planification logistique et la conformité aux régulations – complètent les charges. Chacun de ces éléments contribue à façonner les coûts logistiques globaux, influençant ainsi le prix du GNR pour les secteurs dépendants (Rodrigue, 2024)​.

### 2.2.3 Facteurs démographiques et économiques régionaux

Pour comprendre comment les facteurs démographiques et économiques régionaux influencent les prix du GNR, on peut effectivement tracer un parallèle avec d’autres produits pétroliers, en s’appuyant sur des études de consommation et de distribution d’énergie en général.

#### 2.2.3.1. Densité de population

Des études sur d’autres carburants, comme l'essence et le diesel, montrent que les zones à haute densité de population tendent à avoir des coûts de distribution plus faibles en raison des économies d'échelle. Ces régions, qui concentrent la demande sur une zone géographique réduite, permettent des trajets de distribution plus courts et donc plus économiques. Dans les zones à faible densité, comme les campagnes, où les distances entre les points de livraison sont plus grandes, les coûts de transport augmentent, ce qui se répercute sur les prix de vente au détail des carburants​. Les produits pétroliers destinés aux zones moins peuplées affichent ainsi une prime logistique, comparable à celle observée pour le GNR dans des régions rurales ou éloignées (Brons, Nijkamp, Pels, & Rietveld, 2008).

#### 2.2.3.2. Activité économique locale

Dans le cadre de l’analyse de l’impact de l’activité économique locale sur les prix du GNR, il est pertinent de considérer les zones où ce carburant est le plus utilisé, comme expliqué dans la section 2.2.2 dédiée à l’utilisation sectorielle du GNR. Les régions dominées par des secteurs consommateurs majeurs – en particulier l’agriculture, les travaux publics et l’industrie – montrent une demande particulièrement forte en GNR. Cette demande soutenue provient des besoins en carburant pour les machines agricoles, les engins de construction et autres équipements énergivores.

Ainsi, dans les zones à forte activité agricole ou industrielle, les volumes de GNR requis sont conséquents, ce qui permet aux acheteurs de négocier des prix plus compétitifs en raison de la demande élevée. En revanche, dans les régions où l’activité économique est plus modérée, les coûts de distribution et de transport augmentent souvent, ce qui se traduit par des prix de GNR plus élevés pour les utilisateurs finaux​.

### 2.2.4 Facteurs saisonniers et météorologiques

L'influence des facteurs saisonniers et météorologiques sur le GNR constitue un aspect intrigant de notre analyse. Bien que leur impact précis reste à quantifier par des études empiriques approfondies, plusieurs hypothèses méritent notre attention.

#### 2.2.4.1. Variations saisonnières de la demande

Les fluctuations saisonnières de la demande en GNR pourraient jouer un rôle significatif dans la dynamique des prix. L'hiver semble particulièrement crucial, avec une possible augmentation de la demande liée à deux facteurs principaux : d'une part, l'utilisation accrue du GNR pour le chauffage dans certains secteurs, et d'autre part, la consommation potentiellement plus élevée des engins agricoles et de chantier en conditions froides. À l'inverse, l'été pourrait connaître une baisse de la demande, résultant de la diminution des besoins en chauffage et du ralentissement saisonnier de certaines activités industrielles ou agricoles.

Au-delà de cette dichotomie été-hiver, certains secteurs utilisant le GNR, notamment l'agriculture, pourraient connaître des pics d'activité saisonniers lors des périodes de semis ou de récoltes, influençant ainsi la demande de manière ponctuelle mais potentiellement significative. Ces variations cycliques, si elles se confirment, pourraient se traduire par des fluctuations prévisibles des prix du GNR au fil des saisons.

