

MÉMOIRE

Master 1 Économétrie et Statistique, parcours Économétrie Appliquée

Sensibilité des prix agricoles face aux chocs climatiques

Une analyse de séries temporelles au Brésil de janvier 2000 à décembre 2022

PALACIO BARCO Gloria Isabel

Directeur du master : **DARNE Olivier**

Professeur encadrant : **BENNANI Hamza**

3 juin 2024

Remerciements

Je tiens à exprimer mes remerciements à mon encadrant, M. Hamza Bennani, pour le temps qu'il m'a consacré et ses conseils. Je suis également reconnaissante envers les directeurs du master ECAP, M. Olivier Darné et Mme. Muriel Travers, qui m'ont offert l'opportunité d'intégrer ce master et pouvoir reprendre mes études. Un merci spécial également aux enseignants, dont la qualité et la passion ont été remarquables.

Sur un plan plus personnel, je voudrais remercier mes collègues du master, qui m'ont aidée à m'intégrer et à progresser tout au long de cette année. Enfin, je suis infiniment reconnaissante envers ma famille pour ses encouragements et pour m'avoir toujours rappelé que les choses peuvent toujours s'améliorer. Un merci tout particulier à Claire, qui a non seulement joué un rôle crucial dans la réalisation de ce projet, mais qui a également été un soutien très important pour moi, m'accompagnant à chaque étape de ce parcours.

Résumé

Cette étude analyse les prix des produits agricoles au Brésil, un acteur clé des enjeux agricoles mondiaux, en utilisant l'Indice des Prix à la Production Agricole (IPA) de janvier 2000 à décembre 2022. L'étude a intégré des variables climatiques pour évaluer leur impact sur l'IPA, en utilisant des modèles de régression tels que les Moindres Carrés Ordinaires (MCO) et ARMAX. Dans notre approche par les MCO, nous avons rencontré des difficultés telles que l'autocorrélation des résidus et l'hétéroscédasticité. Malgré plusieurs tentatives, le modèle MCO n'a pas pu être validé et nous avons adopté le modèle ARMAX qui s'est avéré plus efficace. Les résultats montrent que seul le taux de change, le prix du pétrole et l'influence des valeurs passées de l'IPA (effet retardé) ont eu un impact statistiquement significatif. Nos recherches révèlent que, bien que les variables climatiques n'aient pas eu un effet notable, la complexité du système agricole suggère qu'une approche globale et uniforme n'est pas adéquate. Ainsi, nous préconisons, tant dans les études que dans l'élaboration des stratégies, l'adoption de mesures localisées qui répondent mieux aux besoins spécifiques des communautés agricoles.

Abstract

This study analyzes the prices of agricultural products in Brazil, a key player in global agricultural issues, using the Agricultural Production Price Index (IPA) from January 2000 to December 2022. The study incorporated climatic variables to assess their impact on the IPA, using regression models such as Ordinary Least Squares (OLS) and ARMAX. In our approach with OLS, we encountered difficulties such as autocorrelation of residuals and heteroscedasticity. Despite several attempts, the OLS model could not be validated, and we adopted the ARMAX model, which proved to be more effective. The results show that only the exchange rate, oil prices, and the influence of past values of the IPA (lagged effect) had a statistically significant impact. Our research reveals that although climatic variables did not have a noticeable effect, the complexity of the agricultural system suggests that a uniform, global approach is inadequate. Thus, we advocate for the adoption of localized measures in both research and strategy development, which better meet the specific needs of agricultural communities.

Mots clés : Agriculture, Brésil, MCO, ARMAX séries temporelles, R, changement climatique

Sommaire

Glossaire sigles	5
1 Introduction	1
2 Environnement économique et politique	4
3 Méthodologie économétrique	7
4 Présentation des données	14
5 Analyse de Données	30
6 Estimation économétriques et interprétation des résultats	39
7 Conclusion.....	61
8 Discussion	63
9 Bibliographie	66
10 Annexes	70
Source des données	87
Tableau de figures	88
Table des matières	89

Glossaire des sigles

ACF : Autocorrelation Function

ADF : Augmented Dickey-Fuller

AR : AutoRegressive

ARMAX : AutoRegressive Moving Average with eXogenous inputs

AIC : Akaike Information Criterion

BR : Réal brésilien

BRIC : Brésil, Russie, Inde, Chine

CELAC : Communauté des États Latino-Américains et Caraïbes

FAO STATS : Food and Agriculture Organization Statistics

GLS : Generalized Least Squares

IPEA : Institut de Recherche Économique Appliquée

IPA-DI : Índice de Preços ao Produtor Amplo - Disponibilidade Interna

KPSS : Kwiatkowski-Phillips-Schmidt-Shin

MA : Moving Average

MCO : Moindres Carrés Ordinaires

MERCOSUR : Marché commun du Sud

REDD+ : Reducing Emissions from Deforestation and Forest Degradation

SMIC : Salaire minimum nominal actuel

UNASUR : Union des Nations Sud-Américaines

1 Introduction

L'agriculture, dans sa forme actuelle, est au cœur de défis critiques qui affectent non seulement notre bien-être actuel mais aussi l'avenir des sociétés à l'échelle mondiale. S'intéresser à elle devrait donc être une priorité pour les générations futures. Toutefois, il est impératif de repenser notre approche de l'agriculture, non seulement pour répondre aux besoins immédiats en matière de sécurité alimentaire mais aussi pour aborder les enjeux environnementaux et sociaux à long terme. Comprendre l'agriculture dépasse la simple acquisition de connaissances techniques ; c'est aussi comprendre ses implications sociales, car une population qui ne peut assurer sa propre sécurité alimentaire court un risque majeur.

Cette étude s'intéresse à l'agriculture principalement sous un angle économique, en se focalisant sur les prix à la production, qui constituent le premier maillon de la chaîne alimentaire. Ces prix sont déterminants car ils influencent directement le coût des aliments pour les consommateurs et, par extension, le phénomène de l'insécurité alimentaire.

Plus précisément nous nous intéressons aux prix à la production agricole au Brésil, un pays qui se distingue comme un acteur majeur, mais qui est également au cœur de nombreuses controverses.

L'agriculture brésilienne a déjà suscité l'intérêt et fait l'objet de multiples études, par exemple, le travail de (F. Antunes de Araujo, 2020) qui analyse la prévisibilité des principaux produits agricoles, mettant en lumière les dynamiques complexes des marchés agricoles brésiliens. Mais aussi à l'échelle internationale, le rôle du Brésil dans la volatilité des prix agricoles a également été confirmé à travers le travail sur le marché alimentaire mondial, comme celui de (F. Ceballos, 2016).

Plus récemment, en 2024, l'article "Empirical Modeling of Agricultural Climate Risk" (J. Burney, 2024) a exploré les répercussions du changement climatique sur l'agriculture brésilienne. Cette recherche a analysé en détail les variations climatiques pour évaluer leur impact sur les rendements agricoles, les revenus et la performance des prêts agricoles. Comme une grande majorité des recherches se basent sur le rendement

agricole, les résultats indiquent une augmentation de la variabilité des rendements et des revenus agricoles à la mi-siècle, soulignant ainsi l'importance d'analyser conjointement les facteurs climatiques et économiques dès maintenant.

Le Brésil, avec ses 330 millions d'hectares de terres agricoles, détient l'un des plus grands potentiels mondiaux en matière de terres arables et se classe comme la troisième puissance agricole mondiale, après les États-Unis et la Chine. Le pays présente une particularité notable dans sa structure agricole, car elle est divisée en deux systèmes distincts. D'une part, il existe une agriculture familiale, semblable à celle que l'on retrouve dans certaines fermes européennes du sud ou en Afrique, qui, bien que représentant 76,8 % des exploitations, gère seulement 23 % des terres agricoles. D'autre part, l'agrobusiness, qui évoque davantage les grandes exploitations nord-américaines, bien que constituant seulement 9,3 % des exploitations, contrôle 79,5 % des terres (Girardi, 2022).

Cette dualité génère des stratégies de gestion agricole très différentes au sein d'un même pays, qui est par ailleurs marqué par une grande variété climatique avec cinq régions climatiques distinctes. À cette diversité climatique s'ajoutent des conditions topographiques, hydrographiques et une biodiversité unique, qui enrichissent mais complexifient également le panorama agricole brésilien.

L'objectif de cette étude est de faire une analyse globale sur la sensibilité des prix agricoles et comment ils peuvent être influencés par des chocs climatiques. Pour ce faire, nous adoptons deux méthodes de régression : la méthode des Moindres Carrés Ordinaires et la méthode ARMAX. La MCO, bien que traditionnellement moins utilisée pour les séries temporelles en raison de ses limitations dans le traitement de l'autocorrélation et de la non-stationnarité, a été démontrée comme étant utile dans certains contextes par (Aragon, 2011) dans son ouvrage « Séries Temporelles : Méthodes et cas ». L'approche ARMAX est particulièrement adaptée pour notre analyse, car elle représente une variante des méthodes ARMA déjà utilisées par (R. Endré Dahl, 2020) dans ces recherches sur les canaux de transmission du pétrole vers les matières premières agricoles ou dans des études telles que celles (S. Raya, 2023), qui ont employé un modèle ARIMA-LSTM pour prédire la volatilité des séries de prix agricoles. Cette méthode permet d'intégrer les variables exogènes tout en gérant les autocorrélations, offrant ainsi une analyse robuste.

Au début de notre étude, nous présenterons une revue théorique des variables sélectionnées pour le jeu de données, justifiant leur pertinence en relation avec les objectifs de l'étude. Nous explorerons aussi les fondements des méthodes de régression que nous avons choisies.

La partie pratique commencera par une analyse exploratoire des données, comprenant des statistiques descriptives et l'identification des valeurs atypiques. Nous mettrons en œuvre des techniques économétriques telles que la correction de la saisonnalité et l'application de tests de racine unitaire, notamment le test de Dickey-Fuller Augmenté et le test KPSS, pour assurer la stationnarité des séries temporelles.

Les résultats des différents modèles testés seront ensuite présentés, en sélectionnant celui qui respecte toutes les hypothèses requises et fournit des résultats interprétables.

Dans la dernière partie, la conclusion, nous discuterons nos résultats. Nos hypothèses initiales supposaient un impact significatif des variables climatiques telles que les précipitations, le nombre de jours secs consécutifs et la température sur les prix agricoles. Toutefois, les résultats ont montré que seules certaines variables économiques, telles que le pétrole et le taux de change, étaient statistiquement significatives, alors que les variables climatiques n'ont pas montré d'effets notables.

Nous analyserons des hypothèses expliquant cette absence d'impact apparent des variables climatiques, en particulier le rôle des avancées en gestion agricole et en génétique des plantes. Ces innovations, incluant des variétés résistantes à la sécheresse et les systèmes de double culture, ont vraisemblablement atténué les effets négatifs du changement climatique et renforcé la résilience des cultures.

2 Environnement économique et politique

En 2022, plus de 9 % de la population mondiale étaient confrontés à la faim chronique. Les projections pour 2030 ne sont pas plus optimistes, estimant que plus de 600 millions de personnes continueront de souffrir de la faim à travers le monde (Organisation des Nations Unies pour l'alimentation et l'agriculture, 2023). Cette situation met en évidence l'urgence de la crise alimentaire mondiale qui est au cœur des défis depuis plusieurs années.

Selon (Daviron, 2012) et la FAO, la sécurité alimentaire est principalement compromise par les niveaux élevés des prix des denrées alimentaires. Au-delà de l'instabilité immédiate, il existe une préoccupation plus profonde liée à la persistance des prix élevés. Ce phénomène rend particulièrement difficile l'accès à une alimentation adéquate et nutritive, surtout pour les plus vulnérables.

Dans ce contexte, nous nous intéressons au cas du Brésil, qui occupe une position de premier plan dans le secteur agricole mondial. En effet, le Brésil est l'un des plus grands producteurs et exportateurs de divers produits agricoles, tels que le soja, le café, et le bœuf, plaçant le pays au cœur des échanges internationaux.

Le Brésil, bénéficie d'une abondance de ressources naturelles. Cependant, malgré cette richesse, il fait face à de grandes disparités et contradictions. D'une part, il est capable de générer des quantités impressionnantes de produits agricoles, tandis que d'autre part, une partie de sa population lutte contre la pauvreté et la faim. Ces contradictions s'étendent au-delà des aspects matériels et touchent également les choix politiques et économiques.

En termes de politiques agricoles, le Brésil a connu des périodes de grandes avancées et de régressions notables. Depuis les années 2000, le pays était reconnu pour ses politiques progressistes favorisant la sécurité alimentaire, le soutien aux agricultures familiales, et la gestion environnementale. Cependant, à la suite de changements politiques survenus après 2016, une nette régression a été constatée. Il est donc essentiel de comprendre les politiques publiques rurales et environnementales au Brésil, comme décrites par (E.

Sabourin, 2020). Pendant les mandats de Lula da Silva entre 2003 et 2010, la création du programme national de développement durable des territoires ruraux (PRONAT) a marqué un tournant en 2003, établissant une base solide pour le développement de politiques territoriales coordonnées.

Le lancement du programme Territoires de la Citoyenneté (PTC) sous la coordination de la Maison Civile a été une autre étape majeure. Ce programme visait à consolider les actions de différents ministères, y compris le Ministère du Développement Agraire (MDA), pour renforcer les territoires économiquement fragiles.

Sous la présidence de Dilma Rousseff, il y eut un renforcement de certaines politiques agricoles et environnementales initiées précédemment. Toutefois, la période a été marquée par une austérité budgétaire qui a commencé à limiter l'efficacité de ces initiatives, réduisant la capacité du gouvernement à soutenir pleinement le développement agricole et la conservation environnementale.

En mai 2016, la destitution de la présidente Dilma Rousseff a marqué le début d'une période de régression politique avec le gouvernement de Michel Temer.

Sous Michel Temer, il y eut une intensification des stratégies de démantèlement des structures existantes, notamment par la suppression du ministère du Développement Agraire (MDA), un acteur clé de l'agriculture familiale. En décembre 2016, l'introduction de la Mesure Provisoire 759, transformée en loi en juillet 2017, a paralysé la politique de réforme agraire et facilité la privatisation des terres publiques, modifiant profondément le paysage agraire et environnemental du Brésil.

Avec l'arrivée de Jair Bolsonaro, les politiques de démantèlement se sont accélérées et élargies. En juin 2019, la suppression du Conseil National de Développement Rural Durable (CONDRAF) et en 2019, la législation communément appelée « Loi du Poison » a modifié la réglementation des pesticides, autorisant l'utilisation de plus de 500 nouveaux produits, ce qui a suscité de vives critiques au niveau national et international. Cette période a également vu une continuation de l'affaiblissement financier et politique du Programme d'acquisition d'aliments (PAA), renforçant les contradictions entre les politiques de sécurité alimentaire et les pratiques de l'agro-industrie.

Après avoir été emprisonné sous des accusations controversées, qui ont ensuite été annulées, Lula da Silva a fait un retour notable sur la scène politique brésilienne. Sa

réélection en 2022 suggère un possible retour aux politiques qui ont simultanément stimulé le développement agricole et social au Brésil. Cependant, les défis actuels sont conséquents, non seulement par les répercussions de la crise du COVID-19, qui a sévèrement touché le pays, mais également par une grande instabilité climatique.

La position du Brésil comme acteur majeur de la production agricole mondiale coexiste avec son rôle de gardien de l'Amazonie, la plus grande forêt tropicale du monde, cruciale pour la régulation du climat mondial. Le développement économique du pays s'est souvent fait au détriment de l'environnement. Sous l'administration Bolsonaro, par exemple, la déforestation en Amazonie a considérablement augmenté, entraînant la perte d'environ 40 000 km² de forêt, selon Anaïs Fléchet, historienne et spécialiste du Brésil à l'Université Paris-Saclay. Cette déforestation accélérée a de graves implications non seulement pour la biodiversité mais aussi pour le climat, aggravant les effets du changement climatique.

En tant que membre des BRICS — un groupe composé du Brésil, de la Russie, de l'Inde, de la Chine, et de l'Afrique du Sud — le Brésil participe à des discussions cruciales sur des sujets tels que la lutte contre le changement climatique et le développement commercial. Ces nations, toutes à revenu intermédiaire, ont connu une croissance économique rapide et pourraient émerger comme les principales économies mondiales d'ici 2050 (O'Neill, 2008).

Le Brésil est également membre fondateur du MERCOSUR, une union économique significative en Amérique du Sud, et participe activement à l'Union des Nations Sud-Américaines (UNASUR) et à la CELAC (Communauté des États Latino-Américains et Caraïbes), qui visent à renforcer l'intégration et la coopération régionales. De plus, le traité de libre-échange négocié entre l'Union européenne et le MERCOSUR souligne l'importance du Brésil sur la scène du commerce international.

Ces affiliations montrent le rôle stratégique du Brésil dans les discussions économiques et environnementales mondiales, accentuant l'importance de ses politiques intérieures et de leur impact international.

3 Méthodologie économétrique

Dans cette partie, nous aborderons différents concepts économétriques, qui seront essentiels pour approfondir notre réflexion.

Commençons par définir ce qu'est une série temporelle : Il s'agit d'une collection de données enregistrées à des moments précis et à intervalles réguliers. Ces données sont couramment utilisées pour analyser l'évolution de certaines quantités au fil du temps et pour faire des prédictions futures. Elle se compose de quatre parties principales : la tendance, le cycle, la saisonnalité, et la partie irrégulière ou résiduelle.

$$Y_t = T_t + C_t + S_t + I_t$$

où:

- Y_t représente la valeur de la série temporelle au temps t
- T_t est la composante de tendance
- C_t est la composante cyclique
- S_t est la composante saisonnière
- I_t est la composante irrégulière

La tendance illustre l'évolution générale des données sur une période prolongée, signalant une croissance, une décroissance ou une constance. Le cycle, de son côté, indique des fluctuations qui surviennent à des intervalles irréguliers. La saisonnalité désigne les variations qui se répètent de manière régulière et prévisible, comme les changements liés aux saisons. Enfin, la composante irrégulière ou résiduelle comprend les éléments qui ne sont pas expliqués par la tendance, le cycle, ou la saisonnalité, représentant souvent des anomalies ou des événements aléatoires qui influencent les données.

3.1 Stationnarité

La stationnarité est un concept crucial pour notre analyse, car elle conditionne l'efficacité des modèles que nous utiliserons, notamment les méthodes des moindres carrés ordinaires et ARMAX.

Une série temporelle est considérée comme stationnaire si ses propriétés statistiques, telles que la moyenne, la variance et la structure de corrélation, restent invariables au fil du temps.

- $E(Y_t) = m$ indépendant du temps
- $V(Y_t) = \gamma(0) < \infty$, $\gamma(0)$ indépendant du temps
- $Cov(Y_t, Y_{t-h}) = \gamma(h)$ ne dépend pas de t

Cette stationnarité implique qu'il n'y a ni tendance globale ni saisonnalité marquée affectant la série.

Nous pouvons dire que les séries temporelles peuvent devenir non stationnaires pour deux raisons principales : les variances inégales et les tendances, c'est-à-dire un changement systématique dans la moyenne de la série au fil du temps.

La tendance peut prendre deux formes :

Une tendance déterministe, qui est prévisible et peut être modélisée comme une fonction fixe du temps, comme une ligne droite ou une courbe polynomiale. Et une tendance stochastique, aussi connue sous le nom de marche aléatoire, qui évolue de manière imprévisible au fil du temps.

Pour aborder la non-stationnarité, diverses méthodes peuvent être appliquées. Si elle résulte de variances inégales, des transformations comme la transformation logarithmique sont souvent utilisées.

Quant aux tendances, la stratégie de traitement varie selon leur nature : la différenciation est employée pour neutraliser les tendances stochastiques et stabiliser la série, tandis que la régression est préférée pour les tendances déterministes en raison de leur caractère prévisible (Righetti, 2022).

Pour évaluer la stationnarité, plusieurs tests sont disponibles, parmi lesquels nous nous intéresserons principalement à deux :

Le Test de Dickey-Fuller Augmenté (ADF) vérifie l'existence d'une racine unitaire qui indiquerait une tendance stochastique dans les données. L'hypothèse nulle de ce test est la présence d'une racine unitaire, suggérant que la série est non stationnaire.

Le test KPSS (Kwiatkowski-Phillips-Schmidt-Shin), contrairement à l'ADF examine la stationnarité autour d'une tendance, assumant sous l'hypothèse nulle que la série est stationnaire avec une tendance déterministe.

Ces tests offrent des perspectives complémentaires et sont souvent utilisés ensemble pour un diagnostic robuste de la stationnarité.

Pour traiter les cas de non-stationnarité, la méthode de décomposition peut être aussi utilisée. Elle permet d'isoler et de retirer les composantes de tendance, de saisonnalité et les irrégularités des données. Une fois ces éléments identifiés et écartés, il devient possible d'analyser leur impact individuel sur la série.

3.2 Saisonnalité

La décomposition est également un outil pour la correction de la saisonnalité. Cependant, avant de l'appliquer efficacement, il est important de comprendre le type de saisonnalité présente dans la série. La saisonnalité peut se manifester de deux manières principales : additive ou multiplicative.

Dans le modèle additif, la saisonnalité est considérée comme une somme constante qui s'ajoute à la tendance de la série. Ce modèle est approprié lorsque les variations saisonnières sont relativement constantes au fil du temps. L'équation pour le modèle additif est :

$$Y_t = T_t + S_t + e_t$$

Où Y_t est la valeur observée, T est la composante de tendance, S_t est la composante saisonnière, et e_t est la composante irrégulière.

