

MÉMOIRE DE RECHERCHE - DATA SCIENCE
QUANTITATIVE

L'impact des Variables Énergétiques sur la Croissance Économique

Analyse statistique et modélisation des données internationales

Auteurs :
Yoav Cohen, Salma Lahbati



August 4, 2025

Contents

Introduction générale	2
1 Cadre conceptuel et revue de littérature	2
1.1 Concepts clés	2
1.2 Relations théoriques	2
1.3 Travaux empiriques	3
2 Données, transformations et analyse exploratoire	4
2.1 Présentation des données	4
2.2 Prétraitement des données	5
2.3 Exploration des données	5
3 Modélisation et résultats	8
3.1 Modélisation économétrique	8
3.2 Machine learning	12
3.3 Approche deep learning	13
3.4 Comparaison France vs Monde	13
Discussion, limites et perspectives	13
3.5 Synthèse des résultats	13
3.6 Limites	13
3.7 Perspectives	13
Conclusion générale	13
Annexes	13

Introduction générale

Ce mémoire analyse l'impact des variables énergétiques sur la croissance économique mondiale à partir des données de la base OWID.

Problématique

Dans quelle mesure la consommation d'énergie et les émissions associées influencent-elles le PIB par habitant ?

Méthodologie

Nous combinons des approches statistiques et de machine learning pour répondre à cette question.

1 Cadre conceptuel et revue de littérature

1.1 Concepts clés

Le **Produit Intérieur Brut (PIB) par habitant** correspond à la valeur totale de tous les biens et services produits dans un pays au cours d'une année, divisée par sa population. Il constitue une mesure classique du niveau de vie moyen et est souvent utilisé comme indicateur synthétique du développement économique. Dans notre étude, nous utilisons la variable `gdp_per_capita` issue de la base de données mondiale pour comparer les niveaux de richesse entre pays et au cours du temps.

La **consommation énergétique primaire** (`primary_energy_consumption`) représente la quantité totale d'énergie utilisée par un pays, incluant toutes les sources (fossiles, nucléaire, renouvelables). Cette consommation peut être rapportée à la population pour obtenir la **consommation énergétique par habitant** (`energy_per_capita`), qui reflète l'intensité de l'usage de l'énergie dans un pays donné.

L'**intensité carbone de l'énergie** est quant à elle définie comme la quantité de CO₂ émise par unité d'énergie consommée. Elle est mesurée ici à travers la variable `co2_per_unit_energy`. Cet indicateur permet d'appréhender le caractère plus ou moins carboné du mix énergétique d'un pays, et donc son impact environnemental pour une quantité donnée d'énergie utilisée.

Sur le plan économique, ces indicateurs soulèvent plusieurs enjeux. D'une part, la consommation énergétique est traditionnellement corrélée à la croissance économique, l'énergie étant un facteur de production fondamental. D'autre part, une forte intensité carbone pose la question de la soutenabilité de cette croissance, en raison de son impact climatique. Dès lors, comprendre les interactions entre ces variables est essentiel pour orienter les politiques énergétiques dans un cadre de développement durable.

1.2 Relations théoriques

Les relations entre énergie, émissions de CO₂ et croissance économique ont suscité une abondante littérature théorique, mettant en évidence des mécanismes complexes et parfois ambigus.

L'un des cadres les plus étudiés est celui de la **courbe environnementale de Kuznets** (EKC, pour Environmental Kuznets Curve). Inspirée de la courbe de Kuznets

originelle sur les inégalités, cette hypothèse propose une relation en forme de U inversé entre le niveau de développement économique (souvent mesuré par le PIB par habitant) et les émissions polluantes. À un faible niveau de revenu, la croissance économique s’accompagne d’une dégradation environnementale, du fait d’une industrialisation rapide, d’une faible réglementation environnementale et d’une dépendance aux énergies fossiles. Mais au-delà d’un certain seuil de revenu, l’amélioration du capital humain, le développement des technologies propres et la pression sociale pour un environnement sain entraîneraient une diminution des émissions.

Parallèlement, le concept de **croissance verte** a émergé dans le cadre du développement durable. Il postule qu’il est possible de découpler la croissance économique de la dégradation environnementale, notamment par l’innovation technologique, l’amélioration de l’efficacité énergétique et l’adoption de sources d’énergie bas-carbone. Ce paradigme s’oppose à l’idée d’un arbitrage systématique entre prospérité économique et protection de l’environnement.