#### 2.2.4.2. Impact des conditions météorologiques

Au-delà des variations saisonnières régulières, les phénomènes météorologiques ponctuels et parfois extrêmes peuvent exercer une influence significative sur le marché du GNR. Ces événements climatiques intenses affectent à la fois l'offre et la demande du carburant. Du côté de l'offre, les intempéries peuvent perturber la chaîne logistique, entraînant des pénuries localisées et des variations de prix à court terme. Les conditions extrêmes peuvent également impacter la production, avec le froid affectant le fonctionnement des raffineries et les canicules pouvant réduire l'activité de certaines installations industrielles.

Côté demande, les variations climatiques influencent directement la consommation de GNR. Par temps froid, l'efficacité réduite des moteurs augmente la consommation. Le secteur agricole, particulièrement sensible au climat, voit sa demande en GNR fluctuer selon les conditions météorologiques, que ce soit pour l'irrigation en période de sécheresse ou pour s'adapter à des saisons de culture modifiées par le climat.

Ces impacts s'inscrivent dans un contexte énergétique plus large, où des phénomènes météorologiques affectant d'autres sources d'énergie peuvent indirectement influencer la demande en GNR. Pour quantifier précisément ces effets, une analyse croisée des données météorologiques, de production, de distribution et des indicateurs économiques sectoriels est nécessaire, permettant ainsi de distinguer l'impact spécifique des conditions météorologiques sur le marché du GNR.

### 2.2.5 Conclusion

L'analyse des facteurs influençant les prix du Gazole Non Routier (GNR) en France révèle un réseau complexe d'interactions entre variables macroéconomiques, géographiques, logistiques, démographiques, économiques et météorologiques. Cette multiplicité de facteurs explique en grande partie les disparités régionales observées dans les prix du GNR, soulignant la complexité du marché énergétique français.

Les facteurs macroéconomiques, tels que le cours du baril de pétrole et le taux de change USD/EUR, jouent un rôle fondamental dans la détermination des prix de base du GNR. Cependant, ce sont les facteurs géographiques et logistiques qui contribuent significativement aux variations régionales des prix. La localisation des dépôts et la qualité des infrastructures de transport influencent directement les coûts de distribution, créant ainsi des disparités entre les régions urbaines et rurales.

Les aspects démographiques et économiques régionaux ajoutent une couche supplémentaire de complexité. La densité de population et l'activité économique locale, notamment dans les secteurs fortement consommateurs de GNR comme l'agriculture et l'industrie, influencent la demande et, par conséquent, les prix. Ces facteurs peuvent expliquer pourquoi certaines régions bénéficient de prix plus compétitifs malgré leur éloignement géographique des sources d'approvisionnement.

Enfin, l'impact des facteurs saisonniers et météorologiques, bien que moins étudié, semble jouer un rôle non négligeable dans les fluctuations temporelles des prix du GNR. Les variations saisonnières de la demande et les événements météorologiques extrêmes peuvent entraîner des perturbations à court terme dans l'offre et la demande, influençant ainsi les prix de manière ponctuelle mais potentiellement significative.

Cette analyse exhaustive des facteurs influençant les prix du GNR souligne la nécessité d'une approche multidimensionnelle pour comprendre et prédire les disparités régionales. Elle met en évidence l'importance de prendre en compte non seulement les tendances macroéconomiques globales, mais aussi les spécificités locales et temporelles dans toute analyse ou modélisation des prix du GNR en France. Cette compréhension approfondie servira de base solide pour les analyses quantitatives et les modélisations prédictives qui suivront dans cette étude.

## 2.3 Méthodes d'analyse des disparités régionales des prix énergétiques

L'analyse des disparités régionales des prix énergétiques, et plus particulièrement du GNR, nécessite une approche multidimensionnelle combinant diverses techniques statistiques et analytiques. Cette section présente les principales méthodes utilisées dans la littérature récente, qui seront appliquées dans notre étude.