Le modèle multiplicatif est utilisé lorsque les variations saisonnières évoluent proportionnellement à l'évolution de la tendance. Cela est souvent le cas lorsque la série montre une croissance exponentielle. L'équation pour le modèle multiplicatif est :

$$Y_t = T_t \times S_t \times e_t$$

Où les composantes sont multipliées plutôt qu'additionnées.

Après la décomposition, il est également essentiel de vérifier si la série résiduelle est stationnaire.

Une fois que nous avons établi et corrigé la stationnarité et saisonnalité de nos séries si nécessaire, nous pouvons alors appliquer des méthodes de régression. Cela nous permet de distinguer les impacts réels des variables explicatives sur l'IPA de ceux qui pourraient être faussement attribués à des artefacts de non-stationnarité dans les données.

Deux approches de régression seront utilisées pour notre analyse : la méthode des moindres carrés ordinaires, qui est adaptée pour modéliser les relations linéaires dans des données stationnaires, et la méthode ARMAX, qui est conçue pour prendre en compte à la fois les influences exogènes et la dynamique interne des séries temporelles.

3.3 Modèles de régression

3.3.1 Moindres carrés ordinaires (MCO)

La méthode des moindres carrés ordinaires est une technique couramment utilisée pour estimer les relations entre une variable dépendante et une ou plusieurs variables indépendantes . Elle cherche à trouver les valeurs des coefficients qui minimisent la somme des carrés des différences (les "résidus") entre les valeurs observées de la variable dépendante et les valeurs prédites par le modèle linéaire.

La formulation générale de la régression linéaire par la méthode des moindres carrés peut être représentée comme:

$$Y = X\beta + U$$

Où :

- Y est le vecteur de la variable dépendante
- X est la matrice qui contient les données des variables indépendantes, chaque colonne étant une variable différente et chaque ligne une observation différente
- β est le vecteur des coefficients que l'on cherche à estimer, qui quantifie l'impact de chaque variable indépendante sur la variable dépendante
- U est le vecteur des erreurs, représentant la différence entre les valeurs observées et les valeurs prédites par le modèle

Pour que les estimations fournies par la méthode des moindres carrés soient fiables et valides, plusieurs hypothèses doivent être respectées (M.TRAVERS, 2023) :

- Homoscédasticité: La variance des erreurs (différence entre les valeurs réelles et les valeurs prédictives) est constante pour toutes les observations
- Indépendance des erreurs: Les erreurs sont indépendantes les unes des autres, ce qui signifie qu'il n'y a pas de corrélation entre elles
- Normalité des erreurs : Les erreurs suivent une distribution normale
- Erreurs indépendantes des variables explicatives (hypothèse d'exogénéité des variables explicatives)
- Nombre d'observations > Nombre de variables explicatives
- Absence de multicolinéarité entre les variables explicatives
- En moyenne, le modèle est bien spécifié

3.3.2 Régression linéaire et séries temporelles

La méthode des moindres carrés ordinaires appliquée aux séries temporelles cherche à estimer les relations entre une variable dépendante Y_t et plusieurs variables indépendantes $X_{2t}, X_{3t}, \dots, X_{kt}$, où l'indice t représente le temps (par exemple, jours, mois, années). Le modèle linéaire prend la forme :

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \dots + \beta_k X_{kt} + U_t$$

Hypothèses Spécifiques aux Séries Temporelles :

- Non-autocorrélation : Les erreurs U_t à différents moments doivent être non corrélées
- Stationnarité : Pour certains modèles, il est nécessaire que les séries de données soient stationnaires

Les coefficients sont estimés par :

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

Cette estimation vise à minimiser la somme des carrés des résidus entre les valeurs observées et les valeurs prédictives par le modèle. En présence d'autocorrélation ou de non-stationnarité, les estimations standard des MCO peuvent être biaisées ou inefficaces.

3.3.3 Modèle AR

Le modèle autorégressif (AR) est utilisé pour analyser et prévoir les séries temporelles en se basant sur les valeurs passées de la série elle-même. Cela signifie que les valeurs futures d'une série sont prédites à partir de ses valeurs antérieures

Le modèle autorégressif d'ordre 1 peut être exprimé par l'équation suivante :

$$y_t = c + \phi y_{t-1} + z_t$$

Où :

- Y_t est la valeur de la série au temps t
- c est la constante
- ϕ est le coefficient autorégressif qui mesure l'impact de la valeur précédente Y_{t-1} sur la valeur actuelle Y_t
- Z_t est le terme d'erreur aléatoire au temps t , assumé d'être indépendant et normalement distribué avec une moyenne de zéro et une variance σ_z^2

En développant la formule, on obtient une expression où chaque Y_t peut être exprimé comme une somme pondérée de tous les termes d'erreur passés, ajustés par les puissances successives de ϕ :

$$y_t = c + z_t + \phi z_{t-1} + \phi^2 z_{t-2} + \dots$$

Cette expression montre que Y_t est influencé par une "mémoire" des chocs passés (les erreurs), pondérée décroissante par ϕ au fur et à mesure que l'on remonte dans le temps.

3.3.4 Modèle Moyenne Mobile MA

Un modèle moyenne mobile d'ordre q , noté MA(q), exprime chaque valeur de la série temporelle comme la somme d'un terme moyen et d'une combinaison linéaire de termes

$$y_t = \mu + z_t + \theta_1 z_{t-1} + \theta_2 z_{t-2} + \dots + \theta_q z_{t-q}$$

Où :

- μ est la moyenne de la série

- Z_t sont des termes d'erreur blancs (c'est-à-dire des chocs aléatoires indépendants avec une distribution normale)
- $\theta_1, \theta_2, \dots, \theta_q$ sont les coefficients qui mesurent l'impact des erreurs passées sur la valeur actuelle

3.3.5 Modèle ARMA(p, q)

Un modèle ARMA combine les caractéristiques des modèles autorégressifs (AR) et moyenne mobile (MA). Il est défini comme suit :

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + z_t + \theta_1 z_{t-1} + \dots + \theta_q z_{t-q}$$

Où :

- c est la constante
- ϕ_1, \dots, ϕ_p sont les coefficients autorégressifs
- Z_t est le terme d'erreur
- $\theta_1, \dots, \theta_q$ sont les coefficients moyenne mobile

Cette formule montre que chaque valeur de la série dépend à la fois de ses valeurs antérieures (partie AR) et des chocs d'erreur antérieurs (partie MA).

3.3.6 Modèle ARMAX

Le modèle ARMAX est une extension des modèles de séries temporelles qui intègre à la fois les dynamiques internes de la série (comme dans les modèles ARMA) et l'impact des variables explicatives exogènes (similaire à un modèle de régression linéaire).

Un modèle ARMAX combine trois éléments principaux: La partie autorégressive, qui capte la dépendance de la variable actuelle Y_t sur ses propres valeurs passées, la partie moyenne mobile, qui modélise l'impact des erreurs passées sur la valeur actuelle, intégrant les chocs passés dans les prévisions actuelles et les variables exogènes qui sont supposées influencer directement la variable dépendante Y_t .

Le modèle peut être écrit comme :

$$y_t = \beta_0 + \beta_1 x_{1,t} + \dots + \beta_k x_{k,t} + u_t$$

Où U_t est un terme d'erreur qui suit un modèle ARMA :

$$u_t = \frac{1 - \phi_1 B - \dots - \phi_p B^p}{1 + \theta_1 B + \dots + \theta_q B^q} z_t$$

Et Z_t est un bruit blanc avec une variance σ_z^2 .

Dans notre étude sur l'influence des variables climatiques sur l'évolution de l'indice des prix à la production agricole, nous adopterons le modèle ARMAX, en plus de MCO car il est idéal pour intégrer à la fois les effets des variables exogènes et la dynamique interne des séries temporelles. Il permettra d'examiner comment les conditions climatiques passées et actuelles affectent les prix.

Nous commencerons par nous assurer que les séries sont désaisonnalisées et rendues stationnaires pour éviter les régressions fallacieuses. Dans ce processus, nous progresserons par itérations, testant plusieurs approches. Bien que la méthode des moindres carrés ordinaires puisse être adéquate pour analyser des séries temporelles, elle exige une attention particulière pour éviter les résidus autocorrélés qui pourraient invalider les estimations du modèle.

4 Présentation des données

Dans cette partie, nous présenterons les données utilisées pour notre analyse. Ces données sont recueillies avec une fréquence mensuelle, s'étendant de janvier 2000 à décembre 2022.

4.1 Variable d'intérêt : Indice des prix à la production agricole

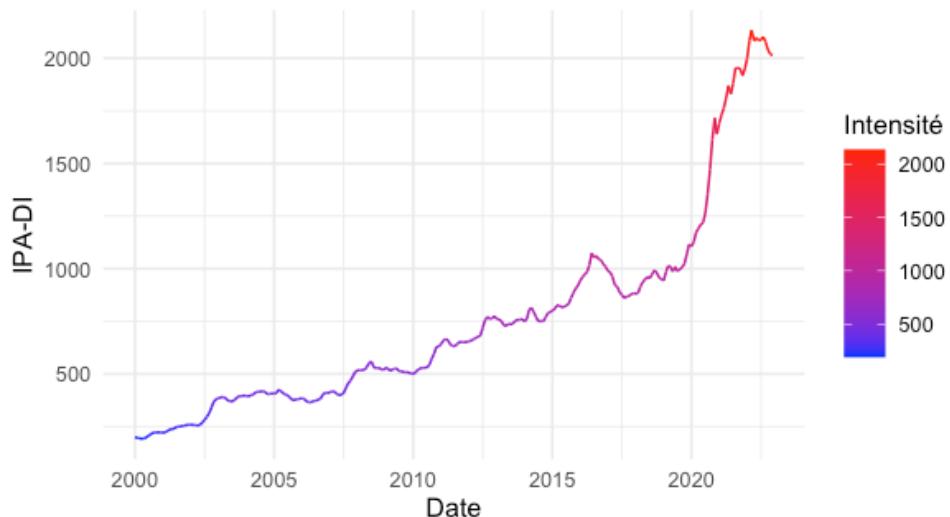
Cet indice est conçu pour mesurer les variations moyennes des prix auxquels les producteurs nationaux vendent leurs produits. Il est significatif car il offre un aperçu des pressions qui peuvent augmenter ou diminuer les prix, du point de vue du producteur

plutôt que du consommateur final. Cela est particulièrement pertinent puisque l'indice est calculé avant l'application des coûts de transport, de transformation et de commercialisation.

Dans ce projet, nous nous intéressons spécifiquement à l'IPA-DI des produits agricoles (englobant les cultures temporaires et permanentes ainsi que l'élevage) au Brésil, qui couvre l'intégralité du mois de référence, c'est-à-dire du premier au dernier jour du mois où les données sont collectées, fournissant ainsi une mesure précise des variations de prix durant cette période spécifique.

L'IPA est essentiellement utilisé pour offrir une perspective détaillée des fluctuations des prix dès le premier stade de la chaîne de production. Il permet de suivre les variations de prix depuis les producteurs agricoles, qui constituent la base de la chaîne, jusqu'aux industries agroalimentaires, et enfin jusqu'aux prix de vente au détail. Cette approche est importante pour évaluer comment les facteurs externes, tels que les coûts des matières premières (semences, engrains, produits énergétiques, etc.), influencent les coûts de production initiale. Étant donné que les effets des variations des prix des matières premières sur les prix à la consommation sont généralement retardés (J. Baleyte, 2023), l'IPA permet d'identifier les principaux facteurs qui influencent ces prix et d'analyser l'ampleur ainsi que la rapidité avec laquelle les augmentations des coûts se répercutent sur les prix en sortie de ferme.

Figure 1 IPA-DI - origine - prod. Agricole - indice (août 1994 = 100)
Jan 2000 à Déc 2022



Dans la Figure 1, nous observons l'évolution de l'indice durant la période étudiée, ce qui permet d'effectuer des constatations initiales. Selon les sites officiels du gouvernement, comme celui de l'Institut de Recherche Économique Appliquée (IPEA Data), le mois et l'année de référence pour l'indice, qui est suivi depuis 1944, est fixée à août 1994 avec un indice de base de 100 (Fundação Getulio Vargas, Conjuntura Econômica, 2024). Toutes les mesures futures de l'indice sont comparées à ce point de référence. Cette méthode permet d'évaluer l'inflation ou la déflation des prix des produits à la production depuis cette date. Il est intéressant de noter que depuis 1994, l'indice n'a jamais été inférieur à 100.

Dans notre graphique l'indice commence en l'année 2000 avec un chiffre de 198,55, révélant une tendance globalement croissante au fil des années, particulièrement notable lors de la période finale. Au cours de la première décennie, l'indice se maintient en dessous de 500, y compris durant la crise économique mondiale de 2008, période durant laquelle on note une légère baisse suivie d'une stabilisation en 2009. Les mandats de Lula da Silva ont souvent été associés à un renforcement des politiques de soutien à l'agriculture familiale et au développement durable, influençant positivement les prix agricoles.

La reprise économique observée entre 2010 et 2012, ainsi que l'augmentation de la demande pour les produits agricoles, ont déclenché une nouvelle hausse des prix agricoles. Cette période fut également marquée par des contraintes budgétaires au Brésil.

En 2016, sous la présidence de Dilma Rousseff, le Brésil a traversé une année particulièrement difficile. Auparavant épargnée par la crise de 2008 grâce à une conjoncture économique favorable, la situation s'est renversée lorsque la demande chinoise, principal client du Brésil, a commencé à ralentir. La chute des prix internationaux des matières premières a directement impacté les revenus en devises du pays, même si la dévaluation du real brésilien a partiellement compensé cet effet.

L'arrivée au pouvoir de Jair Bolsonaro en 2019 a encore aggravé la situation. Bien que certaines de ses politiques aient stimulé l'expansion agricole, notamment par l'assouplissement des réglementations environnementales et le renforcement du soutien à l'agro-industrie, elles ont également suscité des préoccupations internationales et provoqué des perturbations internes. Ces perturbations ont souvent entraîné une hausse des coûts et, par conséquent, des prix.

La dévaluation du real par rapport au dollar américain a rendu les produits agricoles brésiliens plus compétitifs et moins onéreux pour les acheteurs internationaux, dynamisant ainsi les exportations et entraînant une augmentation des prix sur le marché intérieur. En outre, la pandémie de COVID-19 a provoqué d'importantes perturbations dans les chaînes d'approvisionnement mondiales, impactant significativement la production, le transport et la distribution des produits agricoles.

Notre graphique s'arrête en décembre 2022, une année marquée par la fin apparente de la crise du COVID-19, la réélection du président Lula da Silva (qui a pris ses fonctions le 1er janvier 2023), une montée de l'inflation, les répercussions de la guerre en Ukraine, ainsi qu'un durcissement des conditions financières et des conditions météorologiques particulièrement défavorables. Comme nous pouvons le constater, l'indice des prix à la production des produits agricoles est influencé par divers facteurs qui interagissent de manière complexe.

Passons maintenant à la section suivante qui explore les variables explicatives susceptibles d'impacter cet indice, afin de mieux comprendre les mécanismes qui influencent ces fluctuations de prix.

4.2 Variables climatiques

Pour notre modèle, nous avons inclus trois variables climatiques. Avant de les examiner, il est intéressant de comprendre la complexité du climat au Brésil, qui réside en partie dans sa grande étendue, mais aussi dans la diversité des climats qu'il présente. Contrairement à l'Europe, les saisons au Brésil sont inversées. Ainsi, de décembre à février, c'est l'été, la saison la plus chaude et souvent la plus humide, particulièrement dans le sud où des températures élevées s'accompagnent souvent de fortes pluies. De mars à mai, les températures commencent à baisser, marquant une période de transition avec moins de pluies, surtout en avril et mai.

L'hiver, de juin à août, est la saison la plus froide, en particulier dans le sud du Brésil où les températures peuvent parfois avoisiner 0°C. Dans les régions nordiques et centrales, les températures restent relativement chaudes, mais c'est également la période la plus sèche de l'année.

Le printemps, de septembre à novembre, voit les températures remonter et les précipitations augmenter, préparant le terrain pour l'été chaud et humide à suivre.

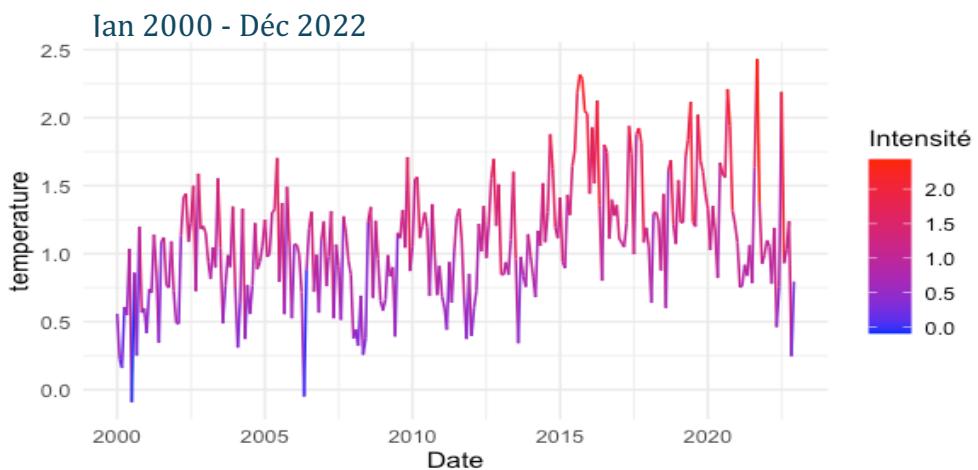
Dans le bassin amazonien, un climat équatorial prévaut, se distinguant par deux saisons distinctes :

- La saison des pluies, de décembre à mai, caractérisée par un pic de précipitations.
- La saison sèche, de juin à novembre, marquée par une diminution relative des pluies.

Les variations climatiques au Brésil sont également influencées par plusieurs facteurs, notamment la géographie, les courants océaniques et les phénomènes météorologiques tels qu' el Niño et la Niña, qui peuvent modifier considérablement les modèles de précipitation et de température à travers le pays.

4.2.1 Variation de température

Figure 2 Variation de températures en °C



Selon le rapport de la Banque mondiale concernant la sécurité alimentaire et le changement climatique de 2022 bien que l'élévation initiale des températures puisse parfois favoriser la croissance des cultures, les conséquences à long terme, comme l'intensification de l'évapotranspiration et la diminution des ressources en eau, se révèlent nuisibles. Les régions sujettes à des températures élevées sont particulièrement vulnérables, avec des menaces directes sur des cultures sensibles à la chaleur telles que

le blé. Lorsque le réchauffement climatique dépasse un seuil de 2 degrés Celsius, les défis d'adaptation se multiplient, rendant la gestion des impacts à la fois coûteuse et complexe (La banque mondiale , 2022).

Une des conséquences directes de la hausse des températures est la réduction des rendements agricoles, causée par un stress hydrique subi par les plantes ainsi qu'un mauvais remplissage des grains, menant à une perte significative de rendement. Cette baisse involontaire de la production influence directement les prix agricoles, en perturbant l'équilibre entre l'offre et la demande. En outre, la productivité agricole est également compromise par les fortes chaleurs qui réduisent la capacité de travail des ouvriers agricoles, affectant leur santé et leur efficacité.

La Figure 2 nous permet d'observer la série de données disponible sur le site de FAO Stats, qui couvre la variation de température mensuelle au Brésil de janvier 2000 à décembre 2022, illustre les anomalies de température. Ces anomalies sont calculées en comparant la température moyenne observée chaque mois à une valeur de référence, ici établie sur la période de 1951 à 1980. Ce procédé permet de mesurer les écarts par rapport à cette norme historique, offrant un aperçu clair de l'évolution des températures au fil du temps. Cette méthode est particulièrement efficace pour observer les tendances climatiques et identifier des périodes de réchauffement ou de refroidissement significatif par rapport à la moyenne de référence.

Dans le graphique nous pouvons constater une tendance générale à la hausse qui s'intensifie après l'année 2015. Selon une étude réalisée par (L.F.A. Curado, 2023), en effet, une augmentation significative a été observée dans les régions du nord et du nord-est en 2015, atteignant un pic en 2016. Pour les régions du centre-ouest, du sud-est et du sud, les températures ont également montré une augmentation, notamment pour les valeurs maximales et minimales, mais ce phénomène ne s'est manifesté qu'à partir de l'année 2019. Il est également à noter une diminution des températures minimales, révélant des variations plus extrêmes.

Une cause directe de ces variations extrêmes a été l'augmentation du nombre d'incendies, comme en Amazonie en 2019, mais aussi dans tout le pays. Par exemple, entre janvier et août, le Brésil a connu 73 000 incendies (France, 2020).

4.2.2 Précipitations

La deuxième variable que nous examinons en lien avec le changement climatique concerne les précipitations cumulées. Si les fortes chaleurs affectent directement la production agricole et donc les prix, le stress hydrique, qui peut aussi résulter d'un excès de pluie, joue un rôle tout aussi important. Ce phénomène, appelé "stress hydrique par excès", dépend non seulement de la quantité mais aussi de la répartition des précipitations.

Dans les climats tropicaux, où les précipitations sont vitales, leurs variations peuvent affecter l'agriculture de multiples façons : elles peuvent devenir plus intenses ou plus rares, avec des impacts qui varient considérablement d'une région à l'autre. Par conséquent, le problème n'est pas uniforme et, dans les situations extrêmes, peut provoquer des dégâts significatifs. En réponse à ces défis, les agriculteurs peuvent être contraints d'investir davantage dans des systèmes d'irrigation pour compenser la variabilité des précipitations. Cet investissement accru dans l'irrigation entraîne une augmentation des coûts de production qui peuvent être répercutés sur les prix.