Des modèles théoriques plus récents, issus de la littérature en économie de l’énergie et en macroéconomie environnementale, intègrent explicitement l’énergie comme facteur de production, aux côtés du capital et du travail. Ces modèles permettent d’explorer les effets de chocs énergétiques, de la taxation carbone ou de la transition énergétique sur la trajectoire de croissance.

Ces différents cadres théoriques fournissent un socle pour notre étude empirique, qui vise à quantifier les liens entre variables énergétiques (consommation, émissions, intensité) et richesse des pays à travers une approche multidimensionnelle.

1.3 Travaux empiriques

Les relations entre consommation énergétique, émissions de gaz à effet de serre et croissance économique ont suscité une abondante littérature empirique. Traditionnellement, ces relations ont été étudiées à l’aide de modèles économétriques linéaires ou à effets fixes dans le cadre de données de panel. Plus récemment, les progrès en apprentissage automatique ont permis d’explorer des dynamiques non-linéaires et des interactions complexes.

Par exemple, Saidi and Hammami (2016) étudient un panel de 58 pays entre 1990 et 2012 et mettent en évidence une relation de causalité bidirectionnelle entre consommation énergétique et croissance économique. Leur analyse révèle également une courbe en U inversé entre le PIB par habitant et les émissions de CO₂, soutenant ainsi l’hypothèse d’une courbe environnementale de Kuznets.

Zhao, J. Wang, and N. Wang (2020) appliquent des algorithmes d’apprentissage automatique tels que le Random Forest et le support vector regression pour prédire à la fois la consommation énergétique et la croissance économique. Leurs résultats montrent que ces méthodes surpassent les approches économétriques classiques en précision prédictive, notamment dans des contextes de données bruitées et hétérogènes.

Dans une approche encore plus innovante, Lin and Liu (2023) proposent un modèle combinant analyse de réseaux sociaux et apprentissage automatique afin d’optimiser le contrôle de la pollution industrielle. Leur méthode synergique repose sur une intégration des données sociales avec les signaux environnementaux pour guider des politiques ciblées, soulignant le potentiel des modèles hybrides dans le champ énergie-environnement.

Enfin, Altınöz, Apergis, and Aslan (2020) utilisent des régressions quantiles en panel afin de montrer que les effets de la consommation énergétique et des émissions varient selon le niveau de développement des pays. Ils constatent qu’à des niveaux de revenu plus élevés, les économies semblent partiellement découpler croissance et pollution, sans

pour autant inverser complètement la tendance.

Ces études révèlent la complémentarité entre approches classiques et outils modernes de data science, et confirment la complexité des relations énergie-économie à l'échelle mondiale.

2 Données, transformations et analyse exploratoire

2.1 Présentation des données

L'analyse repose sur une base de données issue de la compilation réalisée par *Our World in Data*, largement utilisée pour les travaux empiriques portant sur le changement climatique, l'énergie et la croissance économique. Cette base intègre des données provenant de multiples sources, telles que la Banque mondiale, l'Agence internationale de l'énergie (AIE), le Global Carbon Project ou encore le Carbon Dioxide Information Analysis Center (CDIAC).

La période couverte par le jeu de données s'étend de 1850 à 2022, bien que la majorité des variables utilisées dans notre étude soient disponibles à partir des années 1960 pour un nombre significatif de pays. La couverture géographique comprend plus de 190 pays, permettant une analyse globale et comparative. Les identifiants pays sont normalisés selon le code ISO alpha-3 ('iso_code'), et chaque ligne du jeu de données correspond à une observation annuelle ('year') par pays ('country').

