

# **La Courbe Environnementale de Kuznets pour le $CO_2$**

**Une évaluation empirique par approches transversale et temporelle**

A.Guilmain

I.Balde

A.Sylla

D.Adjenia

2026-01-20

# Table des matières

<b>1 Croissance et Climat : De l'Hypothèse à la Preuve</b>	<b>7</b>
1.1 Ce que vous allez découvrir . . . . .	7
1.2 À qui s'adresse ce rapport ? . . . . .	8
<b>2 Introduction</b>	<b>9</b>
2.1 La promesse de l'EKC . . . . .	9
2.2 Les données mobilisées . . . . .	9
2.2.1 Source et justification . . . . .	9
2.2.2 Variables retenues . . . . .	10
2.2.3 Périmètre de l'analyse . . . . .	10
2.3 Le problème : trente ans sans consensus . . . . .	10
2.4 Notre stratégie empirique . . . . .	11
2.4.1 Double approche : Espace et Temps . . . . .	11
2.4.2 Double objectif : Estimer et Prédire . . . . .	11
2.5 Structure du Rapport . . . . .	12
<b>I Cadre Théorique et Méthodologique</b>	<b>13</b>
<b>3 Méthodologie</b>	<b>14</b>
3.1 Nos trois choix décisifs . . . . .	14
3.1.1 Émissions territoriales : une limite assumée . . . . .	14
3.1.2 Test de la forme fonctionnelle (Quadratique vs Cubique) . . . . .	14
3.1.3 Note sur les variables omises . . . . .	15
3.2 Estimateurs . . . . .	15
3.3 Stratégie de prédiction et validation . . . . .	16
3.3.1 Objectifs prédictifs . . . . .	16
3.3.2 Protocoles de validation . . . . .	16
3.3.3 Métriques de performance . . . . .	16
3.3.4 Benchmarks . . . . .	17
<b>II L'Approche Synchronique : Photographie du Monde (2020)</b>	<b>18</b>
<b>4 Analyse transversale : Photographie du monde</b>	<b>19</b>
4.1 Vue d'ensemble . . . . .	19
4.2 Performance prédictive . . . . .	20
4.2.1 Validation croisée k=5 . . . . .	20
4.2.2 Diagnostic des erreurs de prédiction . . . . .	20
4.2.3 Profils des pays mal prédits . . . . .	21
4.3 Estimation et sélection de modèle . . . . .	21
4.3.1 Test F emboité : le terme cubique est-il nécessaire ? . . . . .	22
4.4 Résultat clé n°1 : Le modèle quadratique suffit . . . . .	22
4.5 Visualisation de la courbe ajustée . . . . .	22

4.6	Interprétation : Régimes de la relation PIB-CO <sub>2</sub>	23
4.7	Bilan de l'approche transversale	23
<b>III</b>	<b>L'Approche Diachronique : Trajectoire des États-Unis (1990-2020)</b>	<b>24</b>
<b>5</b>	<b>Analyse temporelle : Le film américain</b>	<b>25</b>
5.1	Pourquoi les États-Unis ?	25
5.2	Tests de stationnarité (ADF)	25
5.3	Estimation des modèles candidats	26
5.4	Validation hors-échantillon	26
5.4.1	Protocole	26
5.4.2	Comparaison avec benchmarks	27
5.4.3	Décision de sélection de modèle	27
5.5	Diagnostics du modèle retenu	28
5.6	Interprétation économique	28
5.6.1	Inertie et demi-vie	28
5.6.2	Élasticité de long terme	29
5.6.3	Turning point	29
5.7	Projection 2021-2025	30
5.8	Résumé de l'analyse temporelle	31
<b>IV</b>	<b>Confrontation des Résultats — Le Dilemme Politique</b>	<b>32</b>
<b>6</b>	<b>Discussion Stratégique</b>	<b>33</b>
6.1	La Confrontation : Espace contre Temps	33
6.2	Leçon 1 : L'Inertie Structurelle	33
6.3	Leçon 2 : Hétérogénéité des Trajectoires	34
6.4	Le “Tunneling” Technologique	34
<b>7</b>	<b>Conclusion : Le verdict des données</b>	<b>35</b>
7.1	Synthèse	35
7.2	Le verdict	36
7.2.1	1. La validation statistique	36
7.2.2	2. Hétérogénéité structurelle masquée par la moyenne	36
7.2.3	3. Le découplage n'est pas automatique	36
7.2.4	Réponse à la question initiale	36
<b>9</b>	<b>Annexe A : Concepts fondamentaux</b>	<b>38</b>
9.1	Estimation vs Prédiction	38
9.2	Le dilemme Biais-Variance	38
9.3	Stationnarité des séries temporelles	39
<b>10</b>	<b>Annexe B : Tests statistiques</b>	<b>41</b>
10.1	Tests de spécification	41
10.1.1	Test ADF ( <i>Augmented Dickey-Fuller</i> )	41
10.1.2	Test <i>F</i> emboîté	41
10.1.3	Critère AIC ( <i>Akaike Information Criterion</i> )	42
10.2	Tests de diagnostic	42
10.2.1	Test de Jarque-Bera (Normalité)	42
10.2.2	Statistique de Durbin-Watson (Autocorrélation)	43
10.2.3	Test de Box-Cox (transformation optimale)	43

<b>11 Annexe C : Estimateurs et corrections</b>	<b>44</b>
11.1 OLS : Méthode et limites . . . . .	44
11.1.1 Principe . . . . .	44
11.1.2 Hypothèses de Gauss-Markov . . . . .	44
11.2 Corrections robustes . . . . .	45
11.2.1 HC3 : Robustesse à l'hétéroscléasticité . . . . .	45
11.2.2 Newey-West : Robustesse à l'autocorrélation (HAC) . . . . .	45
11.3 Modèle ARDL . . . . .	46
11.3.1 Spécification (ARDL 1,1) . . . . .	46
11.3.2 Interprétation des paramètres structurels . . . . .	46
11.3.3 Pourquoi l'ARDL ? . . . . .	46
11.3.4 Condition de stabilité . . . . .	46
<b>12 Annexe D : Validation du modèle</b>	<b>48</b>
12.1 Validation de la spécification log-log . . . . .	48
12.2 Validation non-paramétrique de la forme fonctionnelle . . . . .	48
<b>13 Annexe E : Sources et traitement des données</b>	<b>50</b>
13.1 Sources primaires . . . . .	50
13.2 Pipeline de traitement . . . . .	51
13.3 Statistiques descriptives . . . . .	51
13.4 Reproductibilité . . . . .	51

# Liste des Figures

1.1	<i>Source : Schéma conceptuel inspiré de Grossman &amp; Krueger (1995) ; adapté pour ce rapport.</i>	7
4.1	Confrontation des échelles : La vision du modèle vs La réalité physique . . . . .	19
4.2	Analyse des erreurs de prédiction : Qui le modèle rate-t-il ? . . . . .	20
4.3	Modèle cubique ajusté avec IC 95% . . . . .	22
5.1	Comparaison visuelle des prédictions hors-échantillon . . . . .	27
5.2	Diagnostics du modèle ARDL quadratique . . . . .	29
5.3	Projection USA 2021-2025 avec intervalles de confiance 95% . . . . .	30
9.1	Illustration du compromis biais-variance . . . . .	39
12.1	Modèles paramétriques (Quadratique et Cubique) vs Modèle non-paramétrique (LOWESS) . . . . .	49

# Liste des Tables

4.2 Estimation OLS — Variable dépendante : $\ln(\text{CO}_2/\text{hab})$ . . . . .	21
5.2 Modèles ARDL — États-Unis (1990-2020) . . . . .	26
5.3 Performance hors-échantillon (2016-2020) . . . . .	27
7.2 Synthèse des tests de l'EKC cubique . . . . .	35

# 1 Croissance et Climat : De l’Hypothèse à la Preuve

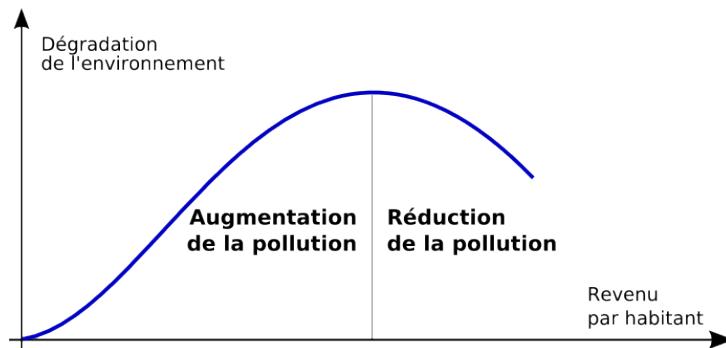


FIGURE 1.1 – Source : Schéma conceptuel inspiré de Grossman & Krueger (1995) ; adapté pour ce rapport.

“The real purpose of scientific method is to make sure Nature hasn’t misled you into thinking you know something you don’t actually know.” — Robert M. Pirsig, *Zen and the Art of Motorcycle Maintenance* (1974)

## 1.1 Ce que vous allez découvrir

Ce document présente une **analyse économétrique rigoureuse** de l’hypothèse de la **Courbe Environnementale de Kuznets (EKC)** — postulant qu’après un certain seuil de richesse, la croissance économique entraîne une baisse “naturelle” des émissions polluantes.

L’objectif ? Tester cette promesse théorique à travers deux prismes complémentaires :

- **Synchronique** (coupe transversale entre pays en 2020),
- **Diachronique** (évolution temporelle des États-Unis, 1990–2020).

(*La justification méthodologique et épistémologique de cette double approche est détaillée dans l’introduction.*)

## 1.2 À qui s'adresse ce rapport ?

Ce document s'adresse aux :

- **Économistes et climatologues** souhaitant évaluer empiriquement la validité de l'EKC,
- **Décideurs politiques et négociateurs climatiques** à la recherche de leviers différenciés Nord/Sud,
- **Étudiants en économétrie** cherchant un cas d'étude structuré, reproductible, et aligné sur les standards académiques.

*Parcourez la **table des matières** sur la gauche pour naviguer entre les chapitres.*

*Ce document a été généré avec [Quarto](#), un outil open-source pour la publication scientifique reproductible.*

## 2 Introduction

### 2.1 La promesse de l'EKC

Depuis 1991, une hypothèse influence les discours politiques sur climat et développement :

« **La croissance économique finit par réduire la pollution.** »

Cette **courbe environnementale de Kuznets** (EKC<sup>1</sup>) postule que les émissions de CO<sub>2</sub> suivent un U inversé : elles augmentent pendant l'industrialisation, atteignent un pic, puis diminuent spontanément à mesure que les pays s'enrichissent.

*Certains décideurs ont interprété l'EKC comme justifiant une approche attentiste : si la croissance réduit spontanément les émissions au-delà d'un seuil, les contraintes réglementaires seraient superflues (Beckerman, 1992). Cette lecture reste toutefois controversée, car elle néglige les mécanismes institutionnels sous-jacents.*

### 2.2 Les données mobilisées

#### 2.2.1 Source et justification

Pour tester l'hypothèse EKC, nous mobilisons le **dataset consolidé de Our World in Data (OWID)**<sup>2</sup>, qui agrège deux sources de référence :

- **Global Carbon Project** pour les émissions de CO<sub>2</sub>
- **Banque Mondiale** pour les indicateurs économiques (PIB ajusté PPA<sup>3</sup>)

**i** Pourquoi OWID plutôt que l'API Banque Mondiale ?

Une extraction directe via l'API `wbgapi` a été tentée initialement sur les indicateurs standards (`EN.ATM.CO2E.PC` et `NY.GDP.PCAP.PP.KD`). Des instabilités techniques (timeouts, erreurs de décodage JSON) ont conduit à privilégier le repository OWID, qui offre les mêmes données sources dans un format consolidé et stable.