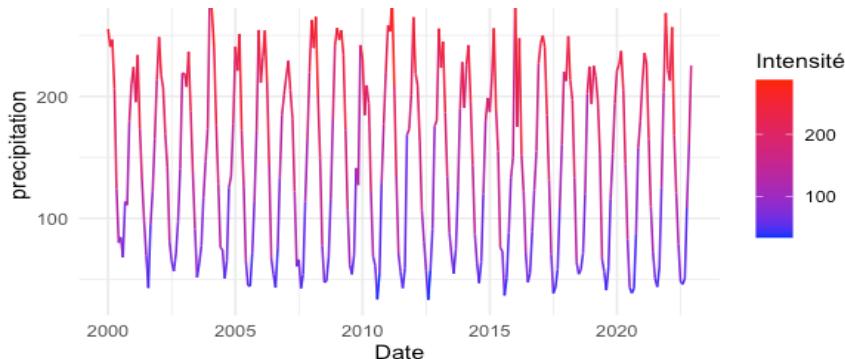
Au-delà de l'impact sur les rendements, les précipitations jouent également un rôle dans la santé publique et la gestion des ressources en eau, affectant ainsi indirectement l'économie agricole. Par exemple, la prolifération de vecteurs de maladies due à l'augmentation des eaux stagnantes peut réduire la disponibilité de la main-d'œuvre agricole, tandis que les défis liés à la gestion de l'eau, augmentent par des précipitations irrégulières, peuvent limiter les capacités d'irrigation, essentielles à la production.

Les fluctuations des précipitations peuvent également inciter à des migrations de la main-d'œuvre agricole vers les zones urbaines, réduisant ainsi les ressources humaines disponibles pour l'agriculture et augmentant potentiellement les coûts de production. De plus, une augmentation de la fréquence des catastrophes naturelles, comme les inondations, peut causer des dommages importants aux infrastructures agricoles, perturber la chaîne d'approvisionnement et, par conséquent, exercer une pression à la hausse sur les prix. Selon la Confédération Nationale des Municipalités, de 2013 à 2022, ces catastrophes ont engendré des pertes financières considérables, estimées à plus de

401,3 milliards de BRL environ 73,3 milliards d'EUR (Service économique régional de Brasilia, 2023), affectant tant le secteur public que privé.

Figure 3 Précipitations en mm

Jan 2000 -Déc 2022



La Figure 3 illustre les données des précipitations enregistrées au Brésil de janvier 2000 à décembre 2022, avec l'intensité des précipitations mesurée en millimètres. Les données proviennent du portail de connaissances sur le changement climatique de la Banque mondiale. On y distingue clairement une variation saisonnière, marquée par des pics récurrents d'intensité élevée, illustrés en rouge, qui apparaissent à intervalles réguliers au fil des ans. Cette observation suggère un modèle de précipitations influencé par des facteurs saisonniers.

4.2.3 Le nombre de jours secs consécutifs (CDD)

Le nombre de jours secs consécutifs (CDD, pour "Consecutive Dry Days") est un indicateur climatologique important qui mesure la durée la plus longue d'une période sans précipitations significatives. Cette durée est calculée en identifiant la plus longue période de sécheresse dans une série chronologique et en notant la valeur maximale observée chaque mois de la période d'étude. Un jour est considéré comme sec si les précipitations quotidiennes accumulées sont inférieures à 1 mm.

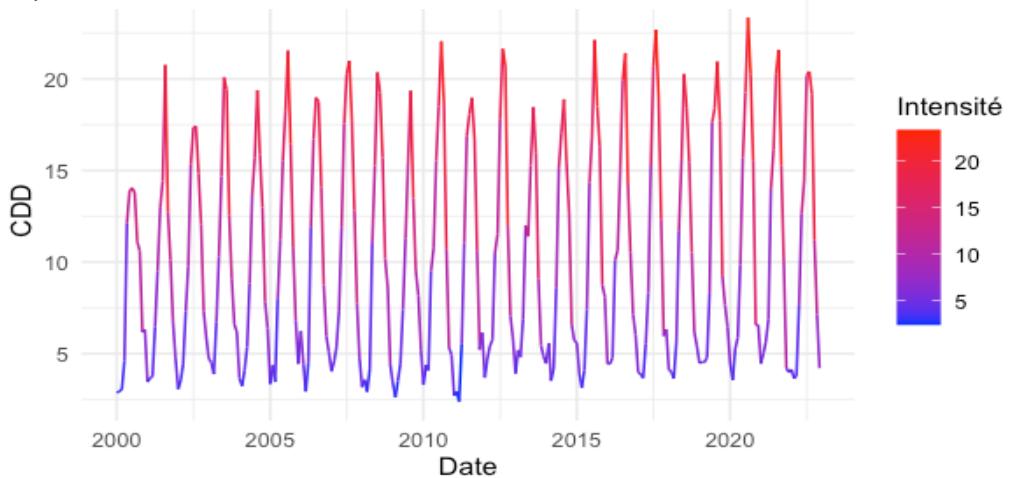
Cette variable est cruciale pour comprendre les risques de sécheresse. En effet, des périodes prolongées sans pluie peuvent impacter significativement les réserves d'eau, nuire aux cultures, et augmenter le risque d'incendies de forêt.

Selon l'article « Analysis of climate extremes indices over northeast Brazil from 1961 to 2014» (R. Lisboa Costa, 2020), une tendance à l'augmentation des vagues de chaleur a été observée, entraînant une hausse du nombre de jours secs consécutifs et intensifiant ainsi la saisonnalité. Cette évolution se manifeste par des saisons sèches de plus en plus arides et des saisons humides plus intenses, altérant profondément les cycles naturels et les conditions environnementales.

La Figure 4 illustre graphiquement les données de la variable CDD, nous pouvons voir cette tendance, montrant une légère augmentation observée au cours des deux dernières décennies.

Figure 4 Nombre de jours secs consécutifs

Jan 2000 – Déc 2022



4.3 Variables économiques

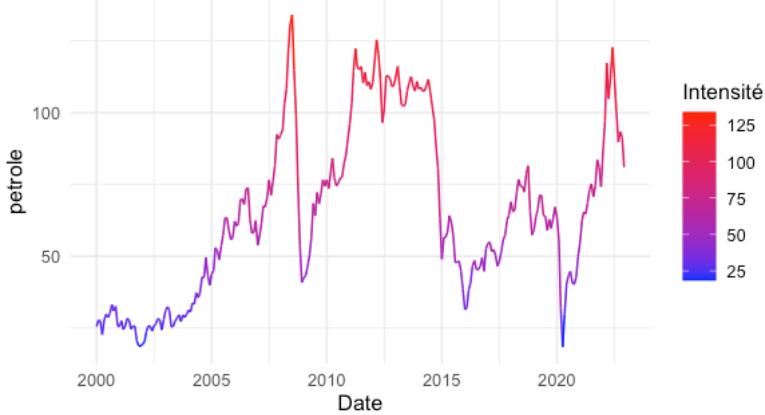
4.3.1 Prix du pétrole

La variable « prix du pétrole » correspond au « Preço por barril do petróleo bruto Brent (FOB) », et les données sont disponibles avec une fréquence quotidienne sur IPEA DATA. Afin de synchroniser cette variable avec les autres indicateurs économiques analysés, nous avons calculé la moyenne mensuelle de ces valeurs quotidiennes.

Ce prix, exprimé en dollars US par baril, est évalué selon la méthode FOB (free on board), qui exclut les coûts de transport et d'assurance. Le Brent, produit en mer du Nord, est une référence essentielle pour le prix international des divers types de pétrole brut.

Figure 5 Prix du pétrole en dollar US

Ian 2000 – Déc 2022



Les recherches économiques établissent souvent un lien entre les variations des prix du pétrole et celles des matières premières agricoles, notant une tendance à la hausse des prix agricoles en réponse à l'augmentation des prix du pétrole. De nombreuses études ont exploré cette dynamique, révélant généralement une causalité à sens unique (A. Kapsuzoglu, 2015): les fluctuations du pétrole influencent directement les marchés agricoles, sans que l'inverse ne soit notable.

Cette relation souligne le rôle important de l'inclusion de l'évolution des prix du pétrole comme variable explicative lorsqu'on s'intéresse aux prix des productions agricoles. Selon la littérature, il existe trois principaux canaux de transmission (G. El Montasserb, 2023). Le premier canal, direct du côté de l'offre, montre comment le pétrole et ses dérivés, en tant que composants essentiels des coûts de production agricole (carburants pour les machines agricoles, chauffage des serres, production d'engrais, etc.), peuvent augmenter les coûts de production agricole lorsqu'ils deviennent plus coûteux, entraînant ainsi une hausse des prix des produits agricoles. Par exemple, au Brésil, où l'agriculture est fortement mécanisée, une hausse des prix du diesel peut directement augmenter les coûts de production de cultures majeures comme le soja et le café.

Le deuxième canal, indirect du côté de l'offre, illustre comment la conversion d'une culture agricole en biocarburant peut affecter les marchés d'autres produits agricoles par des effets de substitution (A. Saucedo, 2015). Au Brésil, l'expansion massive de la culture de la canne à sucre destinée à la production d'éthanol a réduit la disponibilité de terres pour d'autres cultures, telles que les haricots noirs, aliment de base pour la population locale et entraîner une augmentation des prix des haricots .

Le troisième canal, du côté de la demande, intervient lorsque les coûts élevés du pétrole affectent les prix de produits non agricoles, tels que les fibres synthétiques.

La Figure 5 illustre les fluctuations du prix du Brent depuis janvier 2000. Par exemple, nous observons le pic historique des prix du pétrole en 2008, causé par une forte demande mondiale, une spéculation intense sur les marchés et des tensions géopolitiques dans les régions productrices de pétrole. Ce pic a été suivi d'une chute spectaculaire des prix lorsque la crise financière mondiale a éclaté plus tard dans l'année, entraînant une récession économique mondiale.

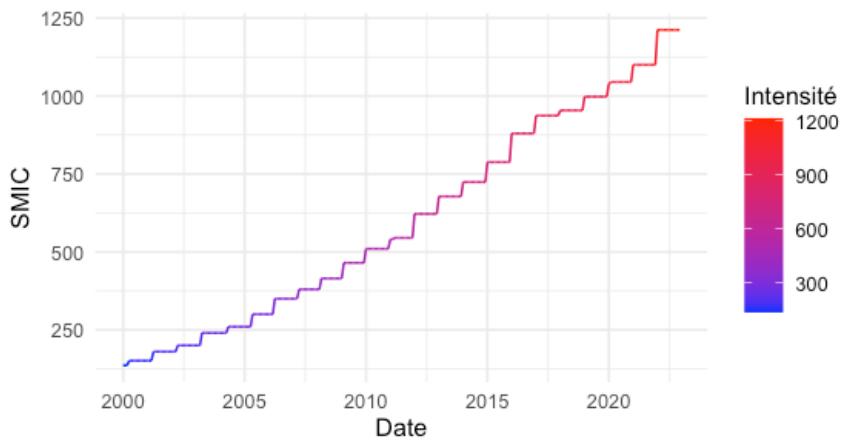
Entre 2010 et 2014, on observe une certaine stabilité des prix du pétrole, une période caractérisée par une reprise économique à la suite de la crise financière de 2008, qui a stimulé la demande mondiale de pétrole. Toutefois, fin 2014 marque le début d'une chute significative des prix, tombant à environ 30 USD en 2015. Cette baisse est principalement attribuable à une surproduction mondiale de pétrole, couplée à un ralentissement économique global. Malgré les tensions géopolitiques persistantes au Moyen-Orient, l'offre excédentaire et les niveaux élevés des stocks de pétrole brut ont maintenu les prix à la baisse.

Cette reprise est entravée par la pandémie de COVID-19 en 2020, provoquant une chute historique des prix en avril 2020. La hausse marquée en 2022 peut être directement liée à l'invasion de l'Ukraine par la Russie, événement qui a suscité d'importantes inquiétudes quant à l'approvisionnement en pétrole russe, un acteur majeur sur le marché mondial.

4.3.2 SMIC

La variable « SMIC », pour « Salário Mínimo Nominal Vigente », désigne le montant minimal de base que les employeurs sont légalement tenus de payer à leurs travailleurs au Brésil. Ce montant, fixé par la loi depuis 1940, constitue le seuil légal de rémunération dans l'ensemble du pays. Le salaire minimum nominal sert de référence pour les salaires dans divers secteurs, influençant ainsi les structures de coûts économiques et les politiques de rémunération. Il est important de noter que le salaire minimum réel diffère du nominal en ce qu'il est ajusté pour tenir compte de l'inflation.

Figure 6 Salaire minimum nominal actuel en R\$
Jan 2000 -Déc 2022



Le Figure 6 démontre une tendance claire à la hausse du SMIC sur la période observée. Le salaire minimum, peut influencer les prix des produits agricoles de plusieurs façons. La plus évidente est à travers le coût de production dans le secteur agricole, où la main-d'œuvre peut représenter une part significative des coûts, surtout dans les cultures intensives en main-d'œuvre comme les fruits et les légumes. Mais aussi une augmentation du SMIC augmente le pouvoir d'achat des ménages, en particulier pour ceux qui gagnent le moins. Cela peut entraîner une hausse de la demande de produits alimentaires, particulièrement des biens de première nécessité et des produits frais.

Cependant, les effets peuvent également se manifester de manière moins directe. D'après la littérature, avec l'augmentation du revenu disponible à la suite de la hausse du SMIC, les consommateurs peuvent opter pour des produits alimentaires de meilleure qualité ou pour des produits plus coûteux, tels que la viande (Regmi A., 2021). Cette modification des habitudes de consommation peut entraîner une réallocation des ressources agricoles, ce qui réduit potentiellement l'offre de certains produits de base et peut, par conséquent, entraîner une augmentation de leurs prix.

Cette dynamique soulève également un problème d'équité où les consommateurs de différents niveaux de revenu se concurrencent pour des ressources limitées. Ceux qui sont moins sensibles aux prix, généralement les plus aisés, disposent d'un avantage concurrentiel. En période de pénurie, les consommateurs les plus pauvres, qui sont beaucoup plus affectés par les hausses de prix, doivent réduire leur consommation de

manière disproportionnée afin de maintenir l'équilibre du marché. Cela crée une dynamique où les conséquences des fluctuations de prix sont plus lourdement subies par les populations les plus vulnérables.

4.3.3 Exportations

La variable «Exportações - agricultura e pecuária - quantum: índice » représente un indice mesurant les quantités physiques des exportations du secteur de l'agriculture et de l'élevage.

L'indice mesure les variations des quantités physiques exportées par le Brésil dans ces secteurs, par rapport à une année de base, qui est ici 2018. Cet indice est établi sur 100 pour l'année de base, ce qui signifie que les valeurs de l'indice reflètent une comparaison par rapport aux niveaux d'exportation de 2018.

L'indice est calculé par un processus nommé "deflacionamento", qui ajuste les valeurs monétaires des exportations en fonction des variations de prix pour séparer l'impact des changements de prix des modifications des volumes physiques exportés. Il révèle la quantité réelle des exportations, indépendamment des fluctuations des prix, ce qui est intéressant pour analyser les exportations agricoles.

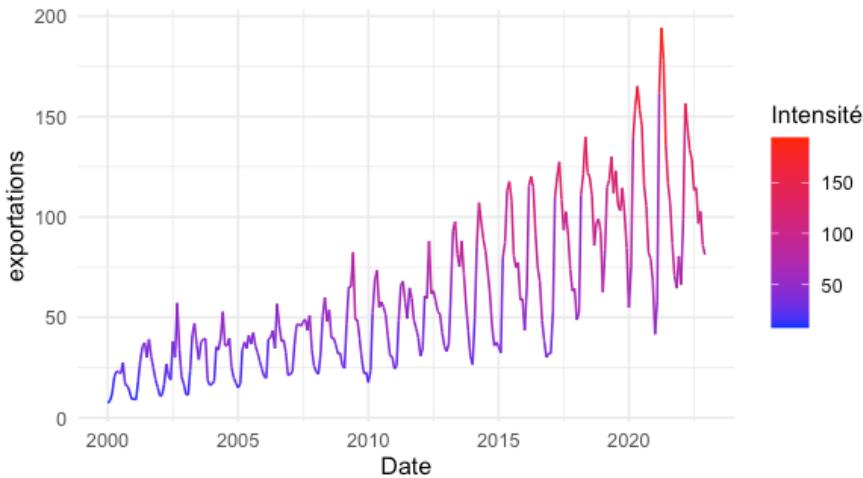
En mettant l'accent sur les volumes plutôt que sur les valeurs monétaires, cet indice permet d'observer la productivité et l'efficacité du secteur agricole, notamment sa capacité à augmenter le volume des produits vendus internationalement.

De plus, les prix de production agricoles sont influencés par ces exportations, car une augmentation des volumes exportés peut indiquer une hausse de la demande internationale pour des produits comme le soja, le café et la viande, dont le Brésil est un exportateur majeur. Cette demande peut pousser les prix de production à la hausse, surtout si l'offre ne parvient pas à s'adapter rapidement à cette nouvelle demande, accentuant ainsi les effets sur l'économie agricole globale.

La Figure 7 montre l'évolution de l'indice des exportations, mettant en évidence une tendance à la hausse clairement marquée depuis 2010.

Figure 7 Exportations des produits agricoles

Jan 2000 – Déc 2022



Cette évolution nous permet de revenir sur la dualité entre l'« agrobusiness » et l'agriculture familiale. Cette situation est le résultat des changements politiques qui ont commencé dans les années 1990, rendant l'agriculture industrielle plus dominante. Les deux formes d'agriculture coexistent, bien que cette cohabitation ne soit pas toujours harmonieuse. L'expansion de l'agro-industrie menace l'existence de l'agriculture familiale, notamment à travers la pression sur les terres, l'accaparement des ressources et les impacts environnementaux qui en résultent.

Le soja illustre parfaitement cette tendance. En 2018, le Brésil produisait 114,6 millions de tonnes, soit 31 % de la production mondiale, avec environ 70 % destinés à l'exportation. Cette dépendance aux marchés internationaux rend le marché intérieur brésilien extrêmement sensible aux variations des prix internationaux. Comme (Gimé C. , 2009) le souligne dans son analyse de la sensibilité des économies du Mercosur, une hausse des prix agricoles mondiaux entraîne généralement une augmentation de la production agricole brésilienne, ce qui peut significativement influencer les prix à la production des produits destinés à l'exportation.

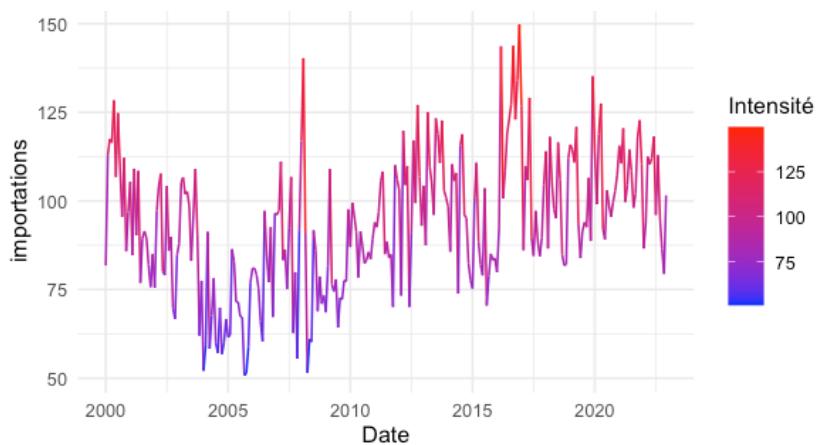
En contraste, l'agriculture familiale brésilienne, qui représente environ 70 % de la consommation alimentaire domestique et inclut principalement la culture de haricots, de manioc, ainsi que la production de lait, de maïs et de porc, est affectée par l'expansion de l'agrobusiness. Bien que moins directement connectée aux marchés internationaux, elle subit la concurrence pour l'accès aux terres et aux ressources en eau, compromettant ainsi sa durabilité et sa capacité à subvenir aux besoins alimentaires locaux. Ces pressions

augmentent les fluctuations des prix à la production, influencées par les dynamiques d'exportation. L'indice des exportations, dès lors, ne reflète pas seulement les variations de la demande internationale, mais expose aussi les tensions internes au sein du secteur agricole brésilien.

4.3.4 Importations pour l'agriculture et l'élevage

Figure 8 Importation des produits agricoles

Jan 2000 - Déc 2022



La Figure 8 illustre les variations des importations de produits agricoles au Brésil de janvier 2000 à décembre 2022. Cet indice, ajusté de manière que la moyenne de 2018 corresponde à 100, montre des fluctuations notables au fil des années, soulignant l'importance des importations dans la complémentation de la production domestique. Parmi les principaux produits importés figurent le blé, en grande partie en provenance des États-Unis et de l'Argentine, ainsi que le riz et des intrants essentiels tels que les engrains et les pesticides, massivement importés des États-Unis, de la Chine, de l'Allemagne et de la Russie pour soutenir l'agriculture intensive du pays.