Variables principales. Parmi les variables disponibles, nous retenons principalement :

- **PIB (gdp)** : mesuré en dollars constants internationaux, il constitue la variable dépendante de notre étude ;
- **Population** (population) : nécessaire à la normalisation par habitant ;
- **Consommation d'énergie primaire** (primary_energy_consumption) : quantité totale d'énergie utilisée (toutes sources confondues) ;
- **Énergie par habitant** (energy_per_capita) et par point de PIB (energy_per_gdp) ;
- **Émissions de CO₂** (co2, co2_per_capita, co2_per_gdp) : totales et rapportées à la population ou au PIB ;
- **Émissions sectorielles** : émissions issues du charbon (coal_co2), du gaz (gas_co2), du pétrole (oil_co2), de la combustion torchère (flaring_co2), de l'industrie cimentière (cement_co2), ou encore du changement d'usage des sols (land_use_change_co2) ;
- **Gaz à effet de serre** : methane, nitrous_oxide, total_ghg, incluant ou excluant les changements d'usage des sols ;
- **Indicateurs d'intensité** : tels que co2_per_unit_energy, indiquant la quantité de CO₂ émise par unité d'énergie consommée.

La base contient également des variables cumulatives (`cumulative_co2`, etc.), des parts relatives mondiales par gaz ou par secteur (`share_global_co2`, etc.), et des estimations d'impact climatique (`temperature_change_from_co2`, etc.).

Un dictionnaire complet des variables est mis à disposition dans la documentation officielle de la base, précisant la définition, l'unité de mesure et les éventuelles transformations appliquées à chaque indicateur. Ce document constitue une ressource essentielle pour garantir une interprétation rigoureuse des indicateurs utilisés dans nos modèles.

Nous appliquons des transformations en valeurs par habitant lorsque cela est pertinent, et normalisons certaines variables pour éviter les biais liés à la taille économique ou démographique des pays. La prochaine section détaille ces étapes de prétraitement.

2.2 Prétraitement des données

L'analyse empirique repose sur des données issues d'un large panel spatio-temporel couvrant plusieurs décennies et plus de 180 pays. Afin d'assurer la cohérence statistique et l'interprétabilité des résultats, un important travail de prétraitement a été mené en amont.

Dans un premier temps, un filtrage des variables a été effectué pour ne conserver que les indicateurs essentiels à l'analyse, notamment : le PIB, les émissions de CO₂, la consommation énergétique par habitant, la population et les émissions totales de gaz à effet de serre (GES). Cependant, au cours des premiers essais de modélisation, certains résultats se sont révélés incohérents. En particulier, les méthodes d'interprétabilité par les valeurs de SHAP ont systématiquement mis en avant une importance excessive de la variable `population`, mais surtout de la variable `year` (année), cette dernière capturant en réalité une simple tendance temporelle plutôt qu'un véritable lien causal.

Cette surinterprétation des effets liés au temps ou à la taille démographique a justifié la décision de supprimer les variables `population` et `year` dans les modèles économétriques et d'apprentissage automatique ultérieurs. Ce choix vise à éviter les biais structurels et à garantir que les relations mises en évidence reflètent des mécanismes économiques et énergétiques sous-jacents, et non des artefacts temporels.

Les autres étapes de prétraitement ont inclus :

- la transformation des variables en valeurs par habitant lorsque cela était pertinent (énergie, CO₂, GES), afin de comparer équitablement les pays indépendamment de leur taille démographique ;
- la suppression des lignes présentant des valeurs manquantes pour les variables principales, garantissant ainsi la robustesse des modèles ;
- le redimensionnement éventuel de certaines grandeurs (ex. : passage en logarithme ou en unités normalisées) dans les phases ultérieures de modélisation.

2.3 Exploration des données

Une première visualisation globale à l'échelle mondiale met en lumière les tendances conjointes du PIB, des émissions de CO₂ totales et de la consommation énergétique par habitant. Contrairement à certains pays développés, on observe une croissance quasi-simultanée de ces trois indicateurs à l'échelle planétaire, traduisant les dynamiques économiques et énergétiques liées à l'industrialisation et à l'urbanisation des pays émergents (voir Figure 1).