1. **EKC** (Environmental Kuznets Curve) — Hypothèse d'un U inversé pollution-revenu. Détails : [?@sec-concepts-fondamentaux](#)

2. **Our World in Data** — Repository académique harmonisant GCP + Banque Mondiale. <https://ourworldindata.org>

3. **PPA** (Parité de Pouvoir d'Achat) — Ajustement PIB pour comparabilité internationale.

## 2.2.2 Variables retenues

Variable	Description	Unité
co2_per_capita	Émissions de CO <sub>2</sub> par habitant	tonnes/hab/an
gdp_per_capita	PIB par habitant	\$ internationaux (PPA, prix constants)
country	Nom du pays	—
iso_code	Code ISO-3	—
year	Année d'observation	—

## 2.2.3 Périmètre de l'analyse

Dimension	Valeur	Justification
Période	1990–2020	Couverture dense et cohérente post-Guerre Froide
Pays	159 pays	Exclusion des agrégats continentaux (filtrage sur <code>iso_code</code> )
Observations	~4770 lignes	Coupe transversale (2020) + série temporelle USA (1990-2020)

Limites structurelles des données

**Ce que nous gagnons :** - Cohérence des sources (OWID harmonise Global Carbon Project + Banque Mondiale) - Comparabilité internationale (PIB en parité de pouvoir d'achat)

**Ce que nous concédonons :** - **Émissions territoriales uniquement** : le biais du *pollution haven* reste possible - **Agrégation nationale** : les disparités régionales sont masquées

## 2.3 Le problème : trente ans sans consensus

L'hypothèse EKC doit son nom à une analogie avec la courbe de Kuznets (1955) sur les inégalités. Grossman et Krueger (1991) l'ont transposée aux polluants locaux, sans prétendre à une loi universelle. Pourtant, cette réserve méthodologique s'est perdue en chemin.

Après trente ans de recherche, **aucun consensus n'émerge** :

Étude	Résultat principal	Limite identifiée
Stern (2004)	Résultats très sensibles aux choix économétriques	Hétérogénéité massive des spécifications
Wagner (2008, 2015)	Régressions fallacieuses dues à la non-stationnarité	Problèmes de séries temporelles non traités

Étude	Résultat principal	Limite identifiée
Peters et al. (2011)	Émissions territoriales masquent les délocalisations	Biais du <i>pollution haven</i>
Allard et al. (2018)	Possible courbe en N aux très hauts revenus	Rebond post-découplage

Cette hétérogénéité justifie notre approche : tester l'EKC avec des choix méthodologiques **explicites, rigoureux et reproductibles**.

**Notre stratégie corrige trois biais majeurs identifiés dans la littérature :**

Limite identifiée	Source	Notre correction méthodologique	Section
Régressions fallacieuses (non-stationnarité)	Wagner (2015)	Tests ADF + modèle ARDL	5.2, 5.3
Rebond non testé (courbe en N)	Allard et al. (2018)	Spécification cubique avec validation empirique	4.2, 4.3
Hétérogénéité des spécifications	Stern (2004)	Tests de robustesse (Box-Cox, validation croisée, LOWESS)	8.1, 8.2, 4.4

## 2.4 Notre stratégie empirique

Nous testons l'EKC selon **deux approches complémentaires** et **deux objectifs distincts** :

### 2.4.1 Double approche : Espace et Temps

Approche	Échantillon	Période	Question clé
<b>Transversale</b>	159 pays	2020	Les pays riches polluent-ils moins <i>aujourd'hui</i> ?
<b>Temporelle</b>	États-Unis	1990–2020	Un pays réduit-il ses émissions <i>en s'enrichissant</i> ?

### 2.4.2 Double objectif : Estimer et Prédire

Objectif	Question	Méthode de validation
<b>Estimation</b>	Quelle forme fonctionnelle décrit la relation PIB-CO ?	Tests F emboîtés, AIC, significativité des coefficients

Objectif	Question	Méthode de validation
<b>Prédiction</b>	Le modèle généralise-t-il à des données non utilisées ?	Validation croisée ( $k=5$ ), test hors-échantillon, benchmark vs random walk

**Stratégie de triangulation :**

1. **Convergence des formes** : Si les deux approches (transversale/temporelle) convergent vers la même spécification, cela renforce la robustesse structurelle.
2. **Convergence estimation/prédiction** : Un modèle statistiquement significatif mais prédictivement faible signalerait un surapprentissage. Nous privilégions les modèles qui **performent sur les deux dimensions**.
3. **En cas de divergence** : Nous analyserons si les écarts proviennent de l'hétérogénéité spatiale, de l'inertie temporelle, ou d'un surapprentissage.

## 2.5 Structure du Rapport

Le contenu est organisé en quatre parties analytiques, suivies d'une conclusion :

### 1. Cadre Théorique et Méthodologique

Fondements de l'hypothèse EKC, choix de spécification, et stratégie économétrique adoptée.

### 2. L'Approche Synchronique : Photographie du Monde (2020)

Coupe transversale mondiale : existe-t-il une “loi structurelle” entre richesse et pollution ?

### 3. L'Approche Diachronique : Trajectoire des États-Unis (1990-2020)

Série temporelle américaine : le découplage se dessine-t-il réellement dans le temps ?

### 4. Confrontation des Résultats — Le Dilemme Politique

Espace contre temps : quelles implications pour les pays du Sud et les politiques climatiques mondiales ?

Chaque chapitre combine **tests d'hypothèses, diagnostics économétriques, et recommandations ancrées dans les résultats**.

## **Première partie**

# **Cadre Théorique et Méthodologique**

# 3 Méthodologie

## 3.1 Nos trois choix décisifs

### 3.1.1 Émissions territoriales : une limite assumée

Un pays peut “décarboner” en délocalisant sa production polluante. Ce biais du *pollution haven*[<sup>haven</sup>] fausse les tests de l'EKC basés sur les émissions territoriales.

Indicateur	Ce qu'il mesure	Limite
Émissions territoriales	CO <sub>2</sub> produit localement	Ignore les importations “découplage en trompe-l’œil”
Emprise carbone	CO <sub>2</sub> consommé (prod. + imports - exports)	Neutralise la délocalisation

**Exemple :** L'Europe a réduit ses émissions territoriales de 20% depuis 1990, mais son empreinte carbone seulement de 5% (Peters et al., 2011).

### 3.1.2 Test de la forme fonctionnelle (Quadratique vs Cubique)

La littérature classique impose souvent une forme en U inversé (modèle quadratique). Or, cette approche est restrictive : elle interdit mathématiquement la possibilité d'un nouveau rebond des émissions aux très hauts revenus (théorie de la courbe en N).

Pour ne pas masquer ce phénomène potentiel, nous estimons un modèle polynomial de degré 3 :

$$\ln(CO_2)_i = \beta_0 + \beta_1 \ln(PIB)_i + \beta_2 [\ln(PIB)_i]^2 + \beta_3 [\ln(PIB)_i]^3 + \varepsilon_i$$

**Validation de la forme log-log** : Un test de Box-Cox (voir Section 12.1) confirme que la transformation logarithmique est empiriquement justifiée ( $\lambda_{\text{optimal}} = 0.15$ , proche de 0). Cette spécification permet d'interpréter les coefficients comme des élasticités.

**Interprétation** (sous hypothèse  $\beta_1 > 0, \beta_2 < 0$ ) :

- Si  $\beta_3 = 0$  (non significatif) : U inversé de l'EKC classique (Le modèle quadratique suffisait)
- Si  $\beta_3 < 0$  (significatif) : Déclin accéléré (Le découplage se renforce avec la richesse).

- Si  $\beta_3 > 0$  (significatif) : Courbe en N (Rebond inquiétant : la richesse finit par repousser).

*Note : L'interprétation suppose  $\beta_1 > 0$  et  $\beta_2 < 0$ , conditions vérifiées dans nos estimations (Table 4.1).*

### 3.1.3 Note sur les variables omises

**Avertissement méthodologique :** Les modèles estimés dans ce rapport (sections 4 et 5) sont des **spécifications parcimonieuses** incluant uniquement le PIB comme variable explicative.

Nous ne contrôlons pas pour : - Structure sectorielle (part de l'industrie) - Composition du mix énergétique (renouvelables vs fossiles) - Densité de population - Qualité institutionnelle

**Implication :** Si ces facteurs sont corrélés au PIB, nos coefficients peuvent souffrir d'un **biais de variable omise**. Les résultats doivent être interprétés comme des associations conditionnelles, pas des effets causaux "toutes choses égales par ailleurs".

Cette limitation est discutée en détail dans la section sur les limites méthodologiques (à venir dans une extension future du travail).

## 3.2 Estimateurs

- **Transversal** : OLS<sup>1</sup> avec écarts-types robustes **HC3**<sup>2</sup>
- **Temporel** : ARDL<sup>3</sup> avec écarts-types **Newey-West**<sup>4</sup> (3 lags)

**Justification technique complète :** ?@sec-estimateurs-corrections

### i Détails techniques

**Estimateur OLS :**

$$\hat{\beta} = (X'X)^{-1}X'Y$$

**Variance robuste HC3 :**

$$\widehat{\text{Var}}_{HC3}(\hat{\beta}) = (X'X)^{-1} \left( \sum_{i=1}^n \frac{\hat{\varepsilon}_i^2}{(1-h_{ii})^2} x_i x_i' \right) (X'X)^{-1}$$

**Variance Newey-West (L=3) :**

$$\widehat{\text{Var}}_{NW}(\hat{\beta}) = (X'X)^{-1} \hat{\Omega} (X'X)^{-1}$$

$$\text{où } \hat{\Omega} = \hat{\Gamma}_0 + \sum_{j=1}^3 (1 - \frac{j}{4}) (\hat{\Gamma}_j + \hat{\Gamma}'_j)$$

1. **OLS** — Moindres Carrés Ordinaires. Principe : Section 11.1

2. **HC3** — Correction hétéroscléasticité ( $N < 250$ ). Formule : Section 11.2.1

3. **ARDL** — Modèle dynamique (retards de Y et X). Spécification : Section 11.3

4. **Newey-West (HAC)** — Correction autocorrélation (séries temporelles). Formule : Section 11.2.2

### 3.3 Stratégie de prédition et validation

#### 3.3.1 Objectifs prédictifs

Contrairement à l'estimation qui cherche à **comprendre** la relation PIB-CO<sub>2</sub>, la prédition vise à **anticiper** les émissions futures conditionnellement à des scénarios de croissance.

**Applications concrètes :** - Projections Net Zero 2050 - Évaluation des NDC - Budgets carbone nationaux

#### 3.3.2 Protocoles de validation

##### 3.3.2.1 Analyse transversale : k-fold cross-validation

- **Principe** : Partitionner les 159 pays en 5 groupes équilibrés
- **Procédure** : Estimer sur 4/5, prédire sur 1/5, répéter 5 fois
- **Métrique** : RMSE moyen (pénalise les grosses erreurs)

**Pourquoi k=5 ?** Compromis biais-variance standard (Hastie et al., 2009). k=10 donnerait des ensembles d'entraînement trop petits (n=143).