L'introduction sur le marché brésilien de produits agricoles importés à des prix compétitifs peut entraîner un effet de substitution, pressant les prix locaux à la baisse et influençant ainsi négativement les revenus des producteurs locaux. Cette dynamique est illustrée par les ajustements des tarifs douaniers sur les importations agricoles en 2022, une réponse à l'inflation provoquée par la pandémie de COVID-19 et les tensions géopolitiques telles que la guerre en Ukraine. Les droits d'importation, précédemment

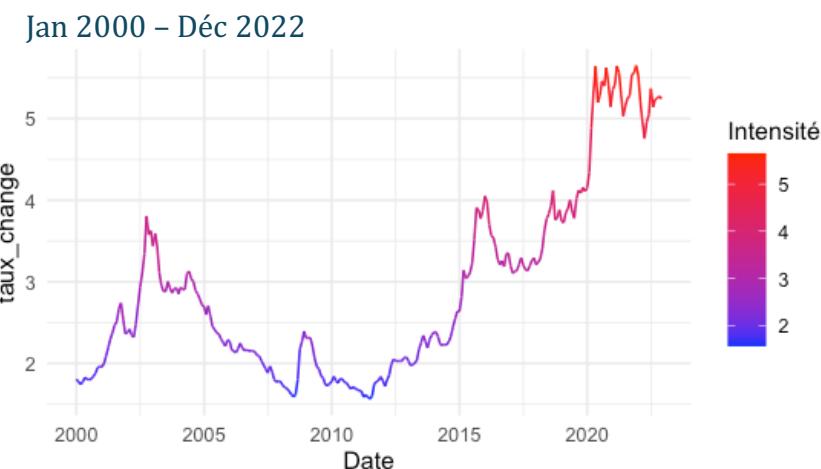
fixés jusqu'à 10 % pour certains produits agricoles, ont été réduits à zéro jusqu'en 2023 pour les importations hors Mercosur, dans un effort pour modérer l'inflation des prix alimentaires (OECD, 2023).

En outre, les importations jouent un rôle crucial dans la stabilisation de l'offre de produits agricoles, particulièrement en périodes de pénurie dues à des conditions climatiques défavorables ou des fluctuations saisonnières. En comblant les déficits de production, les importations aident à stabiliser les prix du marché, offrant ainsi une certaine sécurité alimentaire et économique au pays.

4.3.5 Taux de change

Le taux de change nominal "venda" – moyenne mensuelle, représente le prix moyen auquel le réal brésilien est échangé contre le dollar américain sur le marché des changes durant un mois donné. Ce taux est utilisé pour estimer le coût moyen auquel les entreprises et les consommateurs brésiliens peuvent acheter des dollars américains tout au long du mois.

Figure 9 Taux de change - R\$ / US\$



La Figure 9 nous présente l'évolution du taux de change du real brésilien de janvier 2000 à décembre 2022.

La dépréciation du real a rendu les produits agricoles brésiliens plus compétitifs sur les marchés internationaux, permettant aux producteurs de vendre à des prix plus élevés en monnaie locale tout en restant attractifs en dollars.

Cette situation a favorisé une croissance des exportations, notamment entre 2014 et 2016, malgré une récession globale, en grande partie grâce à la demande croissante de la Chine.

En 2018, une nouvelle hausse des exportations vers la Chine a coïncidé avec une dépréciation continue du real, tandis que l'année 2020, marquée par la crise de la COVID-19, a encore accentué la dépréciation du real et les défis économiques du pays.

Bien que cela ait renforcé la compétitivité des exportations, cela a également accru les coûts des intrants agricoles importés, comme les engrains et les équipements, réduisant ainsi les bénéfices nets des agriculteurs malgré des revenus en hausse en monnaie locale (C. Valdes, 2020).

Les fluctuations du taux de change influencent les prix à la production agricole par divers canaux. Une dépréciation du real rend les produits agricoles brésiliens plus compétitifs à l'international, stimulant les exportations et augmentant la production locale pour répondre à la demande mondiale. Cependant, cette dynamique est tempérée par la hausse des coûts des intrants importés, tels que les engrains et les pesticides, dont les prix sont libellés en dollars. La dépréciation du real rend ces intrants plus coûteux, augmentant ainsi les coûts de production pour les agriculteurs et réduisant les marges bénéficiaires.

5 Analyse de Données

5.1 Statistiques descriptives

Afin de mieux comprendre les observations de nos variables, nous allons analyser et comparer leurs principales statistiques.

Au Tableau 1, nous pouvons observer que notre variable dépendante, varie considérablement, allant de 191.6 à 2133.2. Cette large amplitude indique une grande volatilité et des changements significatifs dans les prix au producteur au fil du temps. La différence entre la médiane 649.3 et la moyenne 748.4 suggère une distribution asymétrique positive, probablement due à des valeurs extrêmes. Ces dernières seront vérifiées plus tard dans cette section et sont observables dans le graphique de boxplot en Annexe 1.

Tableau 1 Statistiques descriptives

Variables	Minimum	1st Qu	Médian	Moyenne	3rd Qu	Maximum
IPA	191,6	401,5	649,3	748,4	944,3	2133,2
Température	-0,09	0,77	1,06	1,08	1,32	2,43
CDD	2,38	4,72	8,44	10,10	15,31	23,34
Précipitations	33,16	72,14	148,39	147,15	212,66	288,10
Exportations	7,50	31,70	48,40	59,26	82,65	194,1
Importations	50,80	81,00	92,60	93,65	107,47	149,80
Prix du pétrole	18,47	42,53	62,81	65,40	85,50	134,03
Taux de change	1,56	2,00	2,47	2,90	3,52	5,65
SMIC	136	300	545	602,5	937	1212

Concernant les variables climatiques, nous observons que les variations des températures ont une amplitude allant de -0,09°C à 2,43°C. Une moyenne de 1,08°C suggère que, pendant la période étudiée, en moyenne, les augmentations de température sont légèrement au-dessus de 1°C. La médiane étant très proche de la moyenne, cela indique une distribution symétrique autour de cette valeur médiane.

Le nombre maximum de jours secs consécutifs, représenté par la variable CDD, varie de 2 à 23 jours, indiquant des variations significatives dans les conditions climatiques d'une année à l'autre. De manière similaire, les précipitations varient de 33,16 mm à 288,10 mm.

Les variables économiques nous permettent d'observer les évolutions significatives des coûts et des politiques économiques. Le prix du pétrole a augmenté considérablement, passant de 18,47 à 134,03. Parallèlement, le salaire minimum a été augmenté de 136 R\$ à 1353,8 R\$, ce qui pourrait nous indiquer une politique de soutien au pouvoir d'achat ou une réaction à l'inflation.

En outre, l'étendue des exportations, variant de 7,50 à 194,1, révèle une volatilité notable qui pourrait être attribuée à des changements dans la demande mondiale, des accords commerciaux, ou des variations de compétitivité des prix.

5.2 Détection et corrections des valeurs atypiques

Les valeurs atypiques, ou outliers, définies par (Grubbs, 1969) comme étant des observations qui « semblent dévier notoirement des autres éléments de l'échantillon auquel elles appartiennent », peuvent altérer de manière significative les estimations des paramètres, rendant les modèles moins précis et potentiellement trompeurs.

Ces points atypiques peuvent aussi fausser les résultats des processus de désaisonnalisation et des méthodes d'estimation. En perturbant les estimations des paramètres, ils introduisent des biais significatifs qui peuvent compromettre l'interprétation des modèles. Par exemple, dans les modèles de régression MCO, la présence non détectée d'outliers peut influer sur la pente et l'interception estimées, conduisant à des conclusions erronées.

Pour identifier et corriger ces anomalies, nous utilisons la fonction `tso` du package `tsoutliers`, conçue pour détecter divers types d'outliers, qu'ils soient transitoires ou permanents (Darné, 2024).

Cette fonction prend en compte cinq types de points atypiques : outliers additifs ("AO"), changements de niveau ("LS"), changements temporaires ("TC"), outliers innovants ("IO") et changements de niveau saisonniers ("SLS").

La fonction opère par itération et utilise par défaut la fonction `auto.arima` pour ajuster des modèles ARIMA qui aident à identifier ces outliers. Cette intégration peut révéler des anomalies que des tests basés strictement sur des mesures statistiques comme le test de Rosner pourraient manquer, surtout si les outliers se manifestent à travers leur impact sur la dynamique temporelle.

Le Tableau 2 présente un exemple de la série de IPA, montrant comment les valeurs sont identifiées et classées selon leur type d'atypicité.

Tableau 2 Outliers série IPA

Type	Dates	Coefhat	Tstat
LS	2020/12	-149,69	-10,973
TC	2021/06	-77,92	-6,611
TC	2021/11	-42,75	-3,655

Le premier point concerne un outlier de type LS (Level Shift), qui signale un changement brusque et significatif dans le niveau de la série. Plusieurs facteurs, observés en décembre 2020, peuvent expliquer ce phénomène. L'impact le plus évident est celui de la pandémie de COVID-19, qui a profondément affecté les économies mondiales.

Le Brésil, comme de nombreux autres pays, a connu des perturbations majeures dans ses chaînes d'approvisionnement. Non seulement l'économie brésilienne a été touchée, mais cela a également été une catastrophe sur le plan social, augmentant le nombre de Brésiliens vivant dans la pauvreté et l'insécurité alimentaire. Les mesures de confinement et les restrictions de déplacement ont perturbé la production et la distribution. Par ailleurs, le grand nombre de contaminations dans le pays a conduit à la fermeture de nombreuses fermes et usines, affectant directement les prix de production. Ces fluctuations ont particulièrement affecté les denrées alimentaires, rendant l'accès à une nourriture abordable encore plus précaire pour les personnes les plus vulnérables de la population (R. Ribeiro-Silva, 2020).

Par ailleurs, la fin de l'année 2020 a été marquée par des conditions météorologiques extrêmes, notamment des sécheresses affectant certaines régions agricoles du Brésil. Ces conditions ont non seulement impacté les rendements, mais ont aussi rendu le transport fluvial impraticable, particulièrement sur le fleuve Paraná, obligeant un retour au transport routier, qui nettement plus coûteux (Carfantan, 2021).

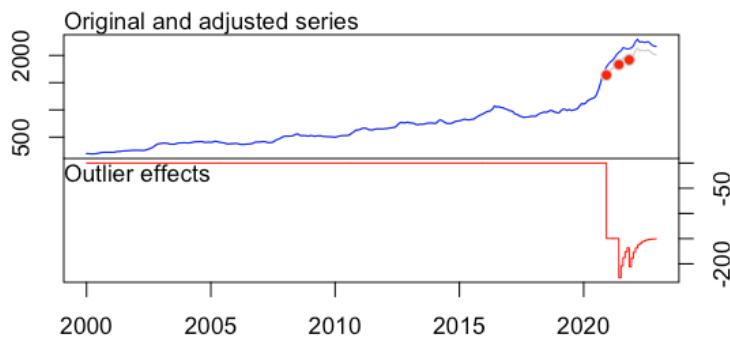
Nous identifions ensuite deux outliers de type TC (Temporary Change) dans la série, correspondant à juin 2021 et novembre 2021. En juin, durant la période de récolte pour plusieurs cultures importantes au Brésil, une sévère sécheresse, déclenchée par le

phénomène climatique la Niña, a causé des fluctuations significatives dans les rendements. Cette sécheresse a également menacé l'approvisionnement en électricité du pays, qui dépend largement de ses centrales hydroélectriques, entraînant une augmentation des coûts énergétiques. En réponse, le gouvernement avait activé son tarif le plus élevé pour le mois de juin (Le Monde avec AFP, 2021).

En fin d'année, en novembre, le pays a été secoué par des tensions politiques et des incertitudes économiques, notamment autour des politiques fiscales et des manifestations. Ces instabilités ont pu influencer les marchés agricoles, engendrant ainsi des fluctuations de prix supplémentaires.

Une fois les valeurs atypiques identifiées, nous procédons à leur correction, comme illustré dans la Figure 10. Les séries sans outliers, telles que celles concernant le CDD, le pétrole et la température, restent inchangées. Les valeurs atypiques détectées dans les autres séries sont détaillées dans l'Annexe 2.

Figure 10 Détection des points atypiques sur la série niveau Série IPA



Nous poursuivons notre analyse avec cette nouvelle série corrigée, désormais sans outliers.

5.2.1 Skewness, kurtosis et test de normalité de la série ajustée

Nous allons maintenant nous intéresser à la distribution de nos séries corrigées, en utilisant plusieurs mesures.

D'abord, le skewness, ou asymétrie, qui mesure la disposition des données autour de leur moyenne. En termes simples, il indique si les données s'étendent plus loin d'un côté de la moyenne que de l'autre. Lorsque le Skewness est positif, cela veut dire que la queue de la distribution s'étend plus loin à droite, indiquant plus de valeurs extrêmes au-dessus de la

moyenne. Un Skewness négatif, nous indique que la queue s'étend plus loin à gauche, avec plus de valeurs extrêmes en dessous de la moyenne et s'il est nul, la distribution est relativement symétrique autour de la moyenne.

Concernant le kurtosis, cette mesure évalue le degré de concentration des données autour de la moyenne de la distribution. Un kurtosis élevé révèle une distribution avec une pointe marquée et des queues épaisses, indiquant une forte concentration de valeurs proches de la moyenne ainsi que la présence fréquente de valeurs extrêmes. À l'inverse, un kurtosis bas suggère une distribution plus aplatie avec moins de valeurs près de la moyenne et moins d'extrêmes.

Enfin, nous utiliserons le test de Shapiro-Wilk pour évaluer si nos données suivent une distribution normale. Ce test vérifie l'hypothèse de normalité des données : une p-value inférieure à 0.05 indique un rejet de l'hypothèse normale, tandis qu'une p-value élevée montre qu'il n'y a pas suffisamment de preuves pour rejeter cette hypothèse, suggérant que les données pourraient être normalement distribuées.

Tableau 3 Skewness – Kurtosis - Normalité

Variables	Skewness	Kurtosis	Shapiro-wilk
	P value		
IPA ($\times 10^{-16}$)	1.587	2.029	2.200
CDD ($\times 10^{-12}$)	0.515	-1.092	2.662
Précipitation ($\times 10^{-10}$)	0.003	-1.399	2.312
Température	0.354	0.246	0.027
Exportations ($\times 10^{-10}$)	0.991	0.835	5.389
Importations	-0.035	-0.489	0.234
Taux change ($\times 10^{-14}$)	0.982	-0.046	3.167
Smic ($\times 10^{-5}$)	0.038	-0.444	2.005
Pétrole ($\times 10^{-8}$)	0.340	-0.940	3.449

Les données de skewness révèlent divers degrés d'asymétrie dans leurs distributions. L'IPA présente une asymétrie positive marquée, avec un skewness de 1.587, caractérisée par une longue queue s'étendant vers des valeurs plus élevées. Cela indique une concentration de données inférieures à la moyenne (748), visible sur l'histogramme des séries présenté en Annexe 3 D'ailleurs, les valeurs en dessous de la moyenne se situent toutes avant 2015 ; après cette année, nous observons une tendance à des valeurs de plus en plus élevées.

Le CDD et la Température affichent des asymétries positives plus légères, avec des skewness de 0.515 et 0.354 respectivement, suggérant des queues étendues vers les valeurs supérieures mais de manière moins prononcée. Les précipitations et les importations, avec des skewness de 0.003 et -0.036, démontrent des distributions presque symétriques, indiquant que les données sont uniformément réparties autour de la moyenne. Les exportations, le taux de change, et le pétrole montrent des asymétries modérées (skewness de 0.991, 0.982, et 0.340 respectivement), indiquant que ces variables ont des distributions avec des queues un peu plus longues vers les valeurs élevées. Enfin, le SMIC montre une distribution pratiquement symétrique avec un skewness de 0.038, indiquant une répartition équilibrée des valeurs autour de la moyenne.

Les valeurs de kurtosis (aplatissement) complètent l'analyse des distributions. L'IPA, avec une kurtosis de 2.029, présente une distribution légèrement plus aplatie que la normale, ce qui indique moins de concentration de données autour de la moyenne et des extrêmes moins prononcés.

Le CDD et les précipitations ont des kurtosis de -1.092 et -1.400 respectivement, indiquant des distributions bien plus plates, signifiant que les données sont plus étaillées et moins concentrées autour de la moyenne. La température, avec une kurtosis de 0.246, et les exportations, avec une kurtosis de 0,836, présentent des distributions proches de la normale mais avec une concentration légèrement supérieure autour de la moyenne pour les exportations.

Les importations, avec une kurtosis de -0.490, et le pétrole, avec une kurtosis de -0.941, présentent des distributions aplatises. Le taux de change, avec une kurtosis de -0.047, montre une distribution qui se rapproche de la normale mais qui est légèrement plus

dispersée. Enfin, le SMIC, avec une kurtosis de -0.444, révèle aussi une distribution un peu aplatie.

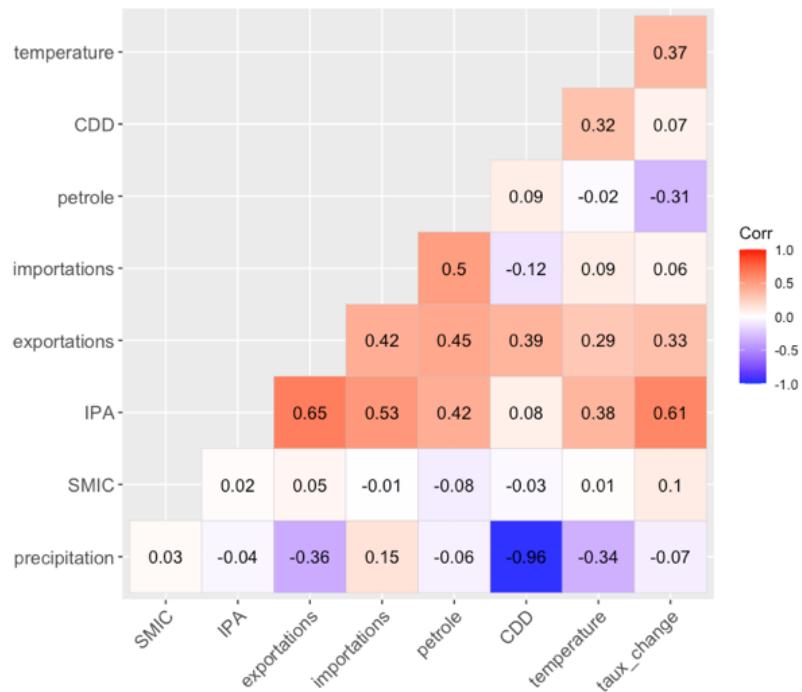
L'analyse de normalité, basée sur les p-values du test de Shapiro-Wilk, confirme que la majorité des variables, à l'exception des importations (p-value de 0.2348), ne correspondent pas à une distribution normale. Des p-values particulièrement basses pour l'IPA (2.2e-16), le CDD (2.662e-12), les précipitations (2.312e-10), la température (0.02765), les exportations (5.389e-10), le taux de change (3.167e-14), le SMIC (2.005e-05) et le pétrole (3.449e-08) signalent des écarts significatifs par rapport à la normalité.

5.3 Corrélation

Nous allons analyser certaines corrélations significatives pour mieux comprendre les interactions entre nos variables, en particulier pour déterminer si elles varient de manière similaire dans le temps.

Étant donné que plusieurs de nos variables ne suivent pas une loi normale, nous utilisons pour le calcul méthode Spearman Annexe 3, car ce type de corrélation, est dite robuste, ne dépend pas de la distribution des données.

**Figure 11 Matrice de corrélation
Séries ajustées**



Pour commencer, nous pouvons constater sur la Figure 11 les corrélations entre nos variables climatiques, telles que CDD et précipitations, avec une très forte corrélation négative de -0.96. Cette corrélation est logique car des périodes prolongées sans pluie coïncident avec une diminution des précipitations globales. Plus il y a de jours secs consécutifs, moins il y a de précipitations durant ces périodes.

Concernant les variables économiques, nous observons que le prix du pétrole et le taux de change ont une corrélation négative, suggérant que lorsque le prix du pétrole augmente, le taux de change tend à baisser. Cela pourrait refléter un renforcement de la monnaie locale dû à une augmentation des revenus d'exportation de pétrole.

La corrélation positive de 0.65 entre l'IPA et les exportations est très significative. Lorsque la demande internationale pour les produits agricoles brésiliens augmente, les prix à la production augmentent aussi. De plus, la corrélation entre l'IPA et le taux de change, à 0.61, est également notable. Une dépréciation du réal brésilien rend les exportations plus compétitives, ce qui peut stimuler la demande extérieure et faire monter les prix agricoles locaux.

Les variations de température et les précipitations montrent une corrélation négative modérée, ce qui pourrait indiquer que les variations plus importantes de température sont associées à des périodes de moindre précipitation, ou que les anomalies de température se constatent plutôt lors de périodes sèches.

Bien que la corrélation fournit des indications sur les relations possibles, l'analyse de corrélation est limitée lorsque nous travaillons sur des séries temporelles influencées par des tendances à long terme (comme la croissance économique ou le changement climatique) et des effets saisonniers. Ces composantes peuvent fausser les corrélations observées. Des analyses causales plus approfondies sont nécessaires pour établir des liens de cause à effet entre nos variables.

6 Estimation économétriques et interprétation des résultats

Pour progresser de manière structurée dans notre analyse, nous adopterons une approche méthodique et détaillée.

D'abord, nous examinerons la série pour déterminer si elle comporte des composantes saisonnières en utilisant un test de saisonnalité. Ensuite, nous emploierons le test ADF pour vérifier la présence d'une tendance stochastique, ce qui nous indiquera si la tendance des données est due à une véritable tendance ou simplement à une dépendance aléatoire dans le temps.