### 2.3.1 Analyse statistique descriptive

Les analyses statistiques descriptives constituent la base de toute étude sur les disparités de prix. Elles permettent de caractériser la distribution des prix et d'identifier les tendances générales. Asche et al. (2013) ont utilisé ces techniques pour analyser les prix du gaz naturel en Europe, mettant en évidence l'importance des mesures de tendance centrale et de dispersion pour comprendre les dynamiques de marché. Dans notre étude, nous emploierons des histogrammes et des boîtes à moustaches pour visualiser la distribution des prix du GNR, suivant l'approche de Mirza et Bergland (2012) dans leur analyse des prix de l'électricité dans les pays nordiques.

### 2.3.2 Analyse spatiale

Les méthodes d'analyse spatiale sont cruciales pour comprendre les disparités géographiques des prix. La cartographie thématique, notamment les cartes choroplèthes, offre une visualisation intuitive des variations spatiales. Ramos et al. (2016) ont utilisé cette approche pour étudier les prix de l'essence au Portugal, révélant des patterns géographiques significatifs. Dans notre étude, nous adopterons une approche similaire pour le GNR en France.

### 2.3.3 Analyse temporelle

L'aspect temporel est essentiel dans l'analyse des prix énergétiques. La décomposition des séries temporelles, comme l'ont fait Zhao et al. (2018) pour les prix du charbon en Chine, permet d'isoler les tendances, les saisonnalités et les composantes irrégulières. Nous appliquerons des techniques similaires pour comprendre l'évolution des prix du GNR au fil du temps.

Les modèles ARIMA (Autoregressive Integrated Moving Average), largement utilisés dans la prévision des prix énergétiques (Nogales et Conejo, 2006), seront également considérés pour modéliser et prévoir les tendances futures des prix du GNR.

# 2.4 Collecte et structuration des données

### 2.4.1 Approche méthodologique de la collecte de données

#### 2.4.1.1. Revue des méthodes de collecte de données dans les études de marché énergétique

Le web scraping est devenu une technique essentielle pour collecter des données sur les prix, soutenant ainsi les statistiques officielles et l’analyse économique. Les instituts de statistique explorent de plus en plus cette méthode pour moderniser leurs processus de collecte, en particulier pour les informations sur les prix (Virgillito & Polidoro, 2017). Comparé à la collecte manuelle, le web scraping présente des avantages importants, notamment une efficacité accrue et la production de statistiques de prix plus précises.

Son accessibilité et son faible coût en font également une option attrayante pour les chercheurs et les entreprises, avec des applications variées, allant des calculs d’indices de prix à l’analyse des tendances du marché (Śpiewanowski, Talavera, & Vi, 2021). Malgré ses nombreux atouts, le web scraping présente certains défis, notamment en matière de réglementation, de risques technologiques et de besoins de formation (Boettcher, 2015). Toutefois, il reste un outil précieux pour les chercheurs et la société civile, permettant une collecte et une interprétation des données socioéconomiques plus rapides et approfondies (Uriarte, Toro, & Larrosa, 2020).

### 2.4.1.2. Critères de sélection des sources de données

Dans le domaine de l'analyse de données, la collecte d'informations provenant de sources diverses ou par le biais du web scraping nécessite une attention particulière à plusieurs critères essentiels. Ces derniers garantissent la qualité, la fiabilité et la pertinence des données recueillies. Parmi ces critères, on peut citer la qualité intrinsèque des données, la crédibilité des sources, les considérations éthiques, la couverture géographique, la fréquence de mise à jour, la fiabilité globale et l'accessibilité via des plateformes de données ouvertes, des interfaces de programmation (API) ou des techniques de web scraping.

La qualité des données constitue le fondement de toute analyse rigoureuse. Elle se caractérise par l'exactitude, l'exhaustivité, la cohérence et l'actualité des informations. Gressler souligne l'importance cruciale de la qualité des données issues du monde réel (RWD) dans le cadre d'évaluations approfondies, notamment dans le domaine des dispositifs médicaux (Gressler, et al., 2024). Dans cette optique, Li et al. proposent une méthodologie robuste pour évaluer la qualité des données dans les environnements de big data, mettant en exergue la nécessité de cadres d'évaluation holistiques (Li, Xu, & Peng, 2021). Par ailleurs, la revue systématique menée par Migueles et al. met en lumière l'impact significatif des critères de collecte et de traitement des données sur l'interprétation des résultats d'activité physique, soulignant ainsi l'importance capitale de méthodes de collecte fiable (Migueles, et al., 2017).