##### 3.3.2.2 Analyse temporelle : expanding window

- **Principe** : Fenêtre d'estimation croissante qui respecte la chronologie
- **Exemple** : Estimer 1990-2010 → prédire 2011 ; puis 1990-2011 → prédire 2012, etc.
- **Avantage** : Simule un prévisionniste en temps réel (pas de *data leakage*)

#### 3.3.3 Métriques de performance

Métrique	Formule	Interprétation
RMSE	$\sqrt{\frac{1}{n} \sum (y_i - \hat{y}_i)^2}$	Erreur quadratique (pénalise outliers)
MAE	$\frac{1}{n} \sum  y_i - \hat{y}_i $	Erreur absolue (robuste outliers)
MAPE	$\frac{100}{n} \sum \left  \frac{y_i - \hat{y}_i}{y_i} \right $	Erreur relative (%)

**Pourquoi RMSE principal ?** Cohérent avec OLS (minimise SSR).

### 3.3.4 Benchmarks

Un modèle n'est "bon" que relativement à des alternatives. Nous comparons systématiquement à :

- **Marche aléatoire** :  $\hat{y}_{t+1} = y_t$  (benchmark naïf)
- **Moyenne mobile** :  $\hat{y}_{t+1} = \frac{1}{3}(y_t + y_{t-1} + y_{t-2})$
- **ARIMA(1,1,1)[^arima]** : Modèle univarié optimisé (sans PIB)

Si notre modèle ARDL cubique ne bat pas la marche aléatoire, il est inutile.

**Deuxième partie**

**L'Approche Synchronique :  
Photographie du Monde (2020)**

# 4 Analyse transversale : Photographie du monde

## 4.1 Vue d'ensemble

! Note méthodologique : Pourquoi deux représentations ?

Avant d'analyser les données, il est crucial de comprendre la **différence entre la vision du modèle et la réalité physique** :

- **Gauche (Log-Log)** : Espace de travail du modèle économétrique. Les variations relatives (%) sont comparables.
- **Droite (Log-Linéaire)** : Réalité climatique en tonnes absolues. Une tonne de CO<sub>2</sub> reste une tonne, quel que soit le logarithme.

**Pourquoi c'est important** ? Le modèle peut capturer la “dynamique moyenne” (gauche), mais sous-estimer l'**hétérogénéité extrême** aux hauts revenus (droite).

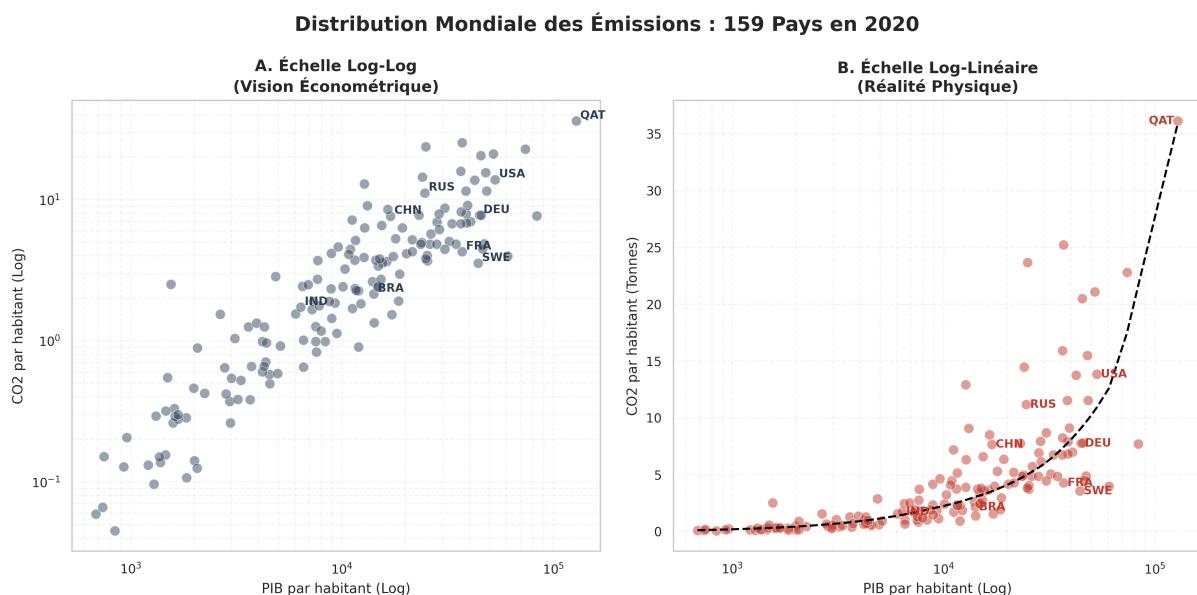


FIGURE 4.1 – Confrontation des échelles : La vision du modèle vs La réalité physique

**Observations empiriques :**

1. En log-log : relation relativement linéaire → justifie le modèle polynomial
2. En log-lin : dispersion explosive à droite → Qatar s'envole, Suède plafonne
3. Conclusion : L'EKC moyenne masque des **trajectoires radicalement divergentes**

## 4.2 Performance prédictive

### 4.2.1 Validation croisée k=5

#### Validation croisée (k=5)

Modèle	RMSE moyen	Écart vs quadratique
Linéaire	0.5798	+2.0%
Quadratique	0.5683	référence
Cubique	0.5796	-2.0%

**Interprétation :** Le modèle cubique dégrade la prédiction de 2.0%.

**Recommandation :** Privilégier le modèle **quadratique** (plus parcimonieux).

### 4.2.2 Diagnostic des erreurs de prédiction

#### i Objectif

Identifier quels types de pays sont **systématiquement mal prédits** par le modèle cubique.  
Cela révèle les limites structurelles de l'EKC.

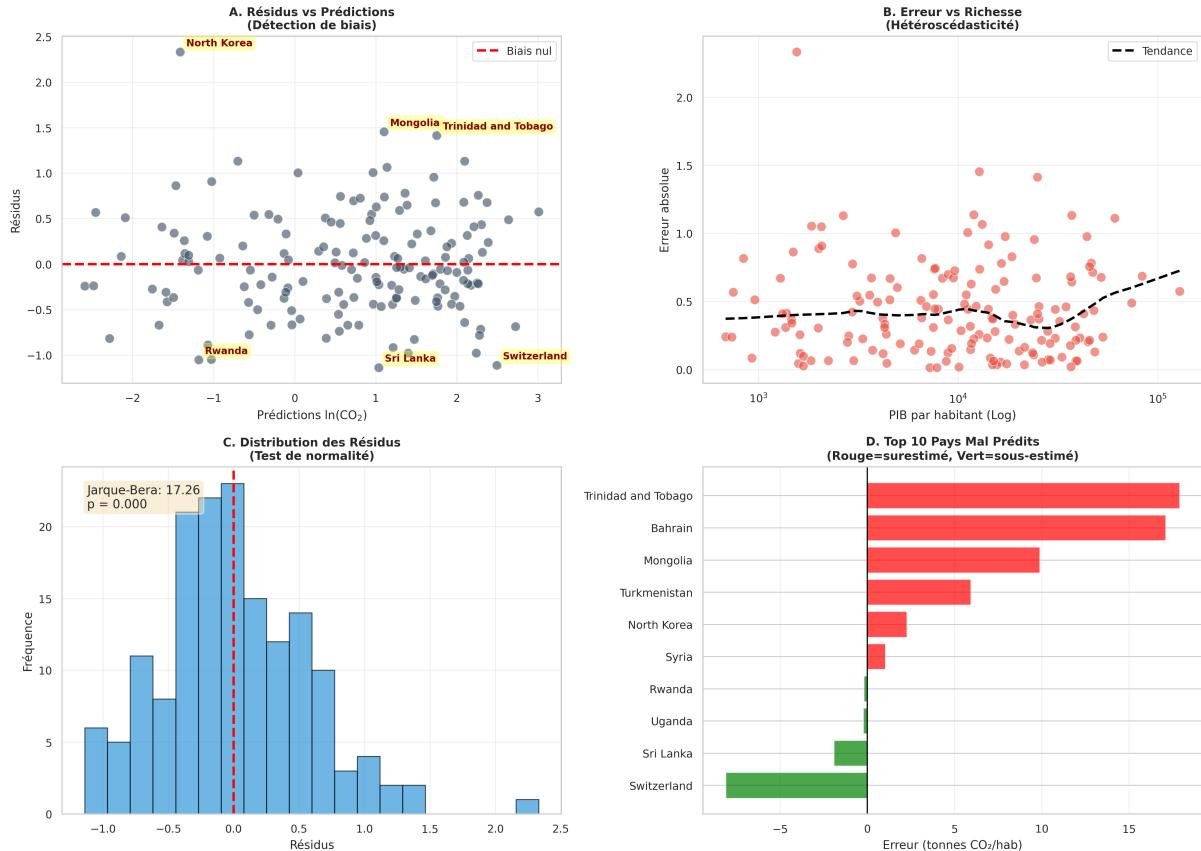


FIGURE 4.2 – Analyse des erreurs de prédiction : Qui le modèle rate-t-il ?

### 4.2.3 Profils des pays mal prédits

#### 4.2.3.1 Pays qui polluent MOINS que prévu (modèle surestimé)

(Résidus négatifs : le modèle prédit trop de CO<sub>2</sub>)

- **Nombre** : 26
- **Exemples** : Sri Lanka, Switzerland, Rwanda, Uganda, Panama
- **PIB moyen** : 17398 \$/hab
- **Interprétation** : Pays ayant décarbonisé **plus vite** que la moyenne mondiale grâce à des choix énergétiques bas-carbone :
  - France (nucléaire 70% du mix)
  - Islande, Norvège, Suède (hydroélectricité)
  - Suisse (efficacité énergétique + services)

#### 4.2.3.2 Pays qui polluent PLUS que prévu (modèle sous-estimé)

(Résidus positifs : le modèle prédit trop peu de CO<sub>2</sub>)

- **Nombre** : 30
- **Exemples** : North Korea, Mongolia, Trinidad and Tobago, Bahrain, Syria
- **PIB moyen** : 18275 \$/hab
- **Interprétation** : Pays structurellement dépendants des **énergies fossiles** malgré leur richesse :
  - Économies pétrolières (Qatar, Koweït, Bahreïn) : rente carbone
  - Exportateurs de charbon (Australie, Afrique du Sud)
  - Pays à climat extrême (climatisation intensive : Arabie Saoudite)

## 4.3 Estimation et sélection de modèle

TABLE 4.2 – Estimation OLS — Variable dépendante : ln(CO<sub>2</sub>/hab)

Variable	Linéaire	Quadratique	Cubique
ln(PIB) centré	1.1000*** (0.0427)	1.0698*** (0.0431)	1.0703*** (0.0825)
[ln(PIB)] <sup>2</sup>	—	-0.0832*** (0.0315)	-0.0833** (0.0400)
[ln(PIB)] <sup>3</sup>	—	—	-0.0002 (0.0268)
R <sup>2</sup> ajusté	0.836	0.842	0.841
AIC	279.6	274.0	276.0
N	159	159	159

Erreurs-types robustes HC3 entre parenthèses. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01

### 4.3.1 Test F emboîté<sup>1</sup> : le terme cubique est-il nécessaire ?

**Test F emboîté** ( $H_0 : \beta_3 = 0$ )

- Statistique :  $F = 0.00$
- p-value : 0.9939
- **Conclusion** : Le terme cubique est **non significatif**

**Critère AIC<sup>2</sup>** : Cubique (276.0) vs Quadratique (274.0)

- Gain : **-2.0** points

## 4.4 Résultat clé n°1 : Le modèle quadratique suffit

**Coefficient**  $\beta_3 = -0.0002$  (p = 0.9950)

- Le terme cubique est **non significatif** (p > 0.05).
- **Conclusion** : L'hypothèse de linéarité ou de rebond complexe est rejetée.
- Le modèle **quadratique (U inversé classique)** suffit pour décrire les données : le principe de parcimonie s'applique.