De plus, nous appliquerons le test KPSS pour évaluer si la série est stationnaire tant au niveau que dans sa tendance, un critère important pour la validité de nos modèles. Si ces tests révèlent des saisonnalités ou des tendances stochastiques, des corrections adéquates seront mises en place. Une fois ces ajustements effectués, nous procéderons à une nouvelle vérification pour nous assurer que la série ajustée est appropriée pour une modélisation.

Le traitement de la tendance déterministe sera abordé de deux manières. La première méthode consistera à corriger individuellement chaque série pour éliminer les tendances déterministes via une régression linéaire. La seconde méthode intégrera une variable temporelle dans le modèle, ce qui nous permettra de modéliser la tendance sans l'éliminer explicitement.

Ces étapes nous aideront à préparer les données pour une interprétation correcte de nos modèles. Si les hypothèses ne sont pas validées après ces corrections, nous réévaluons notre stratégie et ajusterons notre approche.

6.1 Identification des composants

6.1.1 Test de saisonnalité

Nous allons commencer par effectuer des tests de saisonnalité sur nos variables pour mieux identifier et comprendre leur impact. Cela nous permet d'isoler et de corriger les

effets saisonniers, pour une modélisation sans les variations périodiques. À cette fin, nous avons employé deux tests spécifiques du package seastests :

Le premier, le test seasdum, utilise des variables indicatrices (ou dummies) pour chaque période saisonnière de la série temporelle, permettant de détecter la présence de motifs saisonniers réguliers. Ce test évalue si l'amplitude des variations saisonnières est statistiquement significative.

Le second, le test isSeasonal, offre une approche plus directe pour évaluer la saisonnalité. Il analyse la série temporelle pour déterminer si une structure saisonnière est présente, retournant une valeur booléenne qui indique clairement l'existence ou non de la saisonnalité.

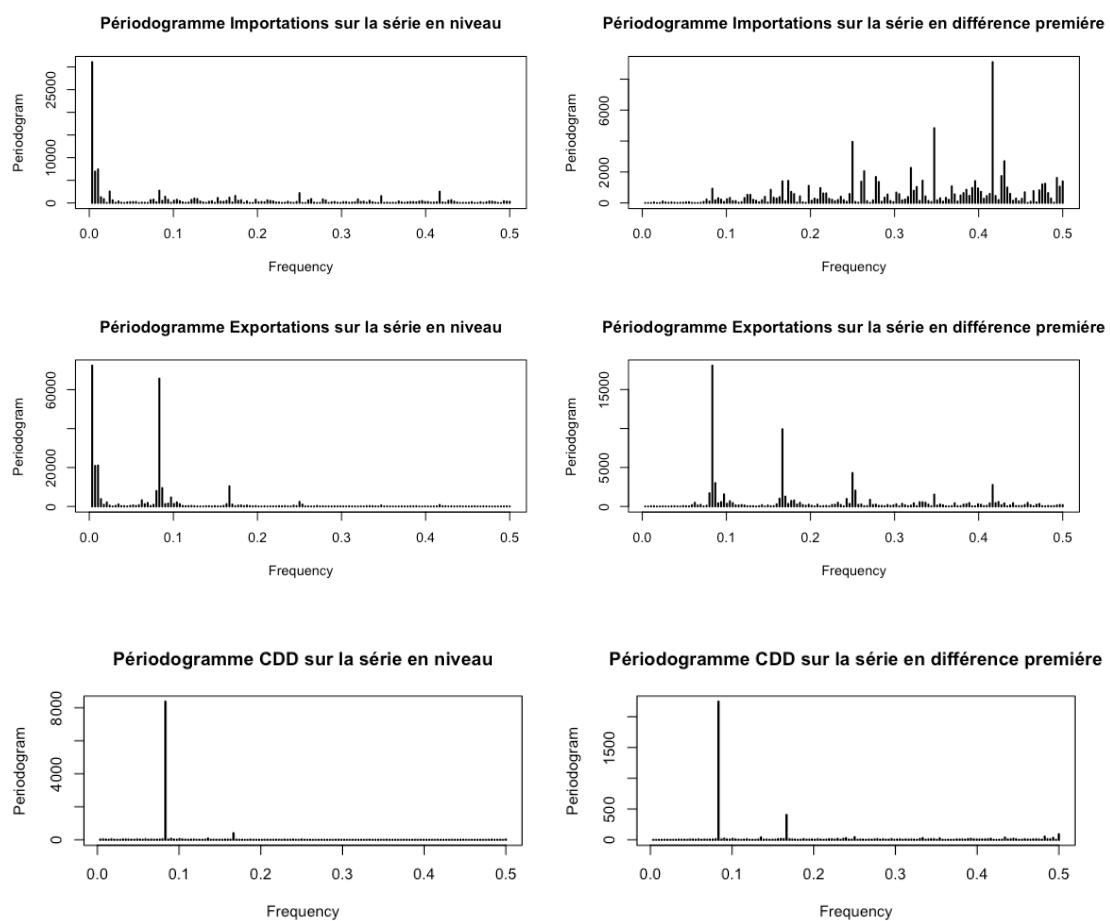
Tableau 4 Tests de saisonnalité

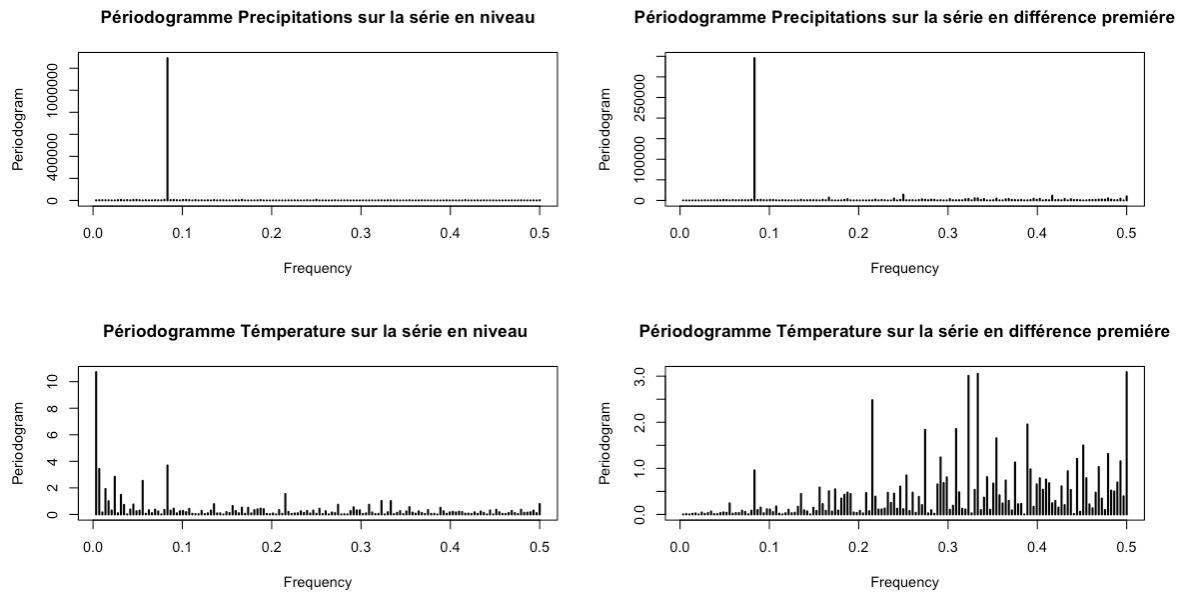
Variables	Test	Seasonal
	isSeasonal	dummies
		P_value
IPA	False	0.24
CDD	True	0
Importations ($\times 10^{-9}$)	True	2.99
Exportations	True	0
Pétrole	False	0.39
Taux change	False	0.21
Smic	False	0.99
Précipitation	True	0
Température (10^{-6})	True	5.41

Les résultats présentés dans le Tableau 4 montrent que les variables telles que CDD, importations, exportations, précipitations et température affichent de forts signes de saisonnalité, nécessitant des ajustements saisonniers avant toute analyse ultérieure. En revanche, l'IPA, le pétrole, le taux de change et le SMIC ne semblent pas présenter de saisonnalité significative.

En complément de nos tests, nous utilisons le corrélogramme des séries (Annexe 4), ainsi que la fonction d'autocorrélation (ACF), qui illustrent la corrélation des données avec elles-mêmes à différents délais. À des fins d'analyse comparative, nous mettons en parallèle l'ACF de la série au niveau brut et celle de la série différenciée. Les pics observés dans le corrélogramme à certains délais indiquent une saisonnalité. Cette comparaison nous aide aussi à réduire le bruit de fond dans le périodogramme, comme le montre Figure 12. Grâce à cette réduction du bruit, les pics de fréquence saisonnière deviennent plus apparents, permettant une observation plus précise des variations saisonnières

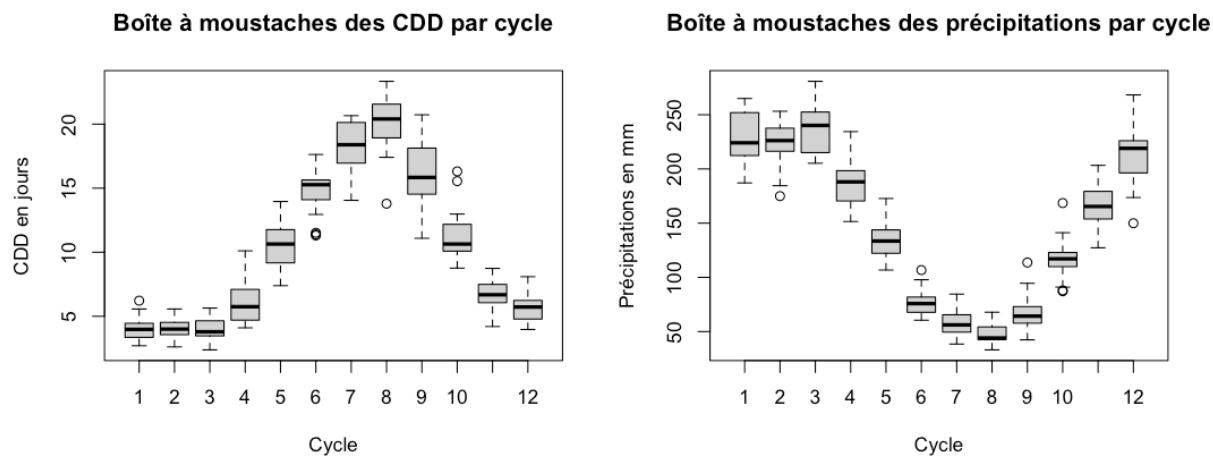
Figure 12 Periodogramme variables saisonnières





Nous pouvons également observer la distribution des séries climatiques par mois pour visualiser les cycles annuels, chaque année étant considérée comme un cycle composé de 12 observations. Cette analyse nous permet d'identifier clairement les périodes sèches, définies précédemment entre les mois 5 à 8, et de constater leurs fluctuations saisonnières. Dans la section dédiée à la présentation des données, nous avons discuté de ces périodes sèches et les graphiques illustrent bien ces variations. Pour la variable des précipitations, on observe également que les médianes varient, ce qui signale une distribution inégale des précipitations tout au long de l'année.

Figure 13 Boxplot variables précipitation et CDD



6.1.2 Analyse de la tendance

Lorsque nous examinons les composants d'une série temporelle, nous identifions souvent ce que l'on appelle une tendance monotone. Pour analyser cette tendance nous employons le test de Mann-Kendall. Il est important de noter que, bien que le test de Mann-Kendall puisse confirmer la présence d'une tendance, il ne fournit pas d'informations directes sur la nature spécifique de cette tendance, qu'elle soit linéaire, non linéaire, déterministe ou stochastique. Toutefois, la mise en évidence d'une tendance significative par ce test indique généralement une variation constante, soit en augmentation, soit en diminution, au fil du temps.

Tableau 5 Test de tendance Mann-Kendall
Série ajustée

Variable	P_value
IPA ($\times 10^{-16}$)	2.200
CDD	0.074
Précipitation	0.810
Température ($\times 10^{-11}$)	8.22
Exportations ($\times 10^{-16}$)	2.200
Importations ($\times 10^{-16}$)	2.200
Taux de change ($\times 10^{-16}$)	2.200
SMIC	0.281
Pétrole ($\times 10^{-16}$)	2.200

Les résultats du Tableau 5 indiquent que la majorité des variables montrent une tendance positive significative, à l'exception des précipitations et du SMIC. Initialement, le SMIC semblait présenter une tendance, mais il s'est avéré que cette tendance était principalement influencée par des valeurs atypiques, après ajustement la série corrigée ne montre plus de tendance significative.

Nous allons maintenant tester la stationnarité en utilisant le test ADF. Ce test vérifie la présence d'une racine unitaire, indiquant une tendance stochastique.

Une p-value inférieure à 0.05 nous permet de rejeter l'hypothèse de non-stationnarité à un niveau de confiance de 95%, suggérant que la série temporelle est stationnaire.

Tableau 6 Résultats du test de stationnarité ADF

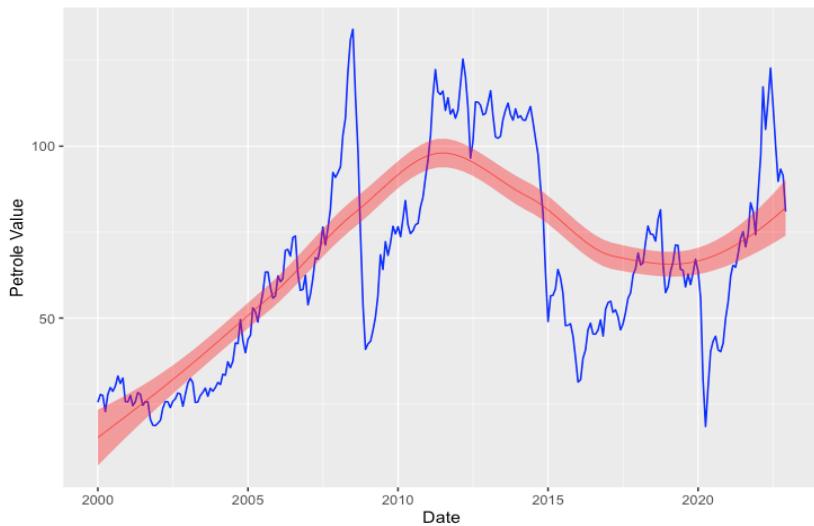
Variable	P-value
IPA	0.54
Température	0.01
CDD	0.01
Précipitations	0.01
Exportations	0.01
Importations	0.02
Prix du pétrole	0.50
Taux de change	0.79
SMIC	0.01

D'après le Tableau 6, les variables IPA, taux de change et pétrole ne sont pas stationnaires, leurs p-values excédant le seuil de 0.05. Cela indique la présence de tendances stochastiques.

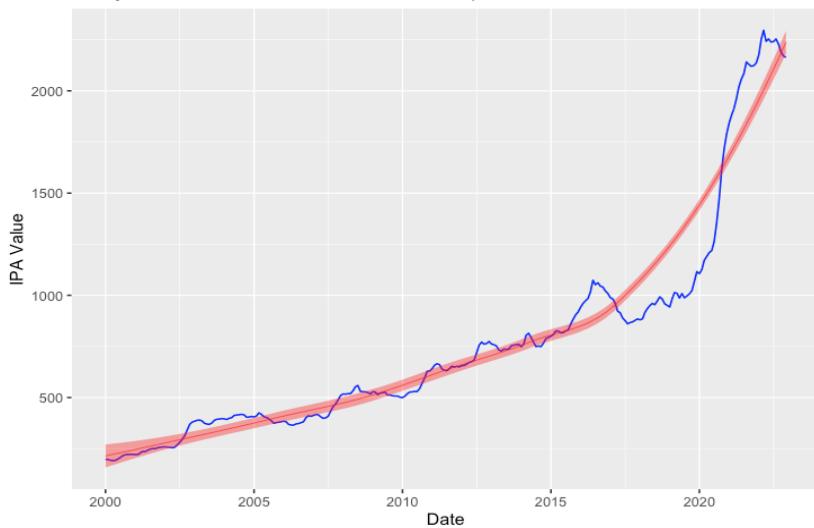
Graphiquement, cette tendance, et donc la non-stationnarité, peuvent être observés dans la Figure 14. La ligne rouge, obtenue par la technique Loess, permet de discerner la tendance globale au fil du temps. Cette méthode lisse les fluctuations mensuelles pour offrir une vue plus nette de l'évolution à long terme de la série.

Figure 14 Tendances - Séries non stationnaires

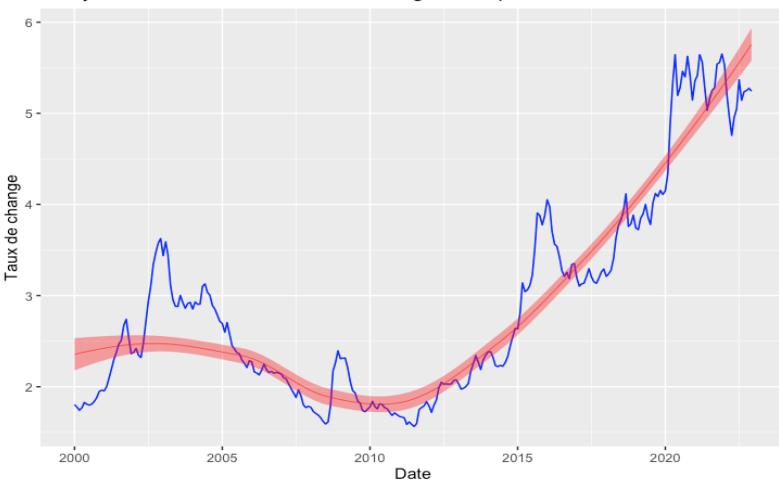
Analyse de la tendance du prix du pétrole sur la période 01/2000 - 12/2022



Analyse de la tendance de l'IPA-DI sur la période 01/2000 - 12/2022



Analyse de la tendance du taux de change sur la période 01/2000 - 12/2022



Poursuivant l'analyse de nos séries, nous allons vérifier deux aspects de la stationnarité à l'aide du test KPSS, qui diffèrent du test ADF par son approche inverse.

Le test KPSS permet d'examiner si une série est stationnaire autour d'une moyenne constante (option "Level") ou d'une tendance déterministe (option "Trend"). Avec l'hypothèse nulle que la série est stationnaire dans chaque cas.

Cependant le test KPSS présente une nuance importante en ce qui concerne l'évaluation de la stationnarité en présence de tendances, et il est important de comprendre ce que les résultats signifient, car l'hypothèse nulle du test lorsqu'il est configuré pour examiner la 'stationnarité en trend', est que la série est stationnaire autour d'une tendance.

Si le test ne rejette pas l'hypothèse nulle (c'est-à-dire, si la p-value est supérieure à un seuil de 0.05), cela signifie que la série peut être considérée comme stationnaire une fois que la tendance déterministe est prise en compte. En d'autres termes, la série présente une tendance, mais les variations autour de cette tendance sont stables et prévisibles à long terme.

Tableau 7 KPSS test

Variables	Kpss test trend	Kpss test level
IPA	0.01	0.01
Température	0.069	0.01
CDD	0.1	0.1
Précipitations	0.1	0.1
Exportations	0.1	0.01
Importations	0.01	0.01
Prix du pétrole	0.01	0.01
Taux de change	0.01	0.01
SMIC	0.1	0.1

6.2 Données corrigées des variations saisonnières et tendances

Maintenant que nous avons une meilleure compréhension de la nature de nos données, nous allons adopter des stratégies adaptées à chaque composante de nos séries. Les variations saisonnières, par exemple, peuvent introduire du bruit et masquer les véritables dynamiques à long terme. Dans un modèle de régression linéaire, la présence de saisonnalité non corrigée peut entraîner une autocorrélation des résidus, violant ainsi l'hypothèse fondamentale d'indépendance des erreurs.

En corrigeant à la fois les effets saisonniers et les tendances, les modèles deviennent plus simples à analyser, sans être perturbés par les fluctuations saisonnières en plus des tendances-cycles plus larges.

Comme le soulignait (Granger, 1978) « la saison est traitée de cette manière, parce qu'elle est économiquement sans importance ». Cela démontre la nécessité de ces corrections pour une analyse économique rigoureuse.

Avant de corriger cette saisonnalité, nous devons déterminer si elle est de nature additive ou multiplicative. Pour ce faire, nous utiliserons le test log-level du package RJDemetra, en employant sa fonction regarima_x13, afin d'évaluer si une transformation logarithmique des données est appropriée. Cette transformation est particulièrement utile pour stabiliser la variance des données.