**Crédibilité des sources :** La crédibilité des sources de données revêt une importance tout aussi primordiale. Les travaux de Yanling et al. (2021) portent sur une méthode novatrice d'évaluation de la crédibilité des données de capteurs par le biais de la fusion d'informations hétérogènes multi-sources, une approche potentiellement transposable à l'évaluation d'autres types de sources. Cohen et al. (2017) mettent en avant les forces et les limites de diverses sources de données nationales dans le secteur de la santé, soulignant la nécessité d'une évaluation minutieuse des méthodologies de collecte.

**Mise-à-jour des sources :** La fréquence de mise à jour des données constitue un critère déterminant pour maintenir leur pertinence et leur capacité à refléter fidèlement les conditions actuelles. L'étude menée par Kempen et al. (2012) compare l'efficacité de la cartographie des sols conventionnelle à celle de la cartographie numérique pour la mise à jour des cartes pédologiques, illustrant ainsi l'importance de mises à jour régulières pour préserver la pertinence des données. Les conclusions de Nwadialor et Obi (2020) démontrent que la fréquence de collecte des données influence positivement la suffisance et l'actualité des rapports d'audit, un constat potentiellement transposable à d'autres domaines.

La couverture géographique des données spatiales s'avère cruciale, notamment pour les analyses dépendant d'informations localisées. Les recherches de Nittel et al. (2012) sur l'interpolation spatiale en temps réel de phénomènes continus à l'aide de flux de données de capteurs mobiles soulignent l'importance du contexte géographique dans la collecte de données. Les travaux de Dugas et al. (2010) mettent en lumière le rôle essentiel des données de routine issues des systèmes d'information hospitaliers dans le recrutement de patients, un processus fortement influencé par les facteurs géographiques. Zahnd et al. (2018) notent que les données accessibles via le CDC offrent une couverture géographique plus étendue que d'autres sources, un atout majeur pour une surveillance sanitaire exhaustive.

La fiabilité des sources de données est un élément clé pour garantir la validité des résultats qui en découlent. Wade et al. (2011) abordent la question de l'intégration des données cliniques et de recherche, soulignant la nécessité de sources fiables pour étayer une prise de décision éclairée dans le domaine de la santé. La revue systématique de Migueles et al. (2017) met en évidence l'impact des critères de collecte et de traitement des données sur l'interprétation des résultats d'activité physique, renforçant ainsi l'importance de méthodes de collecte fiables.

L'accessibilité des données, que ce soit par le biais d'initiatives de données ouvertes, d'interfaces de programmation (API) ou de techniques de web scraping, constitue un facteur crucial dans l'évaluation des sources. Les initiatives de données ouvertes favorisent la transparence et facilitent la recherche en offrant un accès gratuit aux ensembles de données. Les API permettent un accès structuré aux informations, optimisant ainsi le processus de récupération des données pour les chercheurs. Khder (2021) met en avant les avantages du web scraping comme méthode de collecte de données, soulignant sa fiabilité et sa flexibilité d'accès. Li et al. (2019) proposent une approche systématique du web scraping, démontrant comment les informations en ligne peuvent être interrogées et extraites de manière reproductible. L'utilisation des API est également considérée comme une compétence fondamentale en science des données, facilitant une acquisition efficiente des données ("Réduction de l'asymétrie d'information dans le commerce électronique : l'approche du web scraping", 2024).