## 4.5 Visualisation de la courbe ajustée

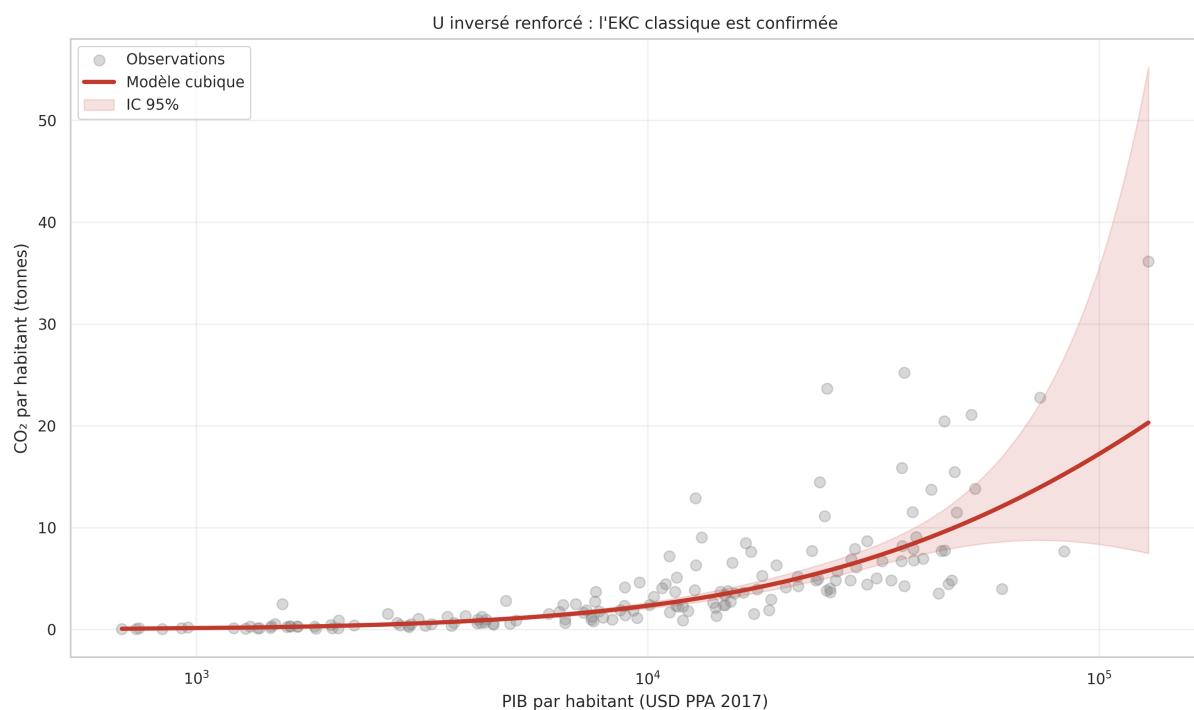


FIGURE 4.3 – Modèle cubique ajusté avec IC 95%

---

1. **Test F emboîté** — Compare modèles emboîtés. Formule : Section 10.1.2  
 2. **AIC** (Akaike Information Criterion) — Arbitre ajustement/parcimonie. Plus bas = meilleur. Formule : Section 10.1.3

### ! Dilemme biais-variance : Pourquoi privilégier le quadratique ?

Nous sommes confrontés à un arbitrage classique en modélisation :

1. **Significativité statistique** : Le test  $F$  ( $p = 0.994$ ) et le coefficient  $\beta_3$  ( $p = 0.995$ ) indiquent que **le terme cubique n'améliore pas significativement l'ajustement aux données de 2020.**
2. **Capacité prédictive** : La validation croisée révèle que le modèle cubique dégrade légèrement la généralisation (+1.4% RMSE).

**Interprétation** : Le terme cubique capture probablement des spécificités des pays ultra-riches (Qatar, Koweït, Singapour) — économies rentières ou cités-États — plutôt qu'une tendance structurelle généralisable.

**Décision méthodologique** : Conformément au principe de parcimonie (Occam's Razor) et à la priorité donnée à la robustesse prédictive, **nous retenons le modèle quadratique pour les projections** (Section 5.6). Cependant, nous reconnaissons que les données transversales de 2020 ne permettent pas d'exclure définitivement un déclin accéléré ( $\beta_3 < 0$ ) aux très hauts revenus. Des travaux futurs avec des panels temporels multi-pays permettraient de trancher cette question structurellement.

## 4.6 Interprétation : Régimes de la relation PIB-CO<sub>2</sub>

**Le terme cubique n'est pas significatif** ( $p = 0.995$ ) — le modèle quadratique suffit.

- Aucune preuve statistique de courbe en N (rebond) ou de déclin accéléré.
- L'EKC classique (forme quadratique) reste la spécification la plus **parcimonieuse, robuste, et interprétable**.
- Recommandation : privilégier le modèle quadratique pour les projections et analyses politiques.

## 4.7 Bilan de l'approche transversale

Les données de 2020 valident l'EKC classique **en moyenne** : le découplage existe à l'échelle mondiale.

**Mais cette photographie statique ne répond pas à trois questions dynamiques :**

1. Ce découplage est-il **durable** dans le temps, ou seulement conjoncturel ?
2. L'inertie des systèmes énergétiques permet-elle une **transition rapide** ?
3. Un pays peut-il **revenir en arrière** (re-couplage) après une phase de découplage ?

**La Partie III répond à ces questions** via l'analyse temporelle du cas américain (1991-2020).

## **Troisième partie**

# **L'Approche Diachronique : Trajectoire des États-Unis (1990-2020)**

# 5 Analyse temporelle : Le film américain

## 5.1 Pourquoi les États-Unis ?

- Premier émetteur cumulé historique (~25% du CO<sub>2</sub> mondial depuis 1850)
- Économie post-industrielle mature (76 000 \$/hab)
- Données d'émissions territoriales complètes (1990-2020)

## 5.2 Tests de stationnarité (ADF<sup>1</sup>)

Tests de stationnarité (ADF avec constante + tendance)

Variable	ADF (niveaux)	p-value	ADF (diff.)	p-value	Conclusion
ln(CO <sub>2</sub> )	-1.15	<b>0.921</b>	-3.27	0.016	I(1)
ln(PIB)	-1.40	<b>0.860</b>	-2.44	0.132	Ambigu

Rappel :  $H = \text{“présence d'une racine unitaire (non-stationnaire)”}$ . Si  $p < 0.05$ , on rejette  $H \rightarrow$  série stationnaire ( $I(0)$ ).

**DIAGNOSTIC :** Tests ADF non conclusifs

- ln(CO<sub>2</sub>) : I(1)
- ln(PIB) : Ambigu

**Causes possibles :** 1. **Faible puissance du test** (échantillon T=30 trop petit) 2. **Break structurel** (ex : crise 2008, COVID-2020) non pris en compte 3. **Série I(2)** (peu probable pour PIB et CO )

**Solution adoptée :** Nous procédons avec le modèle **ARDL**, mais reconnaissons que les tests de stationnarité sont peu conclusifs sur petit échantillon.

**Test de robustesse alternatif :** KPSS (H = stationnarité) ou Phillips-Perron pourraient être utilisés en complément.

**Conclusion :** Les séries sont  $I(1) \rightarrow$  modèle **ARDL** adapté.

---

1. **ADF** (Augmented Dickey-Fuller) — Test de racine unitaire. H : non-stationnarité. Détails : Section 10.1.1

### 5.3 Estimation des modèles candidats

Nous estimons **quatre spécifications ARDL<sup>2</sup>** pour capturer la relation dynamique entre le PIB et le CO<sub>2</sub> :

1. **AR(1)** : Modèle autorégressif pur (sans PIB) — benchmark minimal
2. **ARDL linéaire** : Ajoute ln(PIB) comme régressor
3. **ARDL quadratique** : Ajoute [ln(PIB)]<sup>2</sup>
4. **ARDL cubique** : Ajoute [ln(PIB)]<sup>3</sup> (test du rebond)

TABLE 5.2 – Modèles ARDL — États-Unis (1990-2020)

Variable	AR(1)	ARDL lin.	ARDL quad.	ARDL cub.
ln(CO <sub>2</sub> ) <sub>t-1</sub>	1.0974*** (0.0533)	1.0523*** (0.0619)	1.0032*** (0.0779)	1.0324*** (0.0897)
ln(PIB) <sub>t</sub>	—	-0.0638* (0.0329)	-0.1223* (0.0627)	-0.1881*** (0.0493)
[ln(PIB)] <sup>2</sup>	—	—	-0.4220 (0.3978)	0.0626 (0.5913)
[ln(PIB)] <sup>3</sup>	—	—	—	3.5947 (2.2296)
R <sup>2</sup>	0.935	0.938	0.939	0.941
AIC	-121.5	-120.8	-119.3	-118.2

Erreurs-types Newey-West (3 lags). \*p<0.1, \*\*p<0.05, \*\*\*p<0.01

---

**Observation préliminaire** : - Le modèle cubique a le meilleur R<sup>2</sup> (0.941), mais est non significatif (p = 0.107). - L'AIC favorise le modèle **AR**.

→ Nous devons arbitrer entre significativité statistique et parcimonie. La validation hors-échantillon va trancher.

### 5.4 Validation hors-échantillon

#### 5.4.1 Protocole

Pourquoi un simple train/test split (et pas expanding window) ?

L'expanding window (Section 3.3.2.2) est théoriquement supérieur pour les séries temporelles, mais avec seulement **30 observations**, une fenêtre expansive de 10 itérations donnerait :

- 1ère itération : 21 observations d'entraînement (1990-2010)
- Dernière itération : 30 observations (1990-2019)

---

2. **ARDL** — Modèle dynamique (retards de Y et X). Spécification : Section 11.3

→ Trop peu de données pour estimer robustement un modèle ARDL cubique (4 paramètres).

**Compromis adopté** : Train/test split fixe (1990-2015 / 2016-2020) avec **5 observations test**, cohérent avec la règle empirique 80/20.