Tableau 8 Test Log level

TFUNCTION	ADJUST	AICDIFF
AUTO	None	-2

Le Tableau 8 affiche des résultats uniformes pour toutes nos données saisonnières, indiquant un AICc de -2. Ce résultat suggère une préférence pour un modèle sans transformation logarithmique. En effet, l'absence d'amélioration du modèle par la transformation logarithmique, selon les critères de l'AICc, implique que les variations

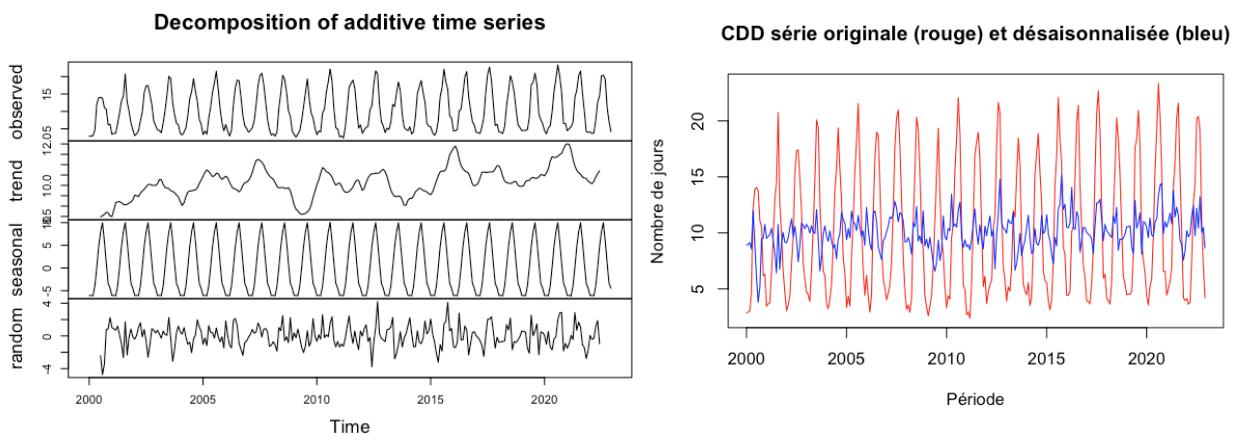
saisonnieres ne sont pas proportionnelles au niveau de la série. Ainsi, nous observons des saisonnalités de type additive, où les effets saisonniers restent constants et ne dépendent pas de l'échelle des données. De plus, le paramètre 'tfunction = Auto' dans notre test signifie que l'évaluation de la nécessité d'une transformation log-level est automatisée, tandis que 'adjust = None' indique que notre analyse ne requiert pas d'ajustements pour les effets des années bissextiles, simplifiant ainsi notre approche.

Parmi les différentes méthodes disponibles pour corriger la saisonnalité dans les séries temporelles, nous optons pour la méthode de décomposition.

Cette technique nous permet de décomposer la série temporelle en ses composantes de base : la tendance, la saisonnalité et le résidu (exemple avec la série CDD Figure 15 – gauche).

En particulier, dans le cas d'une saisonnalité additive, cette approche nous permet de soustraire efficacement la composante saisonnière de la série brute (voir Figure 15 – droite).

Figure 15 Décomposition et correction de la série CDD



Après avoir soustrait la composante saisonnière, la tendance et les irrégularités restent visibles, offrant une vue claire de l'évolution sous-jacente de la série.

Cette technique de désaisonnaliation est appliquée à toutes les séries concernées, assurant que chaque série est analysée dans des conditions comparables et sans les distorsions introduites par les variations saisonnières.

Concernant la non-stationnarité de certaines de nos séries, nous procérons à une différenciation en réponse aux résultats du test ADF afin de les rendre stationnaires.

Tableau 9 Résumé stationnarité

Variable	ADF	KPSS
CDD	0.01	0.1
Température	0.01	0.05
Précipitation	0.01	0.1
Importations	0.07	0.01
Exportations	0.01	0.1
IPA	0.01	0.08
Taux de change	0.01	0.1
Pétrole	0.01	0.1
SMIC	0.01	0.1

Le Tableau 9 résume les résultats après correction, indiquant que nos séries sont désormais stationnaires. Il est important de noter qu'après avoir différencié certaines séries pour atteindre la stationnarité, nous avons dû exclure la première observation des autres séries pour aligner le jeu de données. Cette action assure que toutes les séries temporelles analysées démarrent au même point temporel, facilitant ainsi des comparaisons et analyses cohérentes de l'ensemble des données.

6.3 Sélection des variables

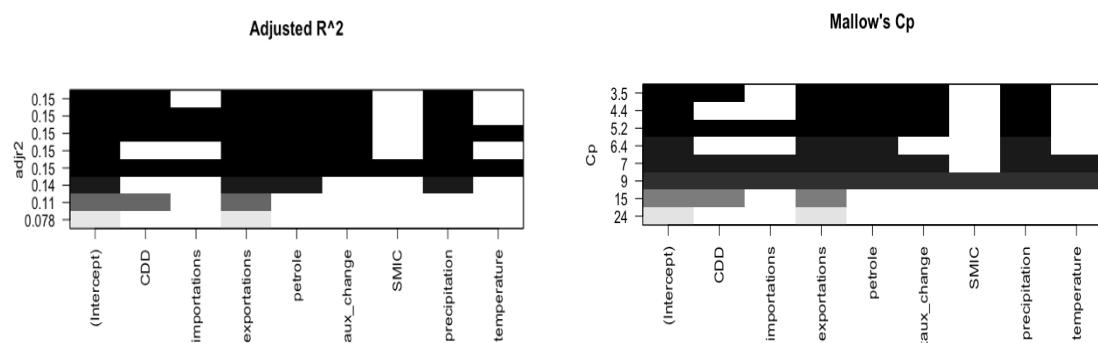
Pour optimiser notre modèle de régression, nous adopterons deux approches distinctes de sélection des variables. La première approche, le Best Subset, consiste à identifier le meilleur modèle de régression linéaire parmi l'ensemble des variables disponibles. Cette méthode, aussi connue sous le nom de sélection de sous-ensembles, explore toutes les

combinaisons possibles de variables afin de déterminer celles qui maximisent la précision de la prédiction de la variable dépendante.

Les critères tels que le R^2 ajusté, le critère de prédiction de Mallows (Cp) et le critère d'information bayésien (BIC) sont utilisés pour juger la performance de chaque modèle.

Selon ces critères, le cinquième modèle Figure 16, qui inclut les variables taux de change, CDD, exportations, pétrole et précipitation, offre le plus haut R^2 ajusté et présente également le plus bas Cp. Ce modèle semble offrir un bon équilibre entre parcimonie et capacité explicative.

Figure 16 Best Subset



La deuxième approche employée est la méthode stepwise avec le critère de l'information Akaike (AIC) comme guide pour la sélection des variables.

Nous utilisons trois stratégies de sélection : ascendante, descendante et double direction. Il apparaît que ces trois méthodes, malgré leurs différences conceptuelles, convergent vers le même modèle final, indiquant une cohérence dans la sélection des variables qui optimisent l'AIC.

Le modèle final retenu comprend aussi les variables exportations, CDD, pétrole, taux change, et précipitation. Avec un AIC de 1653.8, ce modèle permet un équilibre entre la complexité et la capacité à expliquer la variabilité de la variable dépendante, IPA.

6.4 Estimations et validation des modèles MCO

Après avoir sélectionné les variables significatives via les méthodes de sélection, nous passons maintenant à l'étape de modélisation avec le premier modèle retenu utilisant la méthode des moindres carrés ordinaires.

6.4.1 Modèle I

Le modèle inclut les variables exportations, CDD, pétrole, taux change, et précipitation.

D'après l'estimation du 1^{er} modèle les variables sont significatif à plusieurs niveaux de significativité. Le R² est de 0.1691 signifie que 16.91% de la variabilité de l'IPA est expliquée par les variables du modèle. Et avec une p-value de 1.316e-09 il est statistiquement significatif.

Avant de procéder à l'interprétation des résultats de ce modèle, il est crucial de vérifier si les hypothèses sous-jacentes à la régression linéaire sont respectées.

Ces vérifications incluent comme nous l'avons précédemment vu, la normalité des résidus, l'homoscédasticité, l'absence de multicolinéarité entre les prédicteurs, et l'indépendance des erreurs. Ces tests sont essentiels car une violation de ces hypothèses peut compromettre la validité des estimations des coefficients et des tests statistiques associés au modèle.

Test des hypothèses :

1. Le test de Kolmogorov-Smirnov (KS) pour la normalité des résidus donne une statistique D de 0.113 avec une p-value de 0.001. Cela indique que la distribution des résidus s'écarte significativement de la distribution normale. Le test de normalité de Kolmogorov-Smirnov est équivalent au test de Shapiro-Wilk, mais il est plus adapté pour les échantillons de taille ≥ 50 .
2. Le test de Breusch-Pagan pour l'hétéroscédasticité montre une p-value de 0.000, suggérant fortement la présence d'hétéroscédasticité dans les résidus. Cela signifie que les variances des erreurs ne sont pas constantes et varient avec les niveaux des variables prédictives, ce qui peut affecter la fiabilité des estimations standard des erreurs et donc des tests de significativité.
3. Le test RESET pour la spécification du modèle donne une p-value de 0.001. Ce résultat indique des problèmes avec la forme fonctionnelle du modèle.
4. Les valeurs de VIF, les variables sont toutes en dessous de 5, ce qui est considéré comme acceptable pour indiquer l'absence de multicolinéarité
5. Le test de Breusch-Godfrey pour l'autocorrélation dans les résidus montre une p-value extrêmement faible $< 2.2\text{e-}16$, indiquant la présence d'autocorrélation

significative. Cela peut être problématique, car cela signifie que les erreurs ne sont pas indépendantes les unes des autres.

Lorsque nous rencontrons des problèmes d'hétéroscédasticité ou de non-normalité des résidus, une approche efficace peut consister à modifier la forme fonctionnelle du modèle.

6.4.2 Modèle II- Forme semi logarithmique

Il est important de noter, lors de l'interprétation, que l'effet d'un paramètre sur la variable dépendante transformée en logarithme s'exprime comme : $\text{Exp}(\text{coefficient estimé du paramètre analysé}) - 1$ (M.TRAVERS, 2023).

Nous transformerons la variable dépendante en calculant d'abord son logarithme, puis en différenciant la série ainsi transformée, désormais notée diff_log_IPA. Cette méthode vise à stabiliser la variance des résidus et à normaliser leur distribution.

Dans un modèle semi-logarithmique, les coefficients des variables indépendantes quantifient l'effet en pourcentage de chaque variable sur la variable dépendante. Cette transformation est particulièrement utile dans les analyses économiques et financières car elle permet de calculer le taux de croissance périodique des séries, offrant souvent une perspective plus pertinente.

En appliquant à nouveau la sélection de variables avec la méthode stepwise, nous avons identifié un modèle qui inclut les variables taux de change, pétrole, précipitations et exportations, avec un AIC de 2108,42. L'estimation de ce second modèle révèle que toutes les variables incluses sont statistiquement significatives à différents niveaux de significativité.

Le coefficient de détermination R^2 de 0.1048 signifie que 10,48% de la variabilité de l'IPA est expliquée par les variables du modèle. De plus, le modèle affiche une p-value très faible de 4.911e-6, ce qui confirme sa significativité.

Test des hypothèses :

1. Test de Kolmogorov-Smirnov: Le p-value de 0.090, est supérieure à 0.05, indiquant un non-rejet de l'hypothèse de normalité.

2. Test de Breusch-Pagan : Avec un p-value de 0.446, il n'y a pas de preuves d'hétérosécédasticité significative dans les résidus, ce qui est une bonne nouvelle pour l'homogénéité des variances.
3. Test RESET : Le p-value de 0.804 indique qu'il n'y a pas de preuve de mauvaise spécification du modèle, ce qui suggère que la forme fonctionnelle choisie est appropriée.
4. VIF : Tous les VIF sont proches de 1, indiquant une absence de multicollinéarité problématique entre les variables prédictives.

Le modèle semi-logarithmique semble avoir résolu les problèmes de normalité des résidus et d'hétérosécédasticité. Cependant, le problème d'autocorrélation des résidus persiste, comme le montre le test de Breusch-Godfrey, qui indique une p-value extrêmement basse de 4.237e-14, confirmant ainsi la présence significative d'autocorrélation dans les résidus.

6.4.3 Modèle III – Avec variable de temps

Comme nous l'avons constaté, les deux modèles précédents présentent un problème d'autocorrélation des résidus. Nous avons également observé que nos variables sont stationnaires par la tendance, d'après le test KPSS. Cela signifie que si la tendance déterministe est corrigée, les données seraient totalement stationnaires. Cependant, jusqu'à présent, nous n'avons pas pris les mesures nécessaires pour traiter cette tendance déterministe. Nous allons donc utiliser deux approches pour essayer de corriger cette tendance et voir si notre modèle des moindres carrés parvient à mieux capter les dynamiques de nos données.

La première approche consiste à inclure une variable de temps, en incluant le temps comme variable explicative, nous pouvons potentiellement capturer cette dynamique, réduisant ainsi l'autocorrélation résiduelle.

Nous avons donc intégré une variable de temps et procédé à une nouvelle sélection de variables. Le modèle résultant est similaire à celui semi-logarithmique précédemment

établi, mais inclut en plus la variable "time". Ce modèle est donc composé des variables taux de change, pétrole, précipitations, exportations, et temps.

L'estimation du troisième modèle montre que toutes les variables, y compris la variable "time", sont statistiquement significatives à différents niveaux de significativité. Le coefficient de détermination R^2 de 0.1265 signifie que 12,65% de la variabilité de l'IPA est expliquée par les variables du modèle. Et avec une p-value faible 7.282e-07, le modèle, dans son ensemble, est statistiquement significatif.

Tests des hypothèses :

1. Test de Kolmogorov-Smirnov: Le test pour la normalité des résidus montre une p-value de 0.1699, indiquant qu'il n'y a pas de preuve statistique pour rejeter l'hypothèse que les résidus suivent une distribution normale. La normalité des résidus n'est pas un problème ici.
2. Test de Breusch-Pagan: Les résultats du test est une p-value de 0.485 indiquent l'absence d'hétérosécédasticité dans les résidus.
3. Test RESET: Une p-value de 0.5088 suggère qu'il n'y a pas de preuve de mauvaise spécification du modèle, ce qui indique que la forme fonctionnelle du modèle peut être appropriée.
4. Les VIF pour time, exportations et d'autres variables sont en général sous le seuil de 5, à l'exception notable de exportations et time qui sont près de 3, indiquant une possible mais pas sévère multicollinéarité.

Malgré les autres diagnostics favorables, le test de Breusch-Godfrey indique une forte présence d'autocorrélation avec p-value de 4.177e-13, ce qui reste problématique. Cela suggère que la dynamique temporelle pourrait encore être insuffisamment modélisée par la variable time actuelle.

6.4.4 Modèle IV - Avec correction de la tendance déterministe

La deuxième approche que nous avons adoptée pour analyser nos séries temporelles consiste à éliminer la tendance déterministe. Nous procédons en régressant chaque série sur le temps et en utilisant les résidus obtenus comme nouvelles séries temporelles.

À la suite de cette transformation, nous avons effectué une sélection de variables et avons établi le modèle suivant : « IPA_detrended en fonction de CDD_detrended, petrole_detrended, exportations_detrended, taux_change_detrended et precipitation_detrended ». Ce modèle présente un AIC de 1653,66.

L'analyse de ce quatrième modèle révèle que toutes les variables sont statistiquement significatives à plusieurs niveaux. Le coefficient de détermination, R^2 est de 0.1173, indiquant que 11,73 % de la variabilité de l'IPA est expliquée par ces variables. Avec une p-value de 2.673e-06, le modèle est statistiquement significatif.

Tests des hypothèses :

1. Normalité des Résidus : La p_value est de 0.002, ce qui signifie que la distribution des résidus s'écarte significativement d'une distribution normale.
2. Hétéroscédasticité : avec une p-value de 0.110, indique l'absence d'évidence d'hétéroscédasticité significative dans les résidus, ce qui est positif pour la validité des estimations standard des erreurs.
3. Forme Fonctionnelle : Le test RESET a une p-value de 0.557, suggérant que la forme fonctionnelle du modèle est appropriée .
4. Multicollinéarité : Les valeurs VIF sont toutes inférieures à 5, ce qui indique qu'il n'y a pas de problèmes de multicollinéarité sévère parmi les variables explicatives.

Autocorrélation des Résidus : Donne une p-value extrêmement faible (< 2.2e-16), indiquant une autocorrélation significative des résidus.

L'autocorrélation persistante suggère que le modèle pourrait manquer de variables explicatives temporelles ou autres formes de dépendance non capturées. Les tentatives de régression des variables pour éliminer les tendances et l'ajout d'une variable temporelle n'ont pas résolu le problème d'autocorrélation des résidus.

Pour résoudre ce problème, nous pouvons envisager l'utilisation de modèles avec erreurs autorégressives (AR). Ces modèles sont des variantes de la régression linéaire ordinaire qui incorporent un processus autorégressif dans les termes d'erreur, permettant ainsi de modéliser directement l'autocorrélation dans les erreurs.

6.4.5 Modèle V de Moindres carrés généralisés et AR(1)

La fonction gls (Generalized Least Squares) du package nlme permet de réaliser des ajustements de modèles linéaires en utilisant la méthode des moindres carrés généralisés. Cette méthode permet de tenir compte de la corrélation entre les erreurs et/ou de l'hétéroscédasticité.

Le résultat cependant est le même l'autocorrélation est toujours présente (Annexe 13).

6.4.6 Modèle VI – MCO avec « lag »

Nous avons aussi testé l'introduction de lags dans notre modèle pour intégrer des dépendances temporelles, une technique couramment employée pour modéliser la manière dont les valeurs passées influencent les futures.

Ce procédé consiste à introduire des versions décalées de la série originale pour capturer les effets retardés. Cependant, malgré l'apparente précision du modèle, nous avons rencontré un message d'avertissement "essentially perfect fit: summary may be unreliable" lors de notre estimation. Ce message suggère que notre modèle pourrait être surajusté, capturant excessivement le bruit des données au lieu de leurs véritables dynamiques, ce qui met en doute la fiabilité des prédictions et des interprétations issues de ce modèle.

Conclusion

Tous les modèles estimés par la méthode des moindres carrés ordinaires présentent une autocorrélation des résidus, et leur coefficient de détermination R^2 révèle que la capacité explicative des modèles est limitée, comme le montre le Tableau 10). Bien que l'introduction de décalages temporels (lags) ait permis de réduire l'autocorrélation, ces

modèles présentent un problème de surajustement, rendant leur pertinence discutable pour des prédictions fiables. Dans ce contexte, nous allons opter pour une méthode mieux adaptée aux séries temporelles présentant de l'autocorrélation des résidus. L'emploi de modèles tels que ARIMA ou ARMAX est conseillé pour gérer efficacement l'autocorrélation et pour améliorer la précision des estimations.

Les sorties R permettant de vérifier les résultats du Tableau 10 sont présentées de l'Annexe 8 à l'Annexe 13.

Tableau 10 Résultat modèles MCO

Modèle	R²	Normalité	Hétéro	Reset	Autocorr
I	0.169	0.001	8.844e-05	0.001	2.2e-16
II (semi-log)	0.104	0.090	0.446	0.804	4.237e-14
III (Indice temporel)	0.126	0.169		0.508	4.177e-13
			0.485		
IV (correction tendance)	0.117	0.002	0.110	0.557	2.2e-16
N° V MCG					2.2e-16
N° VI (lag)	0.3945	0.743	0.006	0.004	0.161

6.5 Estimations et validation des modèles ARMAX

Nous avons estimé nos modèles en utilisant la fonction auto.arima du package forecast, en appliquant la même approche que celle utilisée précédemment. Après l'estimation, nous procédons aux tests nécessaires pour valider le modèle.

Pour commencer, nous avons testé nos séries temporelles sans corrections, c'est-à-dire dans un état non stationnaire. Le résultat a été un modèle ARIMA(0,0,2)(0,0,2)[12] avec un AIC de 3095.47. Ce modèle présentait des problèmes d'autocorrélation, d'homoscédasticité et de normalité.

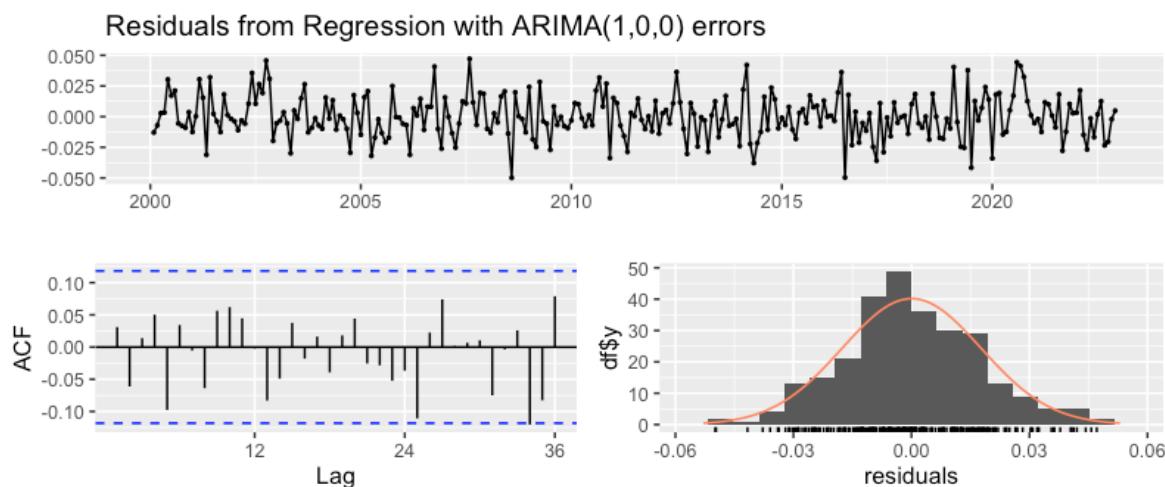
Pour notre deuxième modèle, nous avons utilisé toutes les variables stationnaires, et le résultat a été un ARIMA(1,0,0) avec un AIC de 2313.35. Et bonne nouvelle, l'autocorrélation n'est plus un problème, comme le montre une p-value de 0.178 au test de Ljung-Box. Nous avons ensuite estimé un nouveau modèle avec les variables

sélectionnées pour obtenir un modèle plus parcimonieux. Bien que l'autocorrélation ait été correctement traitée, nous continuons à rencontrer des problèmes d'homoscédasticité et de normalité.