Les considérations éthiques occupent une place prépondérante dans le processus de collecte de données, particulièrement dans le cadre du web scraping. Il est impératif de respecter scrupuleusement les législations relatives à la protection de la vie privée et des données personnelles, en veillant à ce que les méthodes de collecte ne portent aucunement atteinte aux droits individuels. À cet égard, les recommandations formulées par le groupe de travail ISPOR-ISPE (Berger et al., 2017) mettent l'accent sur l'importance des pratiques éthiques dans la collecte de données, afin de renforcer la confiance des décideurs dans les preuves obtenues.

En conclusion, la collecte de données provenant de sources diverses ou par le biais du web scraping nécessite une attention particulière à plusieurs critères cruciaux : la qualité des données, la crédibilité des sources, les considérations éthiques, la couverture géographique, la fréquence de mise à jour, la fiabilité globale et l'accessibilité via les données ouvertes, les API ou le web scraping. L'adhésion rigoureuse à ces critères non seulement renforce la fiabilité des données collectées, mais garantit également que les résultats qui en découlent sont valides, pertinents et exploitables dans un contexte de recherche ou de prise de décision.

### 2.4.2 Techniques de web scraping utilisées

#### 2.4.2.1. Revue des techniques de web scraping

Le web scraping est devenu un outil incontournable dans la recherche économique moderne. Kang et Guo (2023) ont réalisé une revue exhaustive des techniques de web scraping appliquées à l'analyse des marchés énergétiques, soulignant l'importance de l'automatisation et de la gestion des défis techniques tels que les changements de structure des sites web et les mécanismes anti-bot.

### 3.2.2 Architecture du système de scraping

Notre système de scraping a été conçu selon une architecture modulaire, inspirée des meilleures pratiques décrites par Chen et al. (2021). Cette architecture comprend :

* Un module de planification des tâches
* Un module d'extraction de données
* Un module de traitement et de nettoyage
* Un module de stockage des données
* Un module de contrôle et de surveillance

### 3.2.3 Défis et solutions dans le processus de scraping

Nous avons utilisé Python comme langage principal, avec les bibliothèques suivantes :

* BeautifulSoup pour le parsing HTML
* Selenium pour l'interaction avec les pages web dynamiques
* Scrapy pour la gestion des requêtes à grande échelle
* Pandas pour le traitement des données

### 3.2.4 Gestion des défis techniques

Pour surmonter les défis techniques, nous avons implémenté plusieurs stratégies :

* Rotation des User-Agents et des adresses IP pour éviter la détection
* Utilisation de délais variables entre les requêtes
* Mise en place de mécanismes de reprise après erreur
* Surveillance continue de la structure des sites cibles

### 3.2.5 Considérations éthiques et légales du web scraping

Notre approche respecte les directives éthiques proposées par Zook et al. (2017) pour la recherche utilisant des données en ligne. Nous avons veillé à :

* Respecter les fichiers robots.txt
* Limiter la fréquence des requêtes pour ne pas surcharger les serveurs
* Anonymiser les données collectées lorsque nécessaire
* Obtenir les autorisations nécessaires lorsque possible

# 3 Collecte et structuration des données de l’étude

## 3.3 Sources des différentes variables

### 3.3.1 Sources pour les données de prix du GNR

### 3.3.2 Sources pour les facteurs macroéconomiques

### 3.3.3 Sources pour les données géographiques et démographiques

## 3.4 Couverture géographique des données collectées

### 3.4.1 Présentation de la carte des codes postaux scrapés

### 3.4.2 Analyse de la répartition géographique des données

### 3.4.3 Discussion sur la représentativité de l'échantillon

## 3.5 Fréquence et période de collecte des données

## 3.6 Structuration de la base de données

## 3.7 Organisation des données collectées

## 3.8 Description finale de la base de données

# 4 Méthodologie d’analyse

Approche analytique globale

Techniques d'analyse spatiale

Analyse des séries temporelles

# 5 Analyse des disparités régionales des prix du GNR

## 5.1 Cartographie des prix à l'échelle nationale

## 5.2 Identification des facteurs influents par région

## 5.3 Patterns géographiques et clusters de prix

## 5.4 Cas d'étude : focus sur des régions spécifiques

# Annexes

# Références bibliographiques

Office National des Forêts. (2021). *Rapport annuel 2021.* Paris.