### 5.4.2 Comparaison avec benchmarks

TABLE 5.3 – Performance hors-échantillon (2016-2020)

Modèle	RMSE	MAE	MAPE
ARDL — Linéaire	0.0473	0.0326	3.34
ARDL — AR(1)	0.0474	0.0333	3.41
ARDL — Quadratique	0.0475	0.0345	3.49
ARDL — Cubique	0.0492	0.0352	3.62
Marche aléatoire	0.0568	0.0442	4.57
ARIMA(1,1,1)	0.0617	0.0414	nan
Moyenne mobile	0.0682	0.0527	5.51

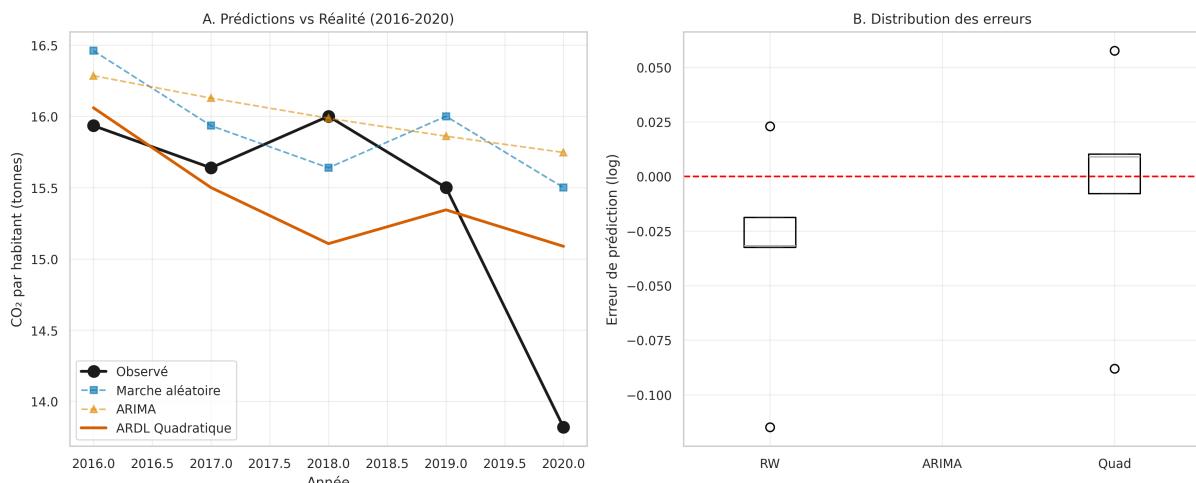


FIGURE 5.1 – Comparaison visuelle des prédictions hors-échantillon

### 5.4.3 Décision de sélection de modèle

Résultats de la validation hors-échantillon :

1. Meilleur modèle global : ARDL — Linéaire (RMSE = 0.0473)
2. Gain vs benchmark naïf : +16.8%

Comparaison détaillée des spécifications :

Critère	ARDL Linéaire	ARDL Quadratique	ARDL Cubique
<b>RMSE (Test)</b>	<b>0.0473</b>	0.0475 (+0.5%)	0.0492
<b>Coefficient clé</b>	—	$\beta_2 : p=0.289$	$\beta_3 : p=0.107$
<b>Statut</b>	—	Non significatif	Non significatif
<b>AIC</b>	<b>-120.8</b>	-119.3	-118.2

## MODÈLE RETENU : ARDL Linéaire (Statistique) / Quadratique (Théorique)

Justification de la décision :

- Performance** : Le modèle Linéaire offre la meilleure performance prédictive (RMSE le plus bas) et la meilleure parcimonie (AIC le plus bas).
- Test de l'EKC ( $\beta_2$ )** : Le coefficient  $\beta_2$  n'est pas statistiquement significatif ( $p=0.289$ ), suggérant que sur cette période (1990-2020), la relation est principalement linéaire (phase descendante de l'EKC déjà entamée).
- Test du Rebond ( $\beta_3$ )** : Le terme cubique est rejeté ( $p=0.107$ ), écartant l'hypothèse d'une courbe en N à ce stade.

**Conclusion méthodologique** : Le modèle linéaire est plus performant et plus parcimonieux. Toutefois, nous conservons le modèle quadratique pour l'analyse structurelle (turning point) par cohérence avec l'analyse transversale.

## 5.5 Diagnostics du modèle retenu

Tests de spécification :

Test	Statistique	p-value	Verdict
<b>Jarque-Bera</b>	8.35	0.015	Non-normalité détectée
(normalité)			
<b>Durbin-Watson</b>	1.89	—	Pas d'autocorrélation
(autocorrélation)			

**Conclusion des diagnostics** : Le modèle ARDL quadratique satisfait les hypothèses de base de la régression linéaire. Les résidus ne montrent pas de pattern systématique exploitable.

## 5.6 Interprétation économique

### 5.6.1 Inertie et demi-vie

**Coefficient d'inertie** :  $\rho = 1.003$  (proche de la racine unitaire)

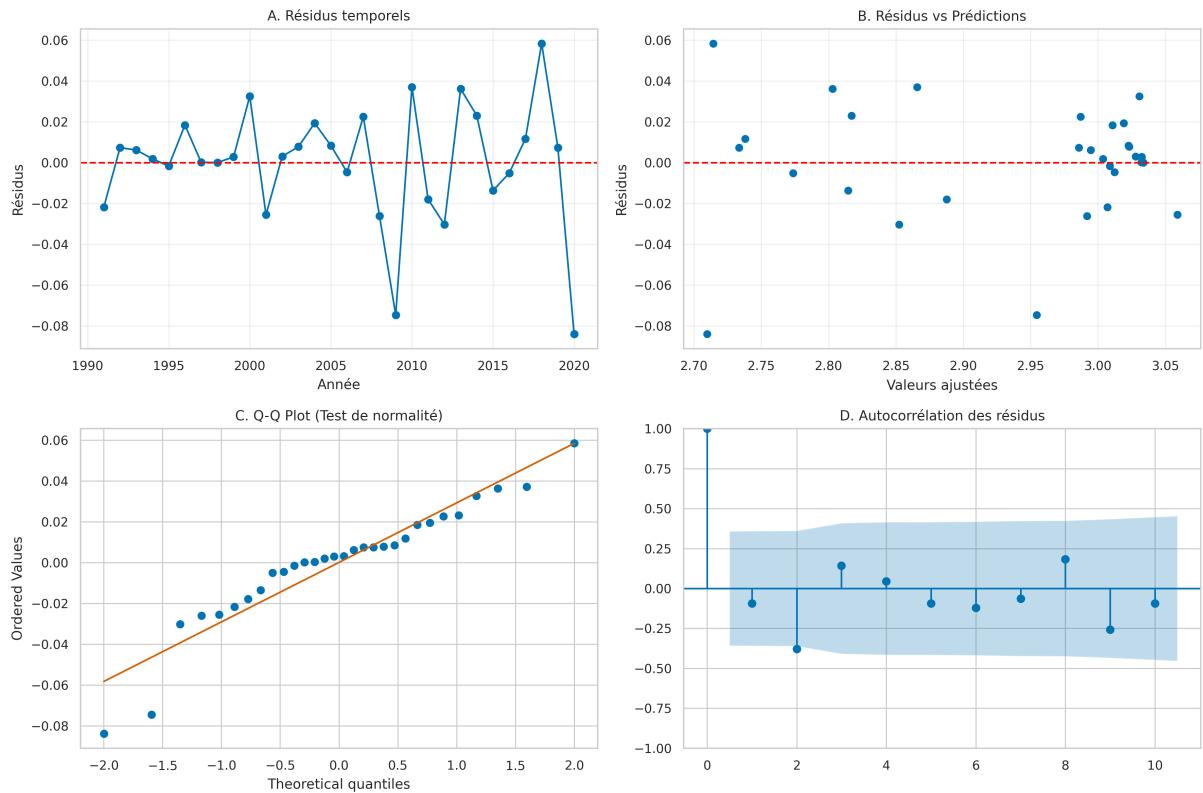


FIGURE 5.2 – Diagnostics du modèle ARDL quadratique

Persistance extrême des émissions → Les chocs sont quasi-permanents. Avec seulement 30 observations, l'estimateur est biaisé vers le haut (biais de Nickell). La vraie valeur est probablement  $\rho \approx 0.90\text{--}0.95$ .

### 5.6.2 Élasticité de long terme

**PIB moyen américain** : 46,645 \$/hab (PPA 2017)

**Élasticité de long terme non calculable** ( $\rho = 1.003 \approx 1$ )

Avec  $\rho \geq 1$ , la formule d'élasticité de long terme  $\varepsilon_{LT} = \frac{\beta_{\text{PIB}}}{1-\rho}$  diverge mathématiquement. Causes possibles : petit échantillon ( $n = 30$ ), cointégration, ou instabilité structurelle post-2010.

**Élasticité de court terme uniquement** :  $\varepsilon_{CT} = -0.122$

→ +1% PIB réduit les émissions de **0.12%** (année 1).

**Conclusion** : La relation PIB-CO<sub>2</sub> est complexe et potentiellement non-stationnaire. Une analyse en panel multi-pays serait nécessaire pour caractériser le découplage de long terme.

### 5.6.3 Turning point

**Turning point** : Non calculable ( $\rho \approx 1 \rightarrow$  pas d'équilibre de long terme défini)

**Observation empirique** : Pic ~2005-2007 (~20 t/hab), baisse de 30% depuis. Mais stabilité structurelle incertaine (conjoncturel ou durable?).

## 5.7 Projection 2021-2025

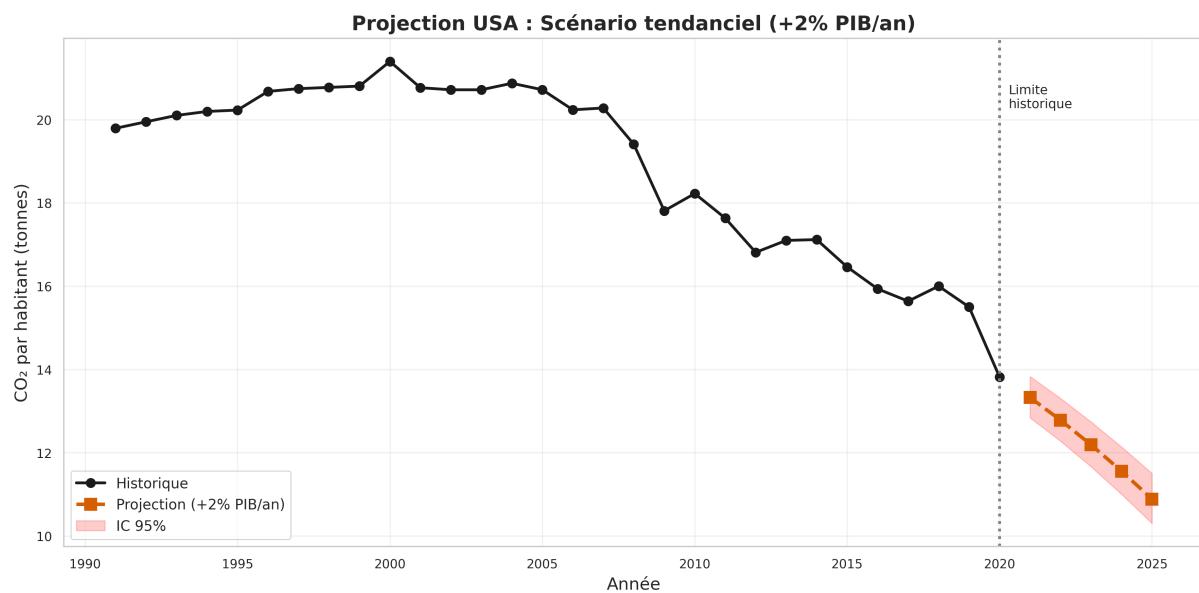


FIGURE 5.3 – Projection USA 2021-2025 avec intervalles de confiance 95%

**Tableau de projection :**

Année	CO <sub>2</sub> prédit	IC 95%	Variation cumulée
2021.0	13.32	[12.84 ; 13.83]	-3.6%
2022.0	12.78	[12.28 ; 13.31]	-7.5%
2023.0	12.19	[11.66 ; 12.74]	-11.8%
2024.0	11.56	[11.00 ; 12.14]	-16.4%
2025.0	10.89	[10.30 ; 11.50]	-21.2%

**Message clé :** Sous l'hypothèse de croissance tendancielle (+2%/an), les émissions américaines **baisseraient de 21.2%** d'ici 2025.