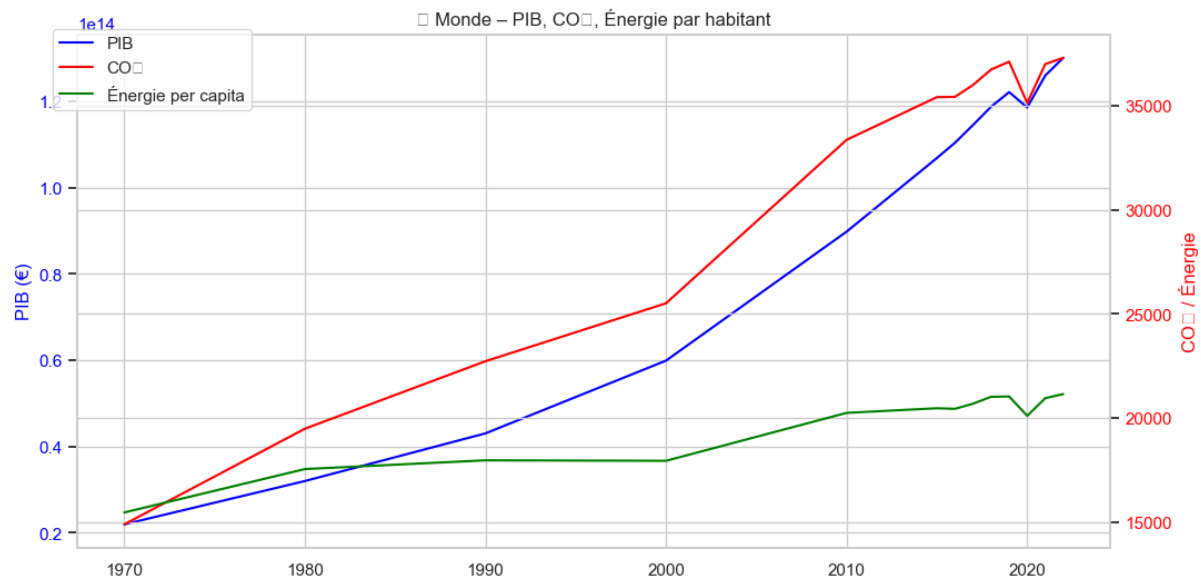


Figure 1: Évolution mondiale du PIB, des émissions de CO₂ et de la consommation énergétique par habitant

En se focalisant sur la France, les visualisations comparatives du PIB avec la consommation énergétique par habitant d'une part, et avec les émissions totales de CO₂ d'autre part, permettent d'appréhender les spécificités nationales.

Bien que le PIB continue de croître, les émissions de CO₂ totales stagnent, voire diminuent légèrement à partir des années 1990, témoignant d'une possible dissociation progressive entre croissance économique et impact carbone.

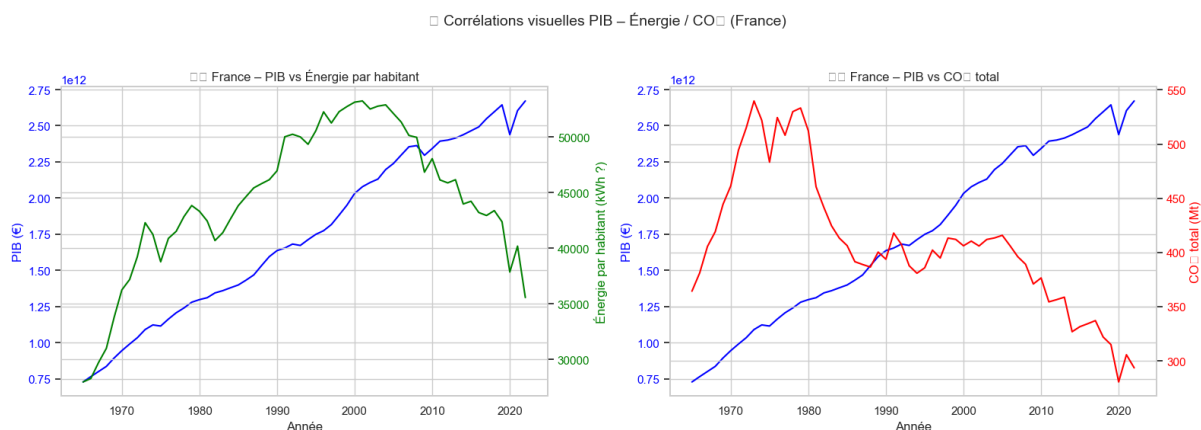


Figure 2: Évolution du PIB vs énergie par habitant (courbe verte) et PIB vs émissions totales de CO₂ (courbe rouge) en France

Pour compléter cette analyse, une matrice de corrélations calculée sur les principales variables françaises met en évidence la force des liens statistiques entre PIB, consommation énergétique, émissions de CO₂ et émissions totales de gaz à effet de serre (Figure 4). Les corrélations positives marquées confirment les interdépendances entre ces variables, tout en laissant apparaître une atténuation relative du lien entre PIB et émissions de CO₂ au fil du temps.