Assemblée nationale. (2021). *LOI n° 2021-1104 du 22 août 2021 portant lutte contre le dérèglement climatique et renforcement de la résilience face à ses effets.* Paris: Légifrance.

Beckmann, J., & Czudaj, R. (2013). Oil Prices and Effective Dollar Exchange Rates. *International Review of Economics & Finance*, 21-31.

Boettcher, I. (2015). *Automatic data collection on the Internet (web scraping).*

Brons, M., Nijkamp, P., Pels, E., & Rietveld, P. (2008). *A Meta-Analysis of the Price Elasticity of Gasoline Demand. A System of Equations Approach.* 2105-2122: Energy Economics.

Commissariat général au développement durable. (2018). *Consommation d'énergie dans le secteur de la construction : Les produits pétroliers, principale énergie dans le secteur de la construction.* Ministère de la Transition écologique. URL : https://www.statistiques.developpement-durable.gouv.fr/media/573/download.

Commission européenne. (2009). *Directive 2009/30/CE du Parlement européen et du Conseil du 23 avril 2009.* Bruxelles: Union européenne.

Fernandois, A., & Medel, C. (2020). Geopolitical tensions, OPEC news, and the oil price: A granger causality analysis. *Economic Analysis Review*, 57-90.

INSEE. (2018). *En décembre 2018, le prix du pétrole chute de nouveau.* Paris: INSEE.

Inspection générale des finances. (2024). *Adaptation de la fiscalité aux exigences de la transition écologique.* Paris: Inspection générale des finances.

Institut technologique Forêt Cellulose Bois-construction. (2021). *Analyse des coûts d'exploitation forestière et impact des prix énergétiques.*

Ministère de la Transition écologique. (2010). Arrêté du 10 décembre 2010 définissant les spécifications techniques du GNR. Paris: Ministère de la Transition écologique.

Ministère de la Transition Écologique. (2022). *Évolution des taxes et subventions sur les carburants non routiers.* Paris: Ministère de la Transition Écologique.

Rodrigue, J.-P. (2024). *The spatial organization of transportation and mobility*. Récupéré sur The Geography of Transport Systems: https://transportgeography.org/contents/chapter3/transport-costs/transport-costs-conditions/

Service de la Donnée et des Études Statistiques. (2023). *Chiffres clés de l'énergie - Édition 2023.* Paris: Ministère de la transition Écologique.

Śpiewanowski, P., Talavera, O., & Vi, L. (2021). *Applications of Web Scraping in Economics and Finance.* Oxford Research Encyclopedia of Economics and Finance.

SSP, I. &. (2024). *Transformations de l'agriculture et des consommations alimentaires.* Ministère de l’Agriculture et de la Souveraineté Alimentaire.

Terre-net. (2020). *Comment a évolué le prix du GNR en 2020*. Récupéré sur https://www.terre-net.fr: https://www.terre-net.fr/carburants/article/175351/prix-du-gnr-quelle-evolution-en-2020-

TotalEnergies. (2024). *Gazole Non Routier : le carburant pour les engins non routiers*. Récupéré sur https://services.totalenergies.fr/: https://services.totalenergies.fr/professionnels/energies-mobilite/carburants/produits/diesel/gazole-non-routier

Trévisiol, A. (2023, août 5). « Le gazole non routier (GNR) reste la principale dépense énergétique des exploitations agricoles ». (B. D. (www.réussir.fr), Intervieweur)

Uriarte, J. I., Toro, G. R., & Larrosa, J. (2020). *Web scraping based online consumer price index: The “IPC Online” case.*

Virgillito, A., & Polidoro, F. (2017). *Big Data Techniques for Supporting Official Statistics: The Use of Web Scraping for Collecting Price Data.*