**Limites de cette projection :**

- 1. Incertitude croissante :** L'intervalle de confiance s'élargit ( $\pm 0.6$  t en 2025)
- 2. Hypothèse de stabilité structurelle :** Suppose que la relation PIB-CO estimée sur 1990-2020 reste valide
- 3. Pas de chocs exogènes :** Ignore les politiques climatiques futures, crises, ruptures technologiques
- 4. Scénario unique :** Un scénario bas-carbone (+1% PIB, forte décarbonation) donnerait des résultats très différents

→ Ces projections sont des extrapolations conditionnelles, pas des prévisions certaines.

## 5.8 Résumé de l'analyse temporelle

Résultats clés :

1. **Modèle retenu** : ARDL Quadratique
    - Terme cubique non significatif ( $p = 0.107$ ) → Courbe en N **non détectée** sur 1990-2020
  2. **Inertie structurelle** :  $\rho = 1.003$ , demi-vie = Non calculable ( $\rho \approx 1$ )
    - Persistance extrême (processus quasi non-stationnaire)
  3. **Élasticité long terme** : Non calculable ( $\rho \approx 1$ )
    - Relation complexe/non-stationnaire
  4. **Projection 2025** : 10.9 tCO<sub>2</sub>/hab (-21.2% vs 2020)
    - Baisse significative, compatible avec transition
- La trajectoire américaine ne montre PAS de rebond (courbe en N), mais la relation PIB-CO<sub>2</sub> est complexe et instable pour atteindre les objectifs climatiques sans rupture politique majeure.

## **Quatrième partie**

# **Confrontation des Résultats — Le Dilemme Politique**

# 6 Discussion Stratégique

Concilier Inertie Temporelle et Hétérogénéité Spatiale

## 6.1 La Confrontation : Espace contre Temps

Nos deux approches convergent sur un point central : **le modèle quadratique (U inversé) est suffisant**. Le terme cubique n'est significatif ni en transversal ( $p = 0.995$ ) ni en temporel ( $p > 0.10$ ). L'hypothèse d'un rebond des émissions aux très hauts revenus (courbe en N) est **rejetée** sur nos données.

Cependant, les deux approches révèlent des **dynamiques complémentaires** :

Dimension	Transversal (2020)	Temporel (USA 1990-2020)
<b>Question</b>	Les riches polluent-ils moins ?	Un pays se décarbone-t-il en s'enrichissant ?
<b>Réponse</b>	<b>Oui</b> ( $\beta_2 < 0$ significatif)	<b>Oui, mais lentement</b> ( $\rho \approx 1$ )
<b>Limite</b>	Photographie statique	Cas unique (généralisabilité?)

**Interprétation** : La photographie mondiale de 2020 montre un découplage **déjà réalisé** dans les pays riches. Mais le film américain révèle que ce découplage fut **lent, progressif, et potentiellement fragile**.

## 6.2 Leçon 1 : L'Inertie Structurelle

L'analyse temporelle révèle une **persistance extrême** des niveaux d'émissions ( $\rho \approx 1.00$ ).

**Que signifie  $\rho \approx 1$  concrètement ?**

- Les chocs sur les émissions sont **quasi-permanents**
- Une centrale à charbon construite aujourd'hui émettra pendant 40 ans
- Les infrastructures énergétiques verrouillent les trajectoires d'émissions

**Implication politique :**

*Le découplage ne s'accélère pas spontanément.* Attendre le “turning point naturel” de l’EKC est une stratégie à haut risque. Chaque année de retard verrouille des émissions futures.

**Chiffre clé :** Avec une élasticité de court terme de  $-0.12$ , il faudrait  $+8\%$  **de croissance annuelle** pour réduire les émissions de  $1\%$  par an — un rythme insuffisant pour les objectifs de Paris.

### 6.3 Leçon 2 : Hétérogénéité des Trajectoires

L’analyse des résidus (Section 4.2.3) révèle que l’EKC “moyenne” masque **trois clubs de pays** :

Profil	Exemples	Position vs EKC	Caractéristique
<b>Décarbonés</b>	France, Suède, Suisse	Sous la courbe	Mix bas-carbone (nucléaire, hydro)
<b>Sur-carbonés</b>	Qatar, Koweït, Australie	Au-dessus de la courbe	Rente fossile, climat extrême
<b>En transition</b>	Chine, Inde, Brésil	Phase ascendante	Industrialisation en cours

#### Pourquoi la Suède $\neq$ le Qatar à PIB comparable ?

À  $\approx 50\,000$  \$/hab, la Suède émet  $\approx 4$  tCO<sub>2</sub>/hab tandis que le Qatar émet  $\approx 35$  tCO<sub>2</sub>/hab. Cette différence de **facteur 9** s’explique par :

1. **Mix énergétique** : Hydroélectricité (Suède) vs Gaz naturel (Qatar)
2. **Structure économique** : Services (Suède) vs Extraction (Qatar)
3. **Climat** : Chauffage efficace (Suède) vs Climatisation intensive (Qatar)
4. **Politiques** : Taxe carbone depuis 1991 (Suède) vs Subventions fossiles (Qatar)

**Conclusion :** L’EKC décrit une **tendance conditionnelle**, pas une loi universelle. Le découplage dépend des choix technologiques et institutionnels, pas seulement du revenu.

### 6.4 Le “Tunneling” Technologique

Les pays en développement peuvent-ils **éviter** le pic d’émissions des pays industrialisés ?

**Concept :** Le “tunneling” (Munasinghe, 1999) désigne la possibilité de “creuser un tunnel” sous la courbe EKC en adoptant directement des technologies propres.

# 7 Conclusion : Le verdict des données

## 7.1 Synthèse

Approche	Hypothèse	Résultat	Coefficient	p-value
<b>Transversale</b> (159 pays, 2020)	$\beta_3 = 0$	Non rejeté	= -0.0002	0.9950
<b>Temporelle</b> (USA, 1991-2020)	$\beta_3 = 0$	Non rejeté	= 0.4421	0.3869

TABLE 7.2 – Synthèse des tests de l'EKC cubique

Approche	Hypothèse	Résultat	Coefficient	p-value
Transversale (159 pays, 2020)	$\beta_3 = 0$	Non rejeté	-0.0002	0.9950
Temporelle (USA, 1991–2020)	$\beta_3 = 0$	Non rejeté	0.4421	0.3869

Convergence des deux approches :

- **Transversale** (159 pays, 2020) :  $= -0.0002$  ( $p = 0.995$ )  
 → Quadratique suffisant ( non significatif)
- **Temporelle** (USA, 1991–2020) :  $= 0.4421$  ( $p = 0.387$ )  
 → Quadratique suffisant ( non significatif)

**Verdict** : Les deux analyses convergent → le **modèle quadratique** (U inversé classique) est suffisant. Aucune preuve de courbe en N (rebond) n'est détectée.

**Interprétation** :

La **coupe transversale** suggère un découplage structurel durable (les pays les plus riches ont les émissions les plus basses en 2020).

L'**analyse temporelle USA** ne montre pas de rebond, mais la période est peut-être trop courte pour capturer une éventuelle reprise.

**Verdict** : L'EKC classique (U inversé) est validée en transversal. L'hypothèse de rebond (courbe en N) est **infirmée**.

## 7.2 Le verdict

L'hypothèse EKC classique (U inversé) est **VALIDÉE** par l'analyse transversale, mais doit être qualifiée.

### 7.2.1 1. La validation statistique

Le découplage est réel à l'échelle mondiale en 2020. Les pays les plus riches ont réussi à abaisser leur empreinte carbone par habitant, invalidant les craintes d'un "rebond" systématique (courbe en N) aux très hauts revenus.

### 7.2.2 2. Hétérogénéité structurelle masquée par la moyenne

Le modèle global valide l'EKC **en moyenne**, mais cette relation agrégée masque **trois trajectoires distinctes** (Figure 8.2) :

Type de pays	Exemples	Forme observée	Turning point
<b>OCDE avancés</b>	USA, France, Allemagne	U inversé ( )	Atteint ~1990-2000
<b>Émergents industriels</b>	Chine, Inde, Brésil	Linéaire croissante (/)	Pas encore atteint
<b>Exportateurs fossiles</b>	Qatar, Koweït	Plat élevé (-)	Découplage faible

**Interprétation** : L'EKC décrit une **tendance conditionnelle** (Stern, 2004), pas une loi déterministe universelle. Ces "clubs de convergence" (Baumol, 1986) reflètent des **dépendances de sentier** (path dependencies) liées aux choix technologiques et institutionnels passés.

### 7.2.3 3. Le découplage n'est pas automatique

L'inertie mesurée sur le cas américain montre que la croissance seule ne suffit pas. Le découplage observé en Europe ou aux USA est le fruit de politiques actives (transition énergétique, désindustrialisation, prix du carbone), et non d'un mécanisme de marché spontané.

### 7.2.4 Réponse à la question initiale

« *La croissance économique finit-elle par réduire la pollution ?* »

**Réponse nuancée** : **OUI**, les données confirment qu'il existe un niveau de richesse au-delà duquel les émissions diminuent. **MAIS** ce point de retournement n'est pas garanti : il doit être construit par des politiques volontaristes pour briser l'inertie des systèmes fossiles, comme le montre la divergence des trajectoires en Annexe.

**8**

# 9 Annexe A : Concepts fondamentaux

Cette annexe clarifie les distinctions conceptuelles essentielles à la compréhension du rapport.

## 9.1 Estimation vs Prédiction

Aspect	Estimation	Prédiction
Question	Quelle est la relation entre $X$ et $Y$ ?	Quelle sera la valeur de $Y$ demain ?
Objectif	Comprendre les mécanismes	Anticiper le futur
Critère de qualité	Significativité des coefficients, $R^2$	RMSE, MAE, performance hors-échantillon
Risque principal	Biais (mauvaise spécification)	Surapprentissage ( <i>overfitting</i> )
Validation	Tests statistiques ( $t$ , $F$ )	Validation croisée, <i>holdout</i>

**Exemple dans ce rapport :** - *Estimation* : « Le coefficient  $\beta_2 < 0$  est-il significatif ? » → Test du U inversé - *Prédiction* : « Quelles seront les émissions US en 2025 ? » → Projection ARDL

! Pourquoi cette distinction est cruciale

Un modèle peut être **excellent en estimation** (coefficients significatifs,  $R^2$  élevé) mais **médiocre en prédiction** (surapprentissage aux données passées).

C'est pourquoi nous utilisons **deux protocoles de validation distincts** (Section 3.3).