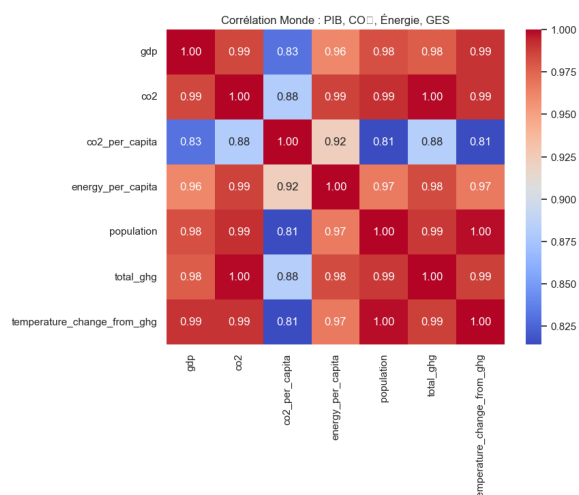


Figure 3: Matrice de corrélation entre PIB, CO₂, énergie et GES – Monde

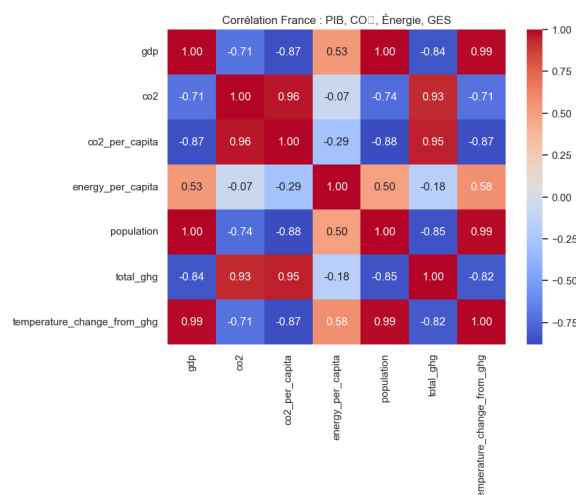


Figure 4: Matrice de corrélation entre PIB, CO₂, énergie et GES – France

Ces analyses exploratoires montrent que la croissance économique reste globalement liée à la consommation d'énergie et aux émissions de CO₂. Au niveau mondial, on observe une co-évolution nette entre le PIB, la consommation énergétique par habitant et les émissions de gaz à effet de serre, traduisant une dépendance persistante aux énergies fossiles dans le développement économique.

En se concentrant sur la France, les relations apparaissent plus nuancées. Les corrélations fortes entre les émissions de CO₂, la consommation d'énergie et le PIB indiquent un lien structurel, mais des tendances récentes laissent entrevoir une dissociation progressive entre croissance économique et émissions. Cela peut refléter des politiques de transition énergétique ou une amélioration de l'efficacité énergétique.

L'analyse de la matrice de corrélation révèle également des liens très forts entre certaines variables, comme la population et le PIB, ou les émissions de GES et les températures. Ces corrélations, bien que cohérentes, soulignent aussi le risque de redondance ou de confusion dans les modèles. C'est notamment en observant des résultats incohérents (comme une importance exagérée attribuée à l'année ou à la population dans les modèles explicatifs) que nous avons choisi de retirer ces variables des analyses économétriques et prédictives suivantes.

Ces résultats confirment donc la nécessité d'approches plus robustes, capables de mieux capter les dynamiques causales et de limiter les biais liés à la structure des données.

3 Modélisation et résultats

3.1 Modélisation économétrique

L'analyse économétrique se concentre exclusivement sur le cas français afin de mieux cerner les dynamiques nationales liant croissance, énergie et pollution. Par rapport à l'étape de prétraitement, plusieurs variables explicatives ont été ajoutées pour enrichir les modèles : émissions de CO₂ sectorielles (gaz, pétrole), GES totaux, consommation d'énergie primaire, intensité carbone du PIB, méthane par habitant, ou encore changement de température lié aux gaz à effet de serre. Nous avons également construit une variable de PIB par habitant.