Nous avons donc repris le modèle semi-logarithmique (Annexe 14) que nous avions utilisé précédemment, résultant à nouveau en un modèle ARIMA(1,0,0). Cette fois, l'AIC a diminué à -1426.15. Les résultats des tests, avec une p-value de 0.897 au test de Ljung-Box et de 0.118 au test de Breusch-Pagan, ainsi qu'une p-value de 0.28 au test de Kolmogorov-Smirnov, confirment que nos résidus sont effectivement du bruit blanc, comme nous pouvons l'observer dans

Figure 17.

Figure 17 Résidus modèle ARMAX



6.6 Interprétation du modèle sélectionné

Nous disposons donc d'un modèle qui satisfait l'ensemble des hypothèses tout en minimisant l'AIC. Nous allons maintenant procéder à la vérification de l'ajustement du modèle. Celui-ci repose sur une régression avec des erreurs ARIMA(1,0,0), intégrant ainsi une composante autorégressive d'ordre 1. Les coefficients de régression pour nos variables explicatives, ainsi que leur significativité, permettent d'identifier les variables qui ont un impact significatif sur l'IPA (Tableau 11).

Tableau 11 Coefficients de régression

Variables	Pr(> z)	Significativité
Ar 1 (x10 ⁶)	2,200	***
Intercept	0.174	
Exportations	0.222	
Pétrole	0.007	**
Précipitation	0.228	
Taux change	0.006	**

Le modèle révèle une forte autocorrélation des erreurs avec un décalage temporel, suggérant que les valeurs futures de notre série temporelle sont significativement influencées par les valeurs immédiatement antérieures, avec un niveau de significativité de 1%. Par ailleurs, le prix du pétrole est statistiquement significatif au seuil de risque de 5%. Le coefficient positif indique que les variations des prix du pétrole ont un impact direct et significatif sur l'indice des prix à la production agricole : lorsque les prix du pétrole augmentent, l'indice augmente également.

Concernant les autres variables du modèle, les exportations ne montrent pas d'impact significatif , avec une p value de 0.222, suggérant qu'elles n'influencent pas notablement la variable dépendante.

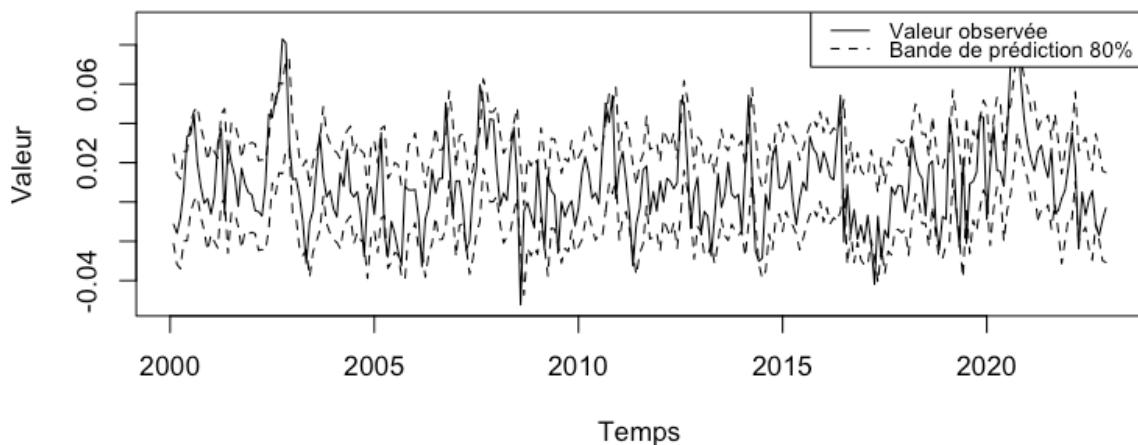
Le taux de change, en revanche, est statistiquement significatif au seuil de risque de 5%, indiquant que ses variations ont un effet positif et notable sur l'indice des prix à la production agricole. Concrètement, cela signifie que lorsque le real brésilien se déprécie, augmentant ainsi le taux de change, les prix à la production agricole tendent à augmenter.

Quant à notre variable climatique, la précipitation, elle s'est révélée non significative (p = 0.228), indiquant que , dans le cadre de cette étude, les variations des précipitations n'ont pas un impact notable sur l'indice des prix à la production agricole de manière statistiquement mesurable

Pour évaluer l'efficacité du modèle ARMAX, il est moins courant de se référer au R^2 , car celui-ci ne reflète pas la complexité temporelle des séries temporelles. Des critères comme l'AIC ou le BIC sont généralement préférés pour évaluer et sélectionner des modèles dans ce contexte. De plus, l'efficacité des prédictions peut être évaluée tant quantitativement que visuellement, en comparant les valeurs prédites par le modèle aux observations réelles. Cette comparaison est illustrée dans la Figure 18, qui montre les valeurs observées aux valeurs ajustées.

Sur le graphique, les valeurs observées correspondent à la série temporelle de notre variable IPA, tandis que les valeurs ajustées sont les prédictions faites par le modèle, basées sur les variables exogènes. Les bandes de prédition qui entourent ces prédictions reflètent l'incertitude associée à ces dernières, illustrant ainsi la marge d'erreur potentielle.

Figure 18 Qualité d'ajustement ARMAX



Le résultat indique que le modèle explique 80,3% de la variance, surpassant légèrement le seuil communément accepté de 80%. Selon (Aragon, 2011)(pag. 160), un score en dessous de 80% pourrait indiquer un ajustement insuffisant, tandis qu'un score trop élevé pourrait être le signe d'un surajustement. Cette analyse nous permet de confirmer la précision de notre modèle et d'évaluer sa fiabilité et ses limites avec plus de précision.

Par ailleurs nous avons également ajusté les résidus de notre modèle II - MCO semi-logarithmique en utilisant un modèle AR(1). Ce processus a permis de corriger l'autocorrélation identifiée précédemment dans les résidus. Ensuite, nous avons intégré

les valeurs ajustées de l'AR(1) dans un nouveau modèle de régression linéaire, ce qui a permis de modéliser plus précisément l'IPA, en tenant compte de cette autocorrélation. Cette approche a révélé que la variable de précipitation était statistiquement significative au seuil de 5% (Annexe 15). Bien que les résidus ajustés ne présentent plus de problèmes d'autocorrélation et semblent suivre une distribution normale, ils ont introduit à nouveau un problème d'hétérosécédasticité et de spécificité. En conséquence, nous avons décidé de conserver les interprétations fournies par le modèle présenté précédemment.

7 Conclusion

L'analyse des résultats du modèle ARMAX révèle que les variables économiques telles que le pétrole et le taux de change exercent des effets statistiquement significatifs. En revanche, les variables telles que les exportations et les variables climatiques, notamment la précipitation, ne présentent pas d'effets significatifs dans ce modèle. Le coefficient autorégressif (ar1) indique une forte dépendance temporelle, démontrant que les valeurs actuelles sont fortement influencées par leurs valeurs antérieures.

La littérature existante a déjà établi un lien entre le pétrole, le taux de change et les fluctuations de l'indice des prix à la production agricole, ce qui rend les résultats de notre étude peu surprenants. En effet, un coefficient positif dans notre modèle indique que, à mesure que la valeur d'une variable exogène comme le pétrole ou le taux de change augmente, la valeur prévue de l'indice des prix à la production agricole tend également à augmenter. Cependant, même si nos analyses n'ont révélé aucun impact statistiquement significatif des variables climatiques — ce qui peut nous sembler contre-intuitif — des recherches sur la modélisation des rendements agricoles pourraient nous fournir des explications sur cette absence de résultats significatifs. Nous explorerons deux hypothèses principales pour mieux comprendre et interpréter nos résultats.

L'article "The Impact of Climate Change on Brazil's Agriculture" (Marcia Zilli, 2020) décrit l'agriculture brésilienne comme extrêmement hétérogène et localisée, soulignant que les effets du changement climatique peuvent varier grandement d'une région à l'autre. Par exemple, tandis que certaines régions peuvent subir une baisse de la production agricole

en raison de conditions climatiques défavorables, d'autres peuvent connaître une hausse ou un déplacement de cette production. De telles dynamiques sont illustrées par le déplacement anticipé de la culture du soja du biome du Cerrado vers le nord de Matopiba, s'étendant sur des pâturages existants. Le maïs, deuxième culture en importance, pourrait également voir sa production déplacée des régions tropicales vers les sous-tropiques en réponse aux changements climatiques.

Ces mouvements montrent comment l'agriculture brésilienne s'adapte et se repositionne face aux problématiques climatiques. Cependant, lorsque ces impacts sont analysés à un niveau national, les variations régionales spécifiques peuvent être masquées. L'agrégation des données au niveau national tend à diluer les effets observables des variables climatiques. Ainsi, si une région compense la diminution de production d'une autre, l'impact global sur les prix nationaux peut apparaître négligeable.

Sous un autre angle, l'impact modérateur des avancées technologiques¹ dans l'agriculture est également à considérer. Bien que ces progrès aient le potentiel d'améliorer significativement les rendements et la gestion des cultures, ils introduisent également une complexité dans l'évaluation des effets directs du climat sur l'agriculture.

Les innovations en génétique des plantes et en gestion agronomique ont permis d'augmenter les rendements dans des zones susceptibles aux stress climatiques, comme le Mato Grosso et le Pará (Anna C. Hampf, 2020). Par exemple, L'introduction de variétés résistantes à la sécheresse, couplée à une optimisation de l'utilisation des intrants tels que les engrains et les pesticides, a significativement renforcé la résilience des cultures face aux aléas climatiques. Ces progrès techniques permettent aux plantations de mieux supporter les conditions défavorables et de maintenir des niveaux de production élevés, ce qui pourrait occulter les impacts négatifs attendus des changements climatiques.

Parallèlement, l'adoption de systèmes de double culture, notamment la production de maïs suivant celle du soja, a optimisé l'utilisation des terres, augmentant ainsi la production globale. Toutefois, cette intensification pose des défis environnementaux, tels

¹ « Toutes les mesures liées à la gestion des cultures tel que l'amélioration des machines, des pesticides et des herbicides, etc., et aux connaissances agronomiques des agriculteurs », (F. Ewert, 2005)

que l'épuisement des sols et une dépendance croissante aux produits chimiques, soulevant des questions quant à la durabilité de ces pratiques à long terme.

En conséquence, malgré les bénéfices apportés par les avancées technologiques pour faire face aux défis du changement climatique, elles peuvent aussi brouiller la compréhension des relations entre le climat et les performances agricoles, rendant les variables climatiques moins pertinentes dans les modèles économétriques.

8 Discussion

Dans cette dernière partie, nous aborderons les difficultés et les limites rencontrées, ainsi que les idées et impressions tirées de notre expérience.

L'accès aux données constitue un défi majeur. De plus, la sélection des variables à intégrer dans le modèle et l'évaluation de leur impact sur les estimations statistiques exigent une réflexion approfondie, surtout dans des domaines où notre expertise est limitée. Par exemple, nous avons opté pour l'inclusion du salaire minimum en vigueur plutôt que du salaire nominal afin de contourner les problèmes d'endogénéité. Concernant les variables climatiques, nous avons considéré la pertinence de l'humidité, mais l'absence de références bibliographiques et le risque de multicollinearité avec d'autres variables climatiques nous ont contraints à l'exclure de notre modèle, une décision similaire ayant été appliquée à d'autres variables.

La correction des valeurs atypiques, bien que souvent nécessaire pour éviter des biais significatifs dans les résultats finaux, peut en effet altérer les informations initiales. Cela peut être problématique, surtout dans des contextes où chaque donnée est importante pour comprendre le phénomène étudié. La décision de corriger ou non ces valeurs doit donc être guidée par une compréhension approfondie du contexte spécifique et des implications potentielles de cette modification sur les résultats de l'étude.

En ce qui concerne la modélisation, bien que la littérature nous ait souvent guidés vers l'utilisation de modèles plus adaptés et complexes pour aborder notre problématique, nous avons finalement opté pour les méthodes étudiées durant notre master.

En pratique, certaines méthodes peuvent sembler confuses et manquent parfois de réponses directes. De plus, il existe un manque de documentation liés aux applications précises, sur des problèmes tels que la stationnarité ou la saisonnalité, ainsi que sur la manière de les traiter. Certaines recherches omettent également d'inclure des tests que nous avons appris à considérer comme essentiels pour l'interprétation des données.

Nos résultats montrent que les modèles ARMA sans composante AR ou MA (spécifiés par `order = c(0,0,0)`) donnent des résultats similaires à ceux obtenus avec la méthode des moindres carrés. Cette similitude s'explique par le fait que ces deux modèles sont essentiellement des régressions linéaires sans composante autorégressive ou de moyenne mobile. En ajoutant une composante AR(1) à ces modèles, nous pouvons capturer la dépendance temporelle présente dans les résidus, améliorant ainsi l'ajustement du modèle. Cependant, l'introduction d'une composante AR dans notre modèle n'a pas produit les résultats attendus. Par ailleurs, nous avons également testé la méthode des moindres carrés avec toutes les variables différencierées (les résultats peuvent être fournis si nécessaire). Nous avons cependant décidé de ne pas les inclure car ils présentent les mêmes problématiques et n'apportent pas de changements significatifs au dossier. Les résultats pour le modèle ARMAX étaient globalement les mêmes ; les variables significatives sont le prix du pétrole et le taux de change.

Analyser les effets climatiques d'un pays sans une connaissance approfondie de celui-ci a été une expérience enrichissante mais complexe. Cette approche a révélé qu'une analyse plus régionalisée pourrait être plus efficace pour saisir les nuances locales. En segmentant les données par région, nous avons l'opportunité de mieux comprendre les variations régionales et d'ajuster nos modèles de manière plus précise. Dans ce contexte, l'utilisation de modèles en panel pourrait être pertinente, car ils sont bien adaptés pour analyser des données collectées à travers différentes régions et périodes. D'autres modèles, tels que les modèles VAR ou de correction d'erreurs, pourraient également enrichir l'étude en explorant les interactions dynamiques entre les variables au fil du temps.

D'un autre côté, progresser dans un projet tel que ce mémoire sans obtenir les résultats statistiques significatifs attendus peut s'avérer frustrant. Toutefois, cette frustration est atténuée lorsque l'on prend du recul pour adopter une perspective plus globale et lorsqu'on analyse les enjeux et la problématique étudiée.

Cette vision élargie nous amène à explorer les conséquences plus larges de notre étude, en particulier sur l'expansion de l'agriculture brésilienne. Motivée par l'ouverture de nouveaux marchés en Asie pour des produits tels que le riz, ainsi que les produits laitiers et la viande de porc, cette expansion suscite des interrogations quant à son coût réel (Céline Broggio, 2021). L'Europe, en tant le premier investisseur étranger (Leyen, 2023) et une de plus grande destination des exportations brésiliennes, joue un rôle très important.

Nous considérons que la pérennité et l'engagement financier envers des programmes tels que REDD+ énoncé dans « Durabilité des Systèmes », p. 202 (A. Thomas, 2024) devraient être au cœur de nos priorités. Dans l'article, les recherches indiquent que, même si les systèmes agricoles diversifiés introduits par ces programmes peuvent initialement modifier les comportements des agriculteurs, beaucoup tendent à reprendre leurs anciennes méthodes une fois ces programmes terminés. Cette régression est souvent due au manque de soutien continu pour ancrer ces changements de manière durable. Il est également important de souligner que les participants à ces programmes ont réalisé des avancées environnementales significatives par rapport à leurs homologues non-participants, prouvant ainsi que des alternatives viables à l'agriculture traditionnelle existent et peuvent contribuer à la protection de nos forêts.

De plus, il est vital que la durabilité des pratiques agricoles soit en harmonie avec la sécurité alimentaire, et non perçue comme un obstacle. Cette perception a malheureusement contribué à l'interruption prématurée du programme en question. Finalement, pour éviter de tels échecs, nous devons envisager des stratégies de financement et d'engagement à long terme qui tiennent compte de toutes ces dimensions.

9 Bibliographie

- A. Kapusuzoglu, M. K. (2015). *The interactions between agricultural commodity and oil prices: an empirical analysis*. Récupéré sur <http://agricecon.agriculturejournals.cz/pdfs/age/2015/09/03.pdf>
- A. Saucedo, B. B. (2015, 01). *The dynamic pattern of volatility spillovers between oil and agricultural markets*. Récupéré sur Department of Agricultural Economics and Rural Development University of Göttingen: www.fp7ulysses.eu
- A. Thomas, A. A.-R. (2024). *DURABILITÉ DES SYSTÈMES POUR LA SÉCURITÉ ALIMENTAIRE Combiner les approches locales et globales*. Versailles: Éditions Quæ.
- Anna C. Hampf, T. S.-M. (2020). *Future yields of double-cropping systems in the Southern Amazon, Brazil, under climate change and technological development*. Récupéré sur Agricultural Systems, Volume 177.
- Aragon, Y. (2011). *Séries temporelles avec R Méthodes et cas*. Springer-Verlag France.
- C. Valdes, K. H. (2020, 10). *Brazil's Currency Depreciation and Changing Macroeconomic Conditions Determine Agricultural Competitiveness and Future Growth*. Récupéré sur Economic Research Service United States Department of Agriculture: AgEcon Search, pag 4
- Carfantan, J. (2021, 09). *La puissance agricole du Brésil est menacée*. Récupéré sur IstoéBrésil: <https://www.istoebresil.org/post/la-puissance-agricole-du-brésil-est-menacée#:~:text=Raréfaction%20des%20pluies%20dans%20le,et%20détérioration%20de%20la%20qualité>
- Céline Broggio, M. D. (2021). *L'agronégoce, enjeu géopolitique au Brésil*. Consulté le 04 2024, sur https://www.cairn.info/revue-herodote-2021-2-page-91.htm?casa_token=glT4LojEk0EAAAAA:iEyfQvP36gmD5KJw57DVlTugDob-5Gnvfl7WlHKQDA2GSKze6B1GIci-m4GLvNizCSEdwj4Opa0wj8zsSA
- Centre d'Information Inter Peuples (CIIP). (2016, 05). Récupéré sur Brésil : l'affrontement de deux modèles de société: <https://www.ritimo.org/Bresil-L-affrontement-de-deux-modeles-de-developpement-agricole>
- Darné, O. (2024, 02). Désaisonnalisation et Décomposition des Séries Temporelles. *Cours Techniques de prévision et conjoncture - Chap 1, pag 62*.