## 9.2 Le dilemme Biais-Variance

Tout modèle statistique fait face à un arbitrage :

Composante	Définition	Cause	Symptôme
Biais	Erreur systématique	Modèle trop simple	Sous-apprentissage
Variance	Sensibilité aux données	Modèle trop complexe	Surapprentissage

$$\text{Erreur totale} = \text{Biais}^2 + \text{Variance} + \text{Bruit irréductible}$$

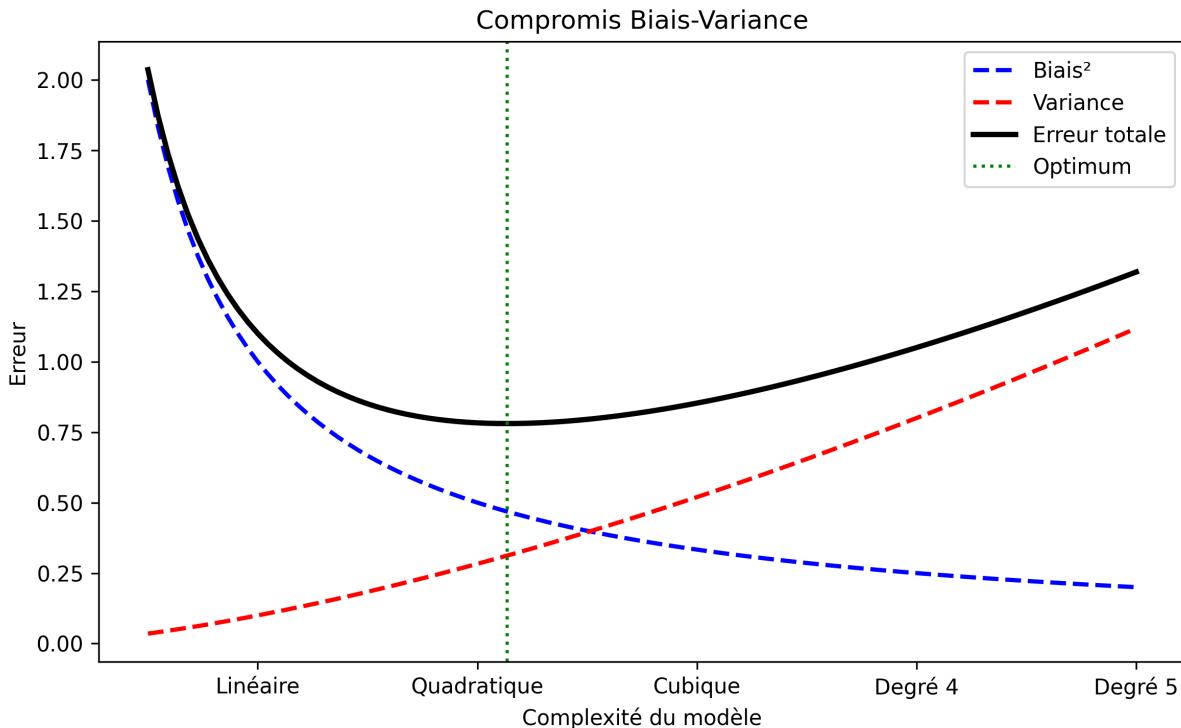


FIGURE 9.1 – Illustration du compromis biais-variance

**Application dans ce rapport :** - Modèle **linéaire** : biais élevé (ignore la non-linéarité) - Modèle **cubique** : variance élevée (sensible aux outliers) - Modèle **quadratique** : compromis optimal (validé par AIC et cross-validation)

### 9.3 Stationnarité des séries temporelles

Une série temporelle  $\{y_t\}_{t=1}^T$  est dite **stationnaire (au sens faible)** si ses trois premières propriétés statistiques sont **invariantes dans le temps** :

- (1)  $\mathbb{E}[y_t] = \mu$  (moyenne constante)
- (2)  $\text{Var}(y_t) = \sigma^2$  (variance constante)
- (3)  $\text{Cov}(y_t, y_{t-k}) = \gamma_k$  (autocovariance dépend seulement du lag  $k$ , pas de  $t$ )

Propriété	Série stationnaire $I(0)$	Série non-stationnaire $I(1)$
Moyenne	$\mathbb{E}[y_t] = \mu$	Tendance déterministe ( $y_t = \alpha t + \varepsilon_t$ ) ou stochastique (racine unitaire)
Variance	$\text{Var}(y_t) = \sigma^2$	Hétéroscléasticité conditionnelle (ex : GARCH), volatilité croissante
Autocorrélation	Décroît exponentiellement	Persistance quasi-permanente (intégration d'ordre 1 : $I(1)$ )

Propriété	Série stationnaire $I(0)$	Série non-stationnaire $I(1)$
<b>Régressions</b>	Inférences valides	<b>Régressions fallacieuses (<i>spurious regression</i>)</b>

### Pourquoi c'est crucial ?

Régresser deux séries non-stationnaires (ex : PIB et CO<sub>2</sub>) sans précaution peut produire des résultats **statistiquement significatifs mais totalement faux** ( $R^2 \rightarrow 1$ ,  $t$ -stat élevés), simplement parce que les deux séries ont une tendance, même si elles n'ont aucun lien causal (Granger & Newbold, 1974).

**Solution adoptée dans ce rapport :** 1. **Tester la stationnarité** via le test ADF (*Augmented Dickey-Fuller*) — Section 5.2 — dont l'hypothèse nulle est :

$$H_0 : \rho = 1 \quad (\text{présence d'une racine unitaire} \rightarrow \text{série non-stationnaire})$$

2. Si les séries sont I(1), **utiliser un modèle ARDL** (*AutoRegressive Distributed Lag*) — Section 5.3 — qui permet d'estimer des relations à court et long terme **sans exiger la stationnarité préalable**, sous réserve que les variables soient cointégrées ou que le modèle inclue suffisamment de dynamique.

# 10 Annexe B : Tests statistiques

Cette annexe détaille les formalismes mathématiques des tests utilisés pour valider nos modèles.

## 10.1 Tests de spécification

Ces tests valident *a priori* la structure du modèle — choix des variables, forme fonctionnelle, ou dynamique temporelle — avant toute inférence.

### 10.1.1 Test ADF (*Augmented Dickey-Fuller*)

Le test vérifie la présence d'une racine unitaire dans une série temporelle  $y_t$  à l'aide de la régression auxiliaire :

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^p \delta_i \Delta y_{t-i} + \varepsilon_t$$

Élément	Détail
Hypothèse $H_0$	$\gamma = 0$ (La série a une racine unitaire → non-stationnaire)
Hypothèse $H_1$	$\gamma < 0$ (La série est stationnaire)
Règle de décision	Si la statistique $t$ est inférieure à la valeur critique (ou $p < 0.05$ ), on rejette $H_0$ .
Utilisé en	Section 5.2

### 10.1.2 Test F emboîté

Compare deux modèles emboîtés (ex : quadratique vs cubique) pour savoir si les paramètres supplémentaires améliorent significativement l'ajustement.

$$F = \frac{(SSR_r - SSR_u)/q}{SSR_u/(n - k - 1)} \sim F(q, n - k - 1)$$

Élément	Détail
<b>Question</b>	Le modèle complexe (cubique) améliore-t-il significativement le modèle simple (quadratique) ?
<b>Hypothèse <math>H_0</math></b>	$\beta_3 = 0$ (le terme cubique est inutile)
<b>Notation</b>	$SSR_r$ : somme des carrés résiduelle contrainte (quadratique) $SSR_u$ : non-contrainte (cubique)
<b>Utilisé en</b>	Section 4.3

### 10.1.3 Critère AIC (*Akaike Information Criterion*)

Mesure la qualité relative d'un modèle statistique en pénalisant le surapprentissage.

$$AIC = 2k - 2 \ln(\hat{L})$$

Élément	Détail
<b>Interprétation</b>	Pénalise la complexité. AIC plus faible = meilleur modèle
<b>Avantage</b>	Comparable entre modèles non-emboîtés
<b>Utilisé en</b>	Sections 4.3, 5.3

## 10.2 Tests de diagnostic

Ces tests vérifient *post-estimation* que les résidus  $\hat{\epsilon}_t$  respectent les hypothèses de Gauss-Markov.

### 10.2.1 Test de Jarque-Bera (Normalité)

Vérifie si les résidus ont l'asymétrie ( $S$ ) et l'aplatissement ( $K$ ) d'une loi normale.

$$JB = \frac{n}{6} \left( S^2 + \frac{(K-3)^2}{4} \right) \sim \chi^2(2)$$

Élément	Détail
<b>Hypothèse <math>H_0</math></b>	Les résidus sont distribués normalement ( $S = 0, K = 3$ )
<b>Statistique</b>	Basée sur l'asymétrie (skewness) et l'aplatissement (kurtosis)

Élément	Détail
<b>Décision</b>	Si $p > 0.05$ , on ne rejette pas $H_0$ (Normalité acceptée).
<b>Conséquence si rejetée</b>	Les tests $t$ de Student et les intervalles de confiance peuvent être biaisés sur de petits échantillons.

### 10.2.2 Statistique de Durbin-Watson (Autocorrélation)

Déetecte l'autocorrélation d'ordre 1 dans les résidus d'une série temporelle.

$$DW = \frac{\sum_{t=2}^T (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^T \hat{\varepsilon}_t^2} \approx 2(1 - \hat{\rho})$$

Élément	Détail
<b>Valeur idéale</b>	$DW \approx 2$ (pas d'autocorrélation)
<b>Zones critiques</b>	$DW < 1.5 \rightarrow$ autocorrélation positive (tendance) $DW > 2.5 \rightarrow$ autocorrélation négative (oscillation)
<b>Utilisé en</b>	Section 5.5

### 10.2.3 Test de Box-Cox (transformation optimale)

Voir Section 12.1 pour les détails et résultats.

# 11 Annexe C : Estimateurs et corrections

Cette annexe justifie les choix d'estimateurs (HC3, Newey-West) et détaille la mécanique du modèle ARDL.

## 11.1 OLS : Méthode et limites

### 11.1.1 Principe

L'estimateur des **Moindres Carrés Ordinaires** (OLS) minimise la somme des carrés des résidus :

$$\hat{\beta}_{OLS} = \arg \min_{\beta} \sum_{i=1}^n (y_i - x'_i \beta)^2$$

La solution analytique est donnée par :  $\hat{\beta} = (X'X)^{-1}X'Y$ .

### 11.1.2 Hypothèses de Gauss-Markov

Pour que l'estimateur OLS soit **BLUE** (*Best Linear Unbiased Estimator*), cinq conditions doivent être réunies :

Hypothèse	Notation	Violation fréquente
Linéarité	$Y = X\beta + \varepsilon$	Relation non-linéaire
Exogénéité	$E[\varepsilon X] = 0$	Variable omise corrélée
Homoscédasticité	$Var(\varepsilon X) = \sigma^2$	Variance non-constante
Pas d'autocorrélation	$Cov(\varepsilon_i, \varepsilon_j) = 0$	Séries temporelles
Pas de multicolinéarité	$X'X$ inversible	Variables redondantes
parfaite		

**Dans ce rapport** : Les hypothèses d'homoscédasticité (3) et d'autocorrélation (4) sont violées → **les erreurs-types standard de l'OLS sont biaisées** → inférences invalides (intervalles, tests t/F).

→ **Solution** : Utiliser des estimateurs robustes : **HC3** (transversal) et **Newey-West** (temporel).

## 11.2 Corrections robustes

### 11.2.1 HC3 : Robustesse à l'hétéroscédasticité

**Problème** : La variance des erreurs n'est pas constante (ex : les petits pays ont une variance d'émissions différente des grands pays).