Trois modèles linéaires multiples ont été estimés pour explorer différentes dimensions de la relation entre économie et environnement.

Modèle 1 – PIB total en fonction de CO₂ et de l'énergie par habitant

Ce premier modèle explore l'effet des émissions de CO₂ totales et de la consommation d'énergie par habitant sur le PIB total en France. Les deux variables explicatives sont statistiquement significatives ($p < 0,001$) et le modèle présente un coefficient de détermination élevé ($R^2 = 0,741$), ce qui indique qu'il explique une part substantielle de la variance du PIB.

Cependant, certains résultats méritent d'être interprétés avec prudence. Le signe négatif du coefficient de CO₂ ($-6,31 \times 10^9$) suggère une relation inverse inattendue avec le PIB, possiblement due à des effets de colinéarité ou à une dynamique inversée dans les données françaises récentes (où la croissance économique s'est parfois accompagnée d'une baisse des émissions). À l'inverse, la consommation d'énergie par habitant présente une relation positive forte avec le PIB, cohérente avec l'intuition selon laquelle une intensité énergétique élevée reflète une activité économique accrue.

Des limites existent néanmoins : les tests de normalité (Omnibus) et l'indice de Durbin-Watson très faible ($DW = 0,116$) indiquent une autocorrélation des résidus, et le condition number élevé ($4,44 \times 10^5$) laisse suspecter une potentielle multicolinéarité. Ces éléments fragilisent en partie la robustesse de l'inférence.

À l'échelle mondiale, le modèle indique une relation positive significative entre les émissions de CO₂ et le PIB, contrairement au cas français. L'énergie par habitant n'apparaît pas significative ici, suggérant qu'au niveau global, ce sont davantage les volumes totaux d'émissions qui sont associés à la croissance économique. Ce contraste met en évidence des dynamiques différentes entre pays industrialisés et agrégats globaux.

Résumé du modèle 1

Dep. Variable:	gdp	R-squared:	0.741
Model:	OLS	Adj. R-squared:	0.732
Method:	Least Squares	F-statistic:	78.82
Date:	Mon, 04 Aug 2025	Prob (F-statistic):	7.08e-17
Time:	12:22:51	Log-Likelihood:	-1614.1
No. Observations:	58	AIC:	3234.
Df Residuals:	55	BIC:	3240.
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
const	2.376e+12	3.94e+11	6.029	0.000	1.59e+12	3.17e+12
co2	-6.314e+09	6.42e+08	-9.842	0.000	-7.6e+09	-5.03e+09
energy_per_capita	4.393e+07	6.19e+06	7.099	0.000	3.15e+07	5.63e+07

Omnibus:	18.893	Durbin-Watson:	0.116
Prob(Omnibus):	0.000	Jarque-Bera (JB):	4.174
Skew:	-0.196	Prob(JB):	0.124
Kurtosis:	1.746	Cond. No.	4.44e+05

Modèle 2 – PIB par habitant en fonction des émissions sectorielles (gaz, pétrole) et GES

Le second modèle affine l'analyse en remplaçant la variable dépendante par le PIB par habitant, et en choisissant des variables explicatives représentant des émissions plus désagrégées de gaz à effet de serre. L'objectif est d'identifier si certains types d'émissions polluantes ont un impact plus marqué sur le PIB par habitant.

Les résultats montrent que, tout en contrôlant pour les émissions totales de gaz à effet de serre (total_ghg) qui ont un effet négatif significatif, les émissions de CO₂ liées au pétrole (oil_co2) et au gaz naturel (gas_co2) sont toutes deux positivement et significativement corrélées au PIB par habitant. Cela suggère une forte dépendance économique historique aux produits pétroliers et gaziers.

Le modèle présente un R-carré ajusté élevé (0.971), témoignant d'un bon ajustement global, même si la condition number relativement élevée indique une possible multicolinéarité entre variables explicatives, ce qui pourrait affecter la stabilité des coefficients.