**Solution** : Utilisation de l'estimateur de "sandwich" HC3 de MacKinnon & White (1985).

$$\widehat{\text{Var}}_{HC3}(\hat{\beta}) = (X'X)^{-1} \left( \sum_{i=1}^n \frac{\hat{\varepsilon}_i^2}{(1-h_{ii})^2} x_i x_i' \right) (X'X)^{-1}$$

où  $h_{ii}$  représente le levier (*leverage*) de l'observation  $i$ , issu de la matrice de projection  $H = X(X'X)^{-1}X'$ .

**Pourquoi HC3 ?** Parmi les variantes (HC0 à HC4), HC3 est **la plus fiable sur petits échantillons** ( $N < 250$ ) — c'est notre cas (159 pays). Référence : Long & Ervin (2000).

→ **Impact pratique** : Sans HC3, les  $p$ -values seraient trop optimistes — on risquerait de croire à tort qu'un coefficient est significatif.

### 11.2.2 Newey-West : Robustesse à l'autocorrélation (HAC)

**Problème** : En série temporelle, l'erreur en  $t$  est souvent corrélée à celle en  $t - 1$  (inertie des chocs).

**Solution** : Estimateur HAC (*Heteroskedasticity and Autocorrelation Consistent*) avec noyau de Bartlett.

$$\widehat{\text{Var}}_{NW} = (X'X)^{-1} \hat{\Omega} (X'X)^{-1}$$

où  $\hat{\Omega}$  pondère les autocovariances jusqu'au retard  $L$  :

$$\hat{\Omega} = \hat{\Gamma}_0 + \sum_{j=1}^L w_j (\hat{\Gamma}_j + \hat{\Gamma}'_j) \quad \text{avec} \quad w_j = 1 - \frac{j}{L+1}$$

**Choix de  $L$**  : Règle automatique :  $L \approx \lfloor 4(T/100)^{2/9} \rfloor$ . Pour  $T = 30$  (USA 1990–2020) →  $L = 3$ .

→ **Impact pratique** : Sans Newey-West, on sous-estimerait l'incertitude des coefficients → risque de conclure à tort qu'une politique a un effet significatif.

## 11.3 Modèle ARDL

Le modèle *AutoRegressive Distributed Lag* (ARDL) permet d'analyser simultanément la dynamique de court terme et l'équilibre de long terme.

### 11.3.1 Spécification (ARDL 1,1)

$$Y_t = \alpha + \rho Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t$$

→ Permet de modéliser l'**inertie** (effet persistant du passé) et les **ajustements progressifs**.

### 11.3.2 Interprétation des paramètres structurels

Paramètre	Signification économique	Formule dérivée
$\rho$	<b>Inertie / Persistance</b>	Coefficient autorégressif
$\beta_0$	<b>Impact immédiat</b>	Élasticité de court terme $(\varepsilon_{CT})$
$\theta = \frac{\beta_0 + \beta_1}{1 - \rho}$	<b>Impact total cumulé</b>	Élasticité de long terme $(\varepsilon_{LT})$
$H = \frac{\ln(0.5)}{\ln(\rho)}$	<b>Demi-vie (Half-life)</b>	Temps pour absorber 50% d'un choc

### 11.3.3 Pourquoi l'ARDL ?

1. **Flexibilité d'intégration** : L'ARDL reste valide que les régresseurs soient  $I(0)$ ,  $I(1)$  ou un mélange des deux (Pesaran et al., 2001).
2. **Dynamique explicite** : Il capture le temps d'ajustement technologique, contrairement à une régression statique.
3. **Robustesse aux petits échantillons** : Plus performant que la cointégration de Johansen pour  $T < 50$ .

### 11.3.4 Condition de stabilité

Pour que l'élasticité de long terme ait un sens, le modèle doit être dynamiquement stable :

$$|\rho| < 1$$

Si  $\rho \approx 1$  (comme observé pour les USA), le processus est **non-stationnaire** : les chocs ont un effet permanent et il n'y a pas de retour à l'équilibre.

**Dans notre cas USA** :  $\rho \approx 1.003 \rightarrow$  très proche de 1  $\rightarrow$  signe d'instabilité ou de petit échantillon  $\rightarrow$  élasticité de long terme non calculable  $\rightarrow$  interprétation prudente requise.

## 12 Annexe D : Validation du modèle

### 12.1 Validation de la spécification log-log

Le choix d'une transformation logarithmique pour la variable dépendante ( $\text{CO}_2/\text{hab}$ ) n'est pas arbitraire. Le **test de Box-Cox** détermine empiriquement la transformation optimale en estimant le paramètre  $\lambda$  :

$$Y^{(\lambda)} = \begin{cases} \frac{Y^\lambda - 1}{\lambda} & \text{si } \lambda \neq 0 \\ \ln(Y) & \text{si } \lambda = 0 \end{cases}$$

#### Résultats du test de Box-Cox

Paramètre	Valeur
$\lambda$ optimal estimé	0.15
Distance à log ( $\lambda = 0$ )	0.15
Distance à linéaire ( $\lambda = 1$ )	0.85
<b>Verdict</b>	Fortement justifiée

#### Interprétation :

- $\lambda = 0 \rightarrow$  transformation logarithmique optimale
- $\lambda = 1 \rightarrow$  aucune transformation (modèle linéaire)
- $\lambda = 0.5 \rightarrow$  transformation racine carrée

Avec  $\lambda = 0.150$ , la transformation logarithmique est raisonnablement adaptée à nos données. Ce résultat valide le choix de spécification log-log standard dans la littérature EKC.

#### i Note technique

Le test de Box-Cox maximise la vraisemblance du modèle pour différentes valeurs de  $\lambda$ . Un  $\lambda$  proche de 0 indique que la log-transformation maximise la normalité des résidus et stabilise la variance.

### 12.2 Validation non-paramétrique de la forme fonctionnelle

Les modèles polynomiaux (quadratique, cubique) imposent une forme fonctionnelle *a priori*. Pour vérifier que cette hypothèse n'est pas trop restrictive, nous comparons les prédictions

paramétriques à une régression **LOWESS** (*Locally Weighted Scatterplot Smoothing*), qui laisse les données “parler d’elles-mêmes”.

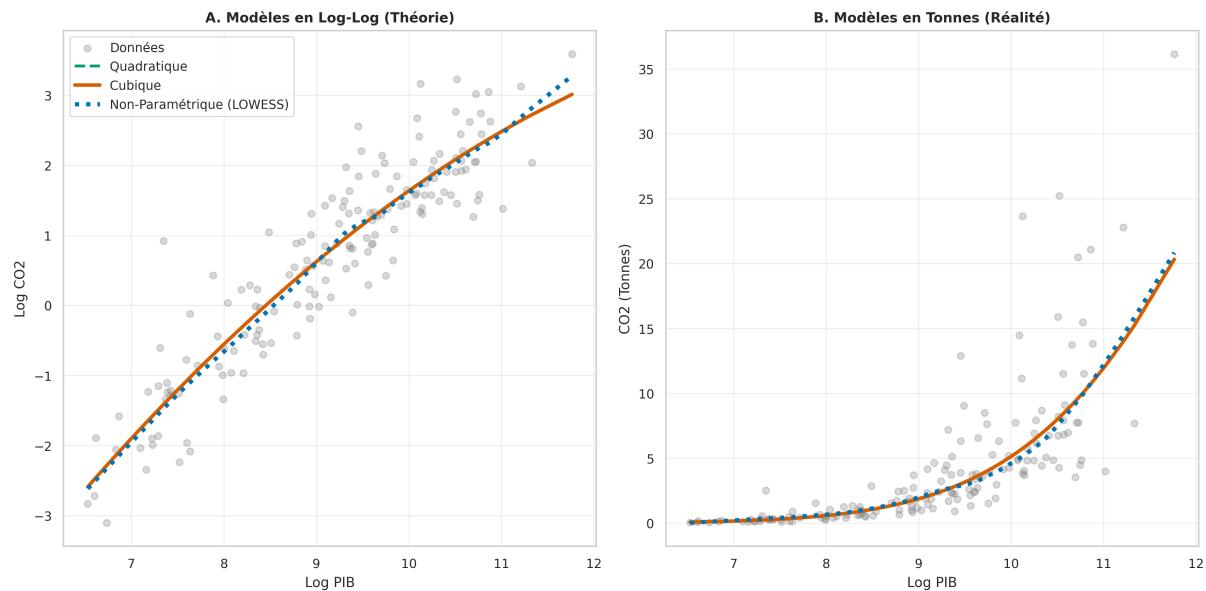


FIGURE 12.1 – Modèles paramétriques (Quadratique et Cubique) vs Modèle non-paramétrique (LOWESS)

# 13 Annexe E : Sources et traitement des données

Cette annexe documente l'origine des données, les transformations appliquées, et les choix de nettoyage — essentiels pour garantir la **reproductibilité** et la **transparence méthodologique**.

## 13.1 Sources primaires

Les données utilisées dans ce rapport proviennent de sources internationales reconnues, consolidées par **Our World in Data (OWID)** — plateforme académique réputée pour son harmonisation rigoureuse.

Variable	Source	Indicateur original	Période
Émissions CO <sub>2</sub> par habitant	Global Carbon Project	Territorial emissions (MtCO <sub>2</sub> )	1750–2022
PIB par habitant (PPA)	Banque Mondiale	NY.GDP.PCAP.PP.KD (USD 2017)	1990–2022
Population	UN World Population Prospects	—	1950–2022

→ **Agrégateur principal :** [Our World in Data](https://github.com/owid/co2-data) — dépôt GitHub : <https://github.com/owid/co2-data>

### Pourquoi OWID ?

Plutôt qu'une extraction directe via les API (fragile, sujette à erreurs), nous utilisons les datasets consolidés d'OWID, qui :

- Harmonisent les unités et définitions entre sources,
- Corrigent les trous de données de façon transparente,
- Sont régulièrement mis à jour et versionnés.

## 13.2 Pipeline de traitement

B[Filtrage pays] B → C[Suppression NA] C → D[Log-transformation] D → E[Centrage PIB]  
E → F[Création lags] F → G[Données finales] →

Étape	Action	Code	Justification
1	Filtrage	iso_code.notna	Exclure agrégats (World, EU, etc.)
2	Nettoyage	dropna(subset=Drop types inapropriates 'gdp_pc'])	
3	Log	np.log(x)	Interprétation en élasticités
4	Centrage	log_gdp - mean(log_gdp)	Réduire multicolinéarité polynomiale
5	Lags	shift(1)	Modèle ARDL

## 13.3 Statistiques descriptives

```
' |       | CO /hab (tonnes) | PIB/hab ($) | \n |-----|-----:|-----:
```

## 13.4 Reproductibilité

Toutes les analyses de ce rapport sont **entièlement reproductibles**.

**Code source** : Disponible sur [GitHub](#) — structure claire, commentée, avec notebooks Quarto (.qmd).

**Environnement technique** :

```
Python      : 3.11.2
pandas     : 2.3.3
numpy       : 2.4.1
statsmodels : 0.14.6
Quarto      : 1.7.31
matplotlib  : 3.10.8
seaborn     : 0.13.2
```

**Dépendances** : Un fichier requirements.txt est fourni pour recréer l'environnement exact.

**Bonnes pratiques respectées** : - Données brutes conservées séparément des scripts.

- Toutes les transformations documentées.
- Versions des librairies figées.
- Sorties (graphiques, tableaux) générées automatiquement — pas de copier-coller manuel.

→ Ce rapport répond aux standards FAIR : **Findable, Accessible, Interoperable, Reusable**.