Résumé du modèle 2

Dep. Variable:	gdp_per_capita	R-squared:	0.972
Model:	OLS	Adj. R-squared:	0.971
Method:	Least Squares	F-statistic:	632.8
Date:	Mon, 04 Aug 2025	Prob (F-statistic):	4.97e-42
Time:	12:35:07	Log-Likelihood:	-495.79
No. Observations:	58	AIC:	999.6
Df Residuals:	54	BIC:	1008.
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
const	3.102e+04	2121.738	14.619	0.000	2.68e+04	3.53e+04
total_ghg	-45.1706	6.077	-7.433	0.000	-57.354	-32.987
oil_co2	38.7064	9.125	4.242	0.000	20.411	57.001
gas_co2	199.6517	11.755	16.984	0.000	176.084	223.220

Omnibus:	3.619	Durbin-Watson:	0.594
Prob(Omnibus):	0.164	Jarque-Bera (JB):	3.520
Skew:	0.570	Prob(JB):	0.172
Kurtosis:	2.604	Cond. No.	7.39e+03

Modèle 3 – PIB par habitant en fonction de la population, de l'énergie primaire, du méthane et du changement climatique

Le modèle 3 étudie la relation entre le PIB par habitant et plusieurs variables explicatives : la population, la consommation d'énergie primaire, les émissions de méthane par habitant, ainsi que le changement de température lié aux gaz à effet de serre.

L'ajustement global du modèle est excellent, avec un R^2 ajusté de 0,993, indiquant que 99,3% de la variance du PIB par habitant est expliquée par les variables incluses. La statistique F (2030) est fortement significative (p-valeur < 0,001), confirmant la pertinence globale du modèle.

L'analyse des coefficients montre que la population ($coef = 0,0013$, $p < 0,001$), la consommation d'énergie primaire ($coef = 3,457$, $p < 0,001$) et les émissions de méthane par habitant ($coef = 4458$, $p = 0,012$) ont un effet positif et significatif sur le PIB par habitant. Ces résultats suggèrent qu'une augmentation de ces variables est associée à une croissance économique, ce qui peut refléter à la fois l'impact de l'industrialisation et des activités humaines énergivores.

En revanche, le coefficient associé au changement de température n'est pas statistiquement significatif ($p = 0,168$), ce qui indique qu'à l'échelle de ce modèle et de cet échantillon, cette variable n'a pas d'effet direct et significatif sur le PIB par habitant.

Cependant, il convient de noter que la condition number extrêmement élevée ($1,7 \times 10^{11}$) signale une forte multicolinéarité entre les variables explicatives, ce qui peut affecter la stabilité et la fiabilité des estimations des coefficients.

Enfin, les tests de normalité des résidus (*Omnibus* et *Jarque-Bera*) ne rejettent pas l'hypothèse de normalité, mais la statistique de Durbin-Watson (0,892) suggère une possible autocorrélation des erreurs à examiner plus en détail.

Résumé du modèle 3

Dep. Variable:	gdp_per_capita	R-squared:	0.994			
Model:	OLS	Adj. R-squared:	0.993			
Method:	Least Squares	F-statistic:	2030.			
Date:	Mon, 04 Aug 2025	Prob (F-statistic):	2.82e-57			
Time:	12:36:45	Log-Likelihood:	-453.72			
No. Observations:	58	AIC:	917.4			
Df Residuals:	53	BIC:	927.7			
Df Model:	4					
Covariance Type:	nonrobust					
	coef	std err	t	P > t	[0.025	0.975]
const	-6.964e+04	9173.544	-7.591	0.000	-8.8e+04	-5.12e+04
population	0.0013	0.000	9.958	0.000	0.001	0.002
primary_energy_consumption	3.4572	0.372	9.281	0.000	2.710	4.204
methane_per_capita	4458.1183	1708.140	2.610	0.012	1032.023	7884.213
temperature_change_from_ghg	3.393e+05	2.43e+05	1.397	0.168	-1.48e+05	8.26e+05
Omnibus:	0.797	Durbin-Watson:	0.892			
Prob(Omnibus):	0.671	Jarque-Bera (JB):	0.579			
Skew:	0.244	Prob(JB):	0.749			
Kurtosis:	2.964	Cond. No.	1.70e+11			

Validation des modèles

Des tests économétriques standards ont été appliqués à chacun des modèles :

- **Normalité des résidus** (test de Shapiro-Wilk) : les valeurs des p-values supérieures à 0.05 indiquent que les résidus sont globalement compatibles avec une distribution normale, ce qui valide l'utilisation des tests statistiques classiques sur les coefficients.
- **Homoscédasticité** (test de Breusch-Pagan) : ce test révèle que la variance des résidus n'est pas toujours constante. En particulier, le Modèle 2 présente une hétéroscédasticité significative (p-value < 0.05), suggérant que les erreurs ont une variance non homogène, ce qui peut biaiser les écarts-types et les tests de significativité.
- **Multicolinéarité** (Facteur d'inflation de la variance - VIF) : certaines variables, notamment la variable "population", présentent des indices de colinéarité élevés ($VIF > 5$) dans certains modèles. Cette multicolinéarité peut affecter la stabilité des estimations des coefficients, justifiant ainsi la suppression ou la transformation de ces variables dans les approches alternatives comme les modèles d'apprentissage automatique.

Tests de causalité de Granger

- **Énergie \rightarrow PIB** : aucune causalité significative n'a été détectée pour les différents retards testés (p-values > 0.5), ce qui indique que les variations passées de l'énergie ne prédisent pas de manière significative la croissance économique.
- **PIB \rightarrow énergie** : une causalité significative est observée pour tous les retards (p-values < 0.001), suggérant que le PIB influence fortement la consommation d'énergie.
- **CO₂ \rightarrow PIB** : aucun effet de causalité de Granger significatif n'a été détecté (p-values > 0.29), indiquant que les émissions de CO₂ ne prédisent pas directement la croissance économique.
- **PIB \rightarrow CO₂** : une causalité significative est observée pour tous les retards testés (p-values < 0.01), montrant que la croissance économique entraîne une augmentation des émissions de CO₂.

Ces résultats soulignent une relation unidirectionnelle où la croissance économique est un facteur moteur des variations de la consommation énergétique et des émissions de CO₂, tandis que l'inverse (énergie ou CO₂ causant le PIB) n'est pas statistiquement vérifié dans ces tests. Cette complexité justifie l'usage de méthodes complémentaires, notamment des approches d'apprentissage automatique (voir sous section suivante), pour explorer d'éventuelles non-linéarités et interactions plus subtiles.

3.2 Machine learning

Présentation et comparaison des modèles Ridge, Random Forest et XGBoost. Analyse de l'importance des variables.

3.3 Approche deep learning

Architecture utilisée, résultats obtenus, comparaison avec autres méthodes.

3.4 Comparaison France vs Monde

Analyse des différences dans les performances et interprétations.

Discussion, limites et perspectives

3.5 Synthèse des résultats

Points clés confirmés ou infirmés, implications pour la compréhension des liens énergie-croissance.

3.6 Limites

Qualité et granularité des données, absence de dimension causale, contraintes méthodologiques.

3.7 Perspectives

Propositions pour approfondir les analyses, intégrer des variables politiques, et modéliser les transitions énergétiques futures.

Conclusion générale

Résumé des apports majeurs du mémoire et réponse à la problématique. Ouverture vers des travaux futurs et implications pratiques.

Annexes

References

- Altınöz, Buket, Nicholas Apergis, and Alper Aslan (2020). “Energy consumption, carbon dioxide emissions and economic growth: Fresh evidence from panel quantile regressions”. In: *Energy Research Letters* 1.3, pp. 1–7. DOI: [10.46557/001c.17075](https://doi.org/10.46557/001c.17075).
- Lin, Yu-Cheng and Yiling Liu (2023). “Contrôle de la pollution industrielle basé sur l’intelligence artificielle : un modèle synergique utilisant l’analyse des réseaux sociaux et l’apprentissage automatique”. In: *Environmental Science and Pollution Research*. DOI: [10.1007/s11356-023-27030-4](https://doi.org/10.1007/s11356-023-27030-4).
- Saidi, Kais and Sami Hammami (2016). “Economic growth, energy consumption and carbon dioxide emissions: recent evidence from panel data analysis for 58 countries”. In: *Quality & Quantity: International Journal of Methodology* 50.1, pp. 361–383. DOI: [10.1007/s11355-014-0153-1](https://doi.org/10.1007/s11355-014-0153-1).
- Zhao, Jing, Jun Wang, and Ning Wang (2020). “Predicting energy consumption and economic growth using machine learning algorithms”. In: *Energy Reports* 6, pp. 1169–1177. DOI: [10.1016/j.egy.2020.05.007](https://doi.org/10.1016/j.egy.2020.05.007).