Daviron, B. (2012, 3). *Prix internationaux des produits alimentaires :volatilité ou hausse durable ?*

Consulté le 2024, sur <https://www.cairn.info/revue-tiers-monde-2012-3-page-91.htm#re4no71>

E. Sabourin, C. G.-F. (2020). *Le démantèlement des politiques publiques rurales et environnementales au Brésil.* Consulté le 2024, sur www.cahiersagricultures.fr:https://agritrop.cirad.fr/596894/1/Demanてlement%20PP%20au%20Brésil%20Sabourin%20et%20al%20C%20agri%202020.pdf

F. Antunes de Araujo, L. B. (2020). *An analysis of Brazilian agricultural commodities using permutation – information theory quantifiers: The influence of food crisis.* Récupéré sur Chaos, Solitons & Fractals, Volume 139: <https://www.sciencedirect.com/science/article/pii/S0960077920304781>

F. Ceballos, M. H. (2016). *Transmission of Food Price Volatility from International to Domestic Markets: Evidence from Africa, Latin America, and South Asia.* Récupéré sur Food Price Volatility and Its Implications for Food Security and Policy, 2016: https://link.springer.com/chapter/10.1007/978-3-319-28201-5_13?utm_source=geftr&utm_medium=geftr&utm_campaign=geftr_pilot

F. Ewert, M. R. (2005). *Future scenarios of European agricultural land use: I. Estimating changes in crop productivity.* Récupéré sur Agriculture, Ecosystems & Environment: <https://www.sciencedirect.com/science/article/pii/S0167880904003627>

France, W. (2020). *Feux en Amazonie.* Récupéré sur World Wide Fund For Nature (WWF): <https://www.wwf.fr/sauver-amazonie>

Fundação Getulio Vargas, Conjuntura Econômica. (2024, 04). *Índice de preços ao produtor amplo - disponibilidade interna (IPA-DI) - origem - produtos agropecuários: índice (ago. 1994 = 100).* Consulté le 2024, sur IPEA DATA: <http://www.ipeadata.gov.br/Default.aspx>

G. El Montasserb, M. M. (2023, 03). *Co-explosivity versus leading effects: Evidence from crude oil and agricultural commodities.* Consulté le 02 2024, sur [https://pdf.sciencedirectassets.com/271807/1-s2.0-S0301420723X00026/1-s2.0-S0301420723000399/main.pdf?X-Amz-Security-Token=IQoJb3JpZ2luX2VjEN7%2F%2F%2F%2F%2F%2F%2F%2F%2F%2F%2F%2F%2FwEaCXVzLWVhc3QtMSJHMEUCIHv877sp5xo%2FnXIE2vEG0jrPgammLMHtPR%2B%2F%2BaHpU4uAiEA%2B6](https://pdf.sciencedirectassets.com/271807/1-s2.0-S0301420723X00026/1-s2.0-S0301420723000399/main.pdf?X-Amz-Security-Token=IQoJb3JpZ2luX2VjEN7%2F%2F%2F%2F%2F%2F%2F%2F%2F%2F%2FwEaCXVzLWVhc3QtMSJHMEUCIHv877sp5xo%2FnXIE2vEG0jrPgammLMHtPR%2B%2F%2BaHpU4uAiEA%2B6)

Gimet, C. (2009). *L'impact des chocs externes dans les économies du Mercosur: un modèle VAR Structurel.* Récupéré sur HAL open science: <https://shs.hal.science/halshs-00356105>, pag 8

- Girardi, E. P. (2022). *Le Brésil, puissance agricole : dynamiques récentes, projections, contradictions et fragilités (2006-2029)*. Récupéré sur Confins revue franco-brésilienne de géographie: <https://journals.openedition.org/confins/47090>
- Granger, C. W. (1978). *National Bureau of Economic Research*. Récupéré sur Chapter Title: Seasonality: Causation, Interpretation, and Implications:
<https://www.nber.org/system/files/chapters/c4321/c4321.pdf>
- Grubbs, F. E. (1969). *Procedures for Detecting Outlying Observations in Samples*. . Technométrie, 11, 1–21.
- J. Baleyté, J. M. (2023, juin 14). *L'inflation reflue, la croissance hésite*. Consulté le 2024, sur Notes de conjoncture: www.insee.fr.
- J. Burney, C. M. (2024, 04). *Empirical modeling of agricultural climate risk*. Récupéré sur The University of British Columbia Institute for Resources Environment and Sustainability.
- L.F.A. Curado, S. d. (2023, 08). *Trends and Patterns of Daily Maximum, Minimum and Mean Temperature in Brazil from 2000 to 2020*. Récupéré sur Climate:
<https://www.mdpi.com/2225-1154/11/8/168#B4-climate-11-00168>
- La banque mondiale . (2022, OCTOBRE 7). *Tout ce que vous devez savoir sur la sécurité alimentaire et le changement climatique*. Récupéré sur
<https://www.banquemonde.org/fr/news/feature/2022/10/17/what-you-need-to-know-about-food-security-and-climate-change>
- Le Monde avec AFP. (2021, 06). *Au Brésil, la sécheresse menace l'approvisionnement en électricité du pays*. Récupéré sur Le Monde :
https://www.lemonde.fr/international/article/2021/06/07/au-bresil-la-secheresse-menace-l approvisionnement-en-electricite-du-pays_6083206_3210.html
- Leyen, V. d. (2023, 06 12). Discours de la Présidente von der Leyen à la Confédération nationale de l'industrie.
- M.TRAVERS. (2023). Hypothèses sous-jacentes à la méthode MCO. *Cours d'économétrie linéaire avancée - Master ECAP*.
- Marcia Zilli, M. S. (2020). *The impact of climate change on Brazil's agriculture*. Récupéré sur Science of The Total Environment, Volume 740:
<https://www.sciencedirect.com/science/article/pii/S0048969720329016>
- OECD. (2023). *Brazil, support to agriculture. pag : contextual information*. Récupéré sur OECD iLibrary: <https://www.oecd-ilibrary.org/sites/5f1b8b6e-en/index.html?itemId=/content/component/5f1b8b6e-en>

O'Neill, J. (2008, 4). *Les BRIC, nouvelles grandes puissances dans le futur ?* Consulté le 2024, sur <https://www.cairn.info/revue-internationale-et-strategique-2008-4-page-231.htm?ref=doi>

Organisation des Nations Unies pour l'alimentation et l'agriculture. (2023). *Objectif 2 : Éliminer la faim, assurer la sécurité alimentaire, améliorer la nutrition et promouvoir l'agriculture durable.* Récupéré sur <https://www.un.org/sustainabledevelopment/fr/hunger/>

R. Endré Dahl, A. O. (2020). *Dynamics of volatility spillover in commodity markets: Linking crude oil to agriculture.* Récupéré sur Volume 20.

R. Lisboa Costa, G. M. (2020, 06). *Analysis of climate extremes indices over northeast Brazil from 1961 to 2014.* Consulté le 05 2024, sur Weather and Climate Extremes, vol 28: [https://pdf.sciencedirectassets.com/283425/1-s2.0-S2212094719X00079/1-s2.0-S2212094719300763/main.pdf?X-Amz-Security-Token=IQoJb3JpZ2luX2VjENr%2F%2F%2F%2F%2F%2F%2F%2F%2F%2F%2F%2FwEaCXVzLWVhc3QtMSJIMEYCIQDaqDOTTaX81syYqIZP1VQ9QZKIBgiBH8X%2BxkFJVSDr4QIhAOZzB0hHL3](https://pdf.sciencedirectassets.com/283425/1-s2.0-S2212094719X00079/1-s2.0-S2212094719300763/main.pdf?X-Amz-Security-Token=IQoJb3JpZ2luX2VjENr%2F%2F%2F%2F%2F%2F%2F%2F%2F%2FwEaCXVzLWVhc3QtMSJIMEYCIQDaqDOTTaX81syYqIZP1VQ9QZKIBgiBH8X%2BxkFJVSDr4QIhAOZzB0hHL3)

R. Ribeiro-Silva, M. P. (2020, 08). *Covid-19 pandemic implications for food and nutrition security in Brazil.* Récupéré sur Ciência & Saúde Coletiva, pag 3424 Chart 1. Effects of COVID-19: <https://www.scielo.br/j/csc/a/mFBrPHcbPdQCPdsJYN4ncLy/?lang=en&format=html>

Regmi A., D. M.-s. (2021). *Cross-Country Analysis of Food Consumption Patterns.* Récupéré sur https://www.ers.usda.gov/webdocs/outlooks/40303/14973_wrs011d_1_.pdf?v=529

Righetti, N. (2022, 12). *Time Series Analysis With R.* Récupéré sur Structural Decomposition: <https://nicolarighetti.github.io/Time-Series-Analysis-With-R/structural-decomposition.html#white-noise-and-stationarity>

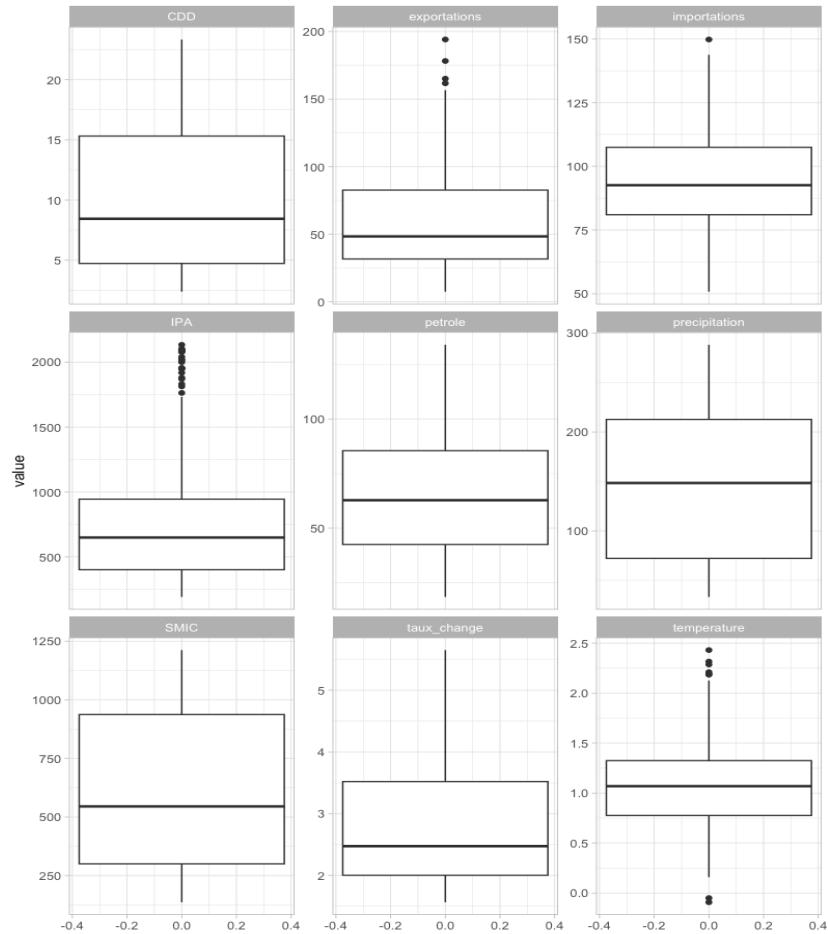
S. Raya, A. L. (2023). *An ARIMA-LSTM model for predicting volatile agricultural price series with random forest technique.* Récupéré sur Volume 149, Part A: (<https://www.sciencedirect.com/science/article/pii/S1568494623009572>)

Service économique régional de Brasilia. (2023, 07). *BRÈVES ÉCONOMIQUES DU BRÉSIL.* Consulté le 05 2024, sur Ministère de l'économie des finances et de la souveraineté industrielle et numérique: <https://www.tresor.economie.gouv.fr/Articles/a9e4445be337-4bac-ad27-1d32cba08c98/files/4d7741e0-be07-4cca-a28a-77d226c5320a>

Variation de température sur la superficie des terres. (s.d.). Consulté le 2024, sur Organisation des Nations unies pour l'alimentation et l'agriculture: <https://www.fao.org/faostat/fr/#data/ET>

10 Annexes

Annexe 1. Boxplot pour l'ensemble des variables



Annexe 2. Variables avec des valeurs atypiques – Outliers

```

Series: ts_exportations
Regression with ARIMA(1,0,0)(0,1,1)[12] errors

Coefficients:
      ar1     sma1     A033     A0114     LS185     A0254
    0.7625 -0.6238  25.6306  31.0681  26.0407 -27.9435
  s.e.  0.0426   0.0510   6.9851   6.8995   6.8729   7.1521

sigma^2 = 94.47: log likelihood = -975.31
AIC=1964.61  AICc=1965.05  BIC=1989.65

Outliers:
  type ind    time coefhat tstat
1  A0  33 2002:09   25.63  3.669
2  A0 114 2009:06   31.07  4.503
3  LS 185 2015:05   26.04  3.789
4  A0 254 2021:02 -27.94 -3.907

```

Series: ts_importations
 Regression with ARIMA(1,1,2)(1,0,0)[12] errors

 Coefficients:
 ar1 ma1 ma2 sar1 LS47 A098
 0.0432 -0.7165 -0.1532 0.2900 -29.4529 47.8804
 s.e. 0.3300 0.3263 0.2703 0.0597 7.4462 12.0018
 LS195 LS206
 40.0092 -30.3560
 s.e. 7.3329 7.3662

$\sigma^2 = 175.7$: log likelihood = -1097.94
 AIC=2213.88 AIc=2214.56 BIC=2246.43

Outliers:
 type ind time coefhat tstat
 1 LS 47 2003:11 -29.45 -3.955
 2 A0 98 2008:02 47.88 3.989
 3 LS 195 2016:03 40.01 5.456
 4 LS 206 2017:02 -30.36 -4.121

Series: ts_taux_change
 Regression with ARIMA(0,1,1) errors

Coefficients:
 ma1 A034
 0.3165 0.3361
 s.e. 0.0573 0.0697

$\sigma^2 = 0.0143$: log likelihood = 194.75
 AIC=-383.5 AIc=-383.41 BIC=-372.65

Outliers:
 type ind time coefhat tstat
 1 A0 34 2002:10 0.3361 4.824

Series: ts_precipitation
 Regression with ARIMA(0,0,1)(2,1,0)[12] errors

Coefficients:
 ma1 sar1 sar2 TC49 A0193
 0.2476 -0.6126 -0.3736 59.2558 86.5774
 s.e. 0.0571 0.0580 0.0589 14.5850 18.3329

$\sigma^2 = 424.8$: log likelihood = -1174.07
 AIC=2360.14 AIc=2360.46 BIC=2381.59

Outliers:
 type ind time coefhat tstat
 1 TC 49 2004:01 59.26 4.063
 2 A0 193 2016:01 86.58 4.723

```

Series: ts_SMIC
Regression with ARIMA(0,0,5)(0,0,2)[12] errors

Coefficients:
    ma1     ma2     ma3     ma4     ma5     sma1     sma2   intercept     LS16
    0.8636  0.6544  0.5548  0.3908  0.1828 -0.9308  0.3063  144.4274  44.9926
s.e.  0.0632  0.0802  0.0792  0.0747  0.0548  0.0625  0.0594  4.3551  5.7660
      LS40     LS65     LS88    A0121    LS122    LS145    LS157    LS169
    61.3801 71.3287 94.4402 76.0256 120.9066 88.0599 44.6969 77.6692
s.e.  4.2806  3.8698  3.8999  5.0530  4.4605  8.0303 10.7124  7.9620
      LS193    A0229    LS230    LS253    LS265
    177.1326 95.2608 104.2519 74.4118 107.2180
s.e.  4.3219  5.1179  4.4431  8.0232 11.0379

sigma^2 = 69.78: log likelihood = -972.05
AIC=1990.11  AICc=1994.49  BIC=2073.38

```

```

Outliers:
  type ind   time coefhat   tstat
1   LS  16 2001:04   44.99  7.803
2   LS  40 2003:04   61.38 14.339
3   LS  65 2005:05   71.33 18.432
4   LS  88 2007:04   94.44 24.216
5   A0 121 2010:01   76.03 15.046
6   LS 122 2010:02 120.91 27.106
7   LS 145 2012:01   88.06 10.966
8   LS 157 2013:01   44.70  4.172
9   LS 169 2014:01   77.67  9.755
10  LS 193 2016:01 177.13 40.985
11  A0 229 2019:01   95.26 18.613
12  LS 230 2019:02 104.25 23.464
13  LS 253 2021:01   74.41  9.275
14  LS 265 2022:01 107.22  9.714

```

- Les variables sans valeurs atypiques sont :

```

> fit_temperature
Series:
ARIMA(0,1,2)

Coefficients:
    ma1     ma2
    -0.6600 -0.1641
s.e.  0.0588  0.0651

sigma^2 = 0.1426: log likelihood = -121.87
AIC=249.74  AICc=249.83  BIC=260.59

No outliers were detected.

```

```

> fit_CDD
Series:
ARIMA(0,0,2)(2,1,2)[12]

Coefficients:
      ma1     ma2    sar1    sar2    sma1    sma2
      0.4183  0.2297 -0.8083 -0.0933 -0.0277 -0.6980
s.e.  0.0635  0.0596  0.1733  0.0830  0.1735  0.1586

sigma^2 = 2.206: log likelihood = -484.75
AIC=983.5   AICc=983.94   BIC=1008.53

```

No outliers were detected.

```

> fit_petrole
Series:
ARIMA(1,1,0)

```

```

Coefficients:
      ar1
      0.3497
s.e.  0.0567

sigma^2 = 32.51: log likelihood = -868.48
AIC=1740.95   AICc=1741   BIC=1748.19

```

No outliers were detected.

- Variable IPA, détection des trois valeurs atypiques

```

Series: ts_IPA-DI
Regression with ARIMA(1,1,0) errors

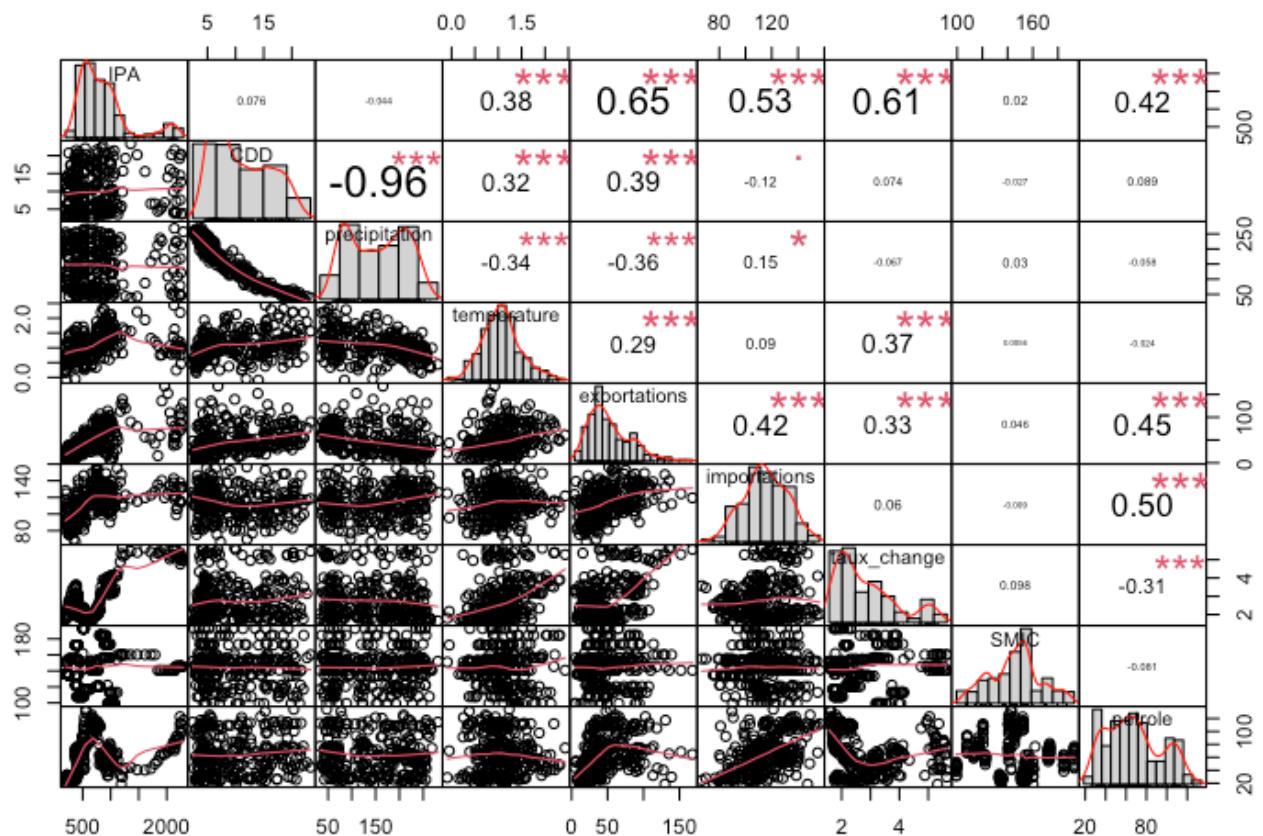
Coefficients:
      ar1      LS252      TC258      TC263
      0.6912  -149.6890  -77.9185  -42.7527
s.e.  0.0435   13.6418   11.6972   11.6982

sigma^2 = 275.1: log likelihood = -1160.87
AIC=2331.73   AICc=2331.96   BIC=2349.82

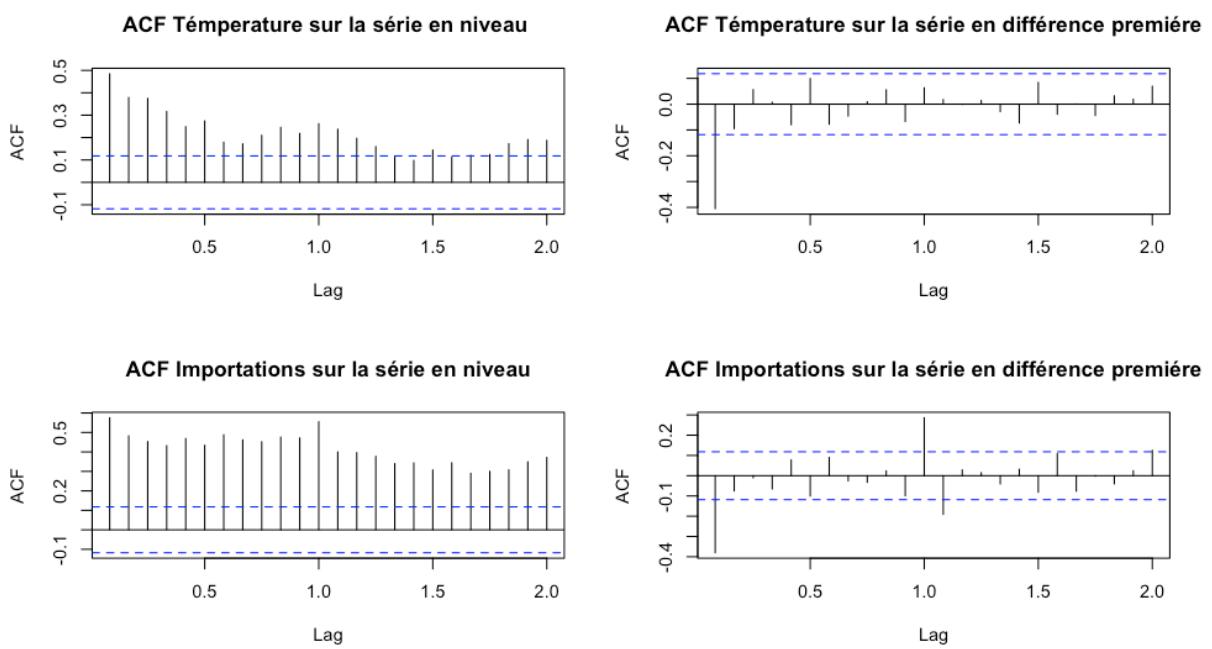
Outliers:
  type ind   time coefhat   tstat
1  LS 252 2020:12 -149.69 -10.973
2  TC 258 2021:06  -77.92  -6.661
3  TC 263 2021:11  -42.75  -3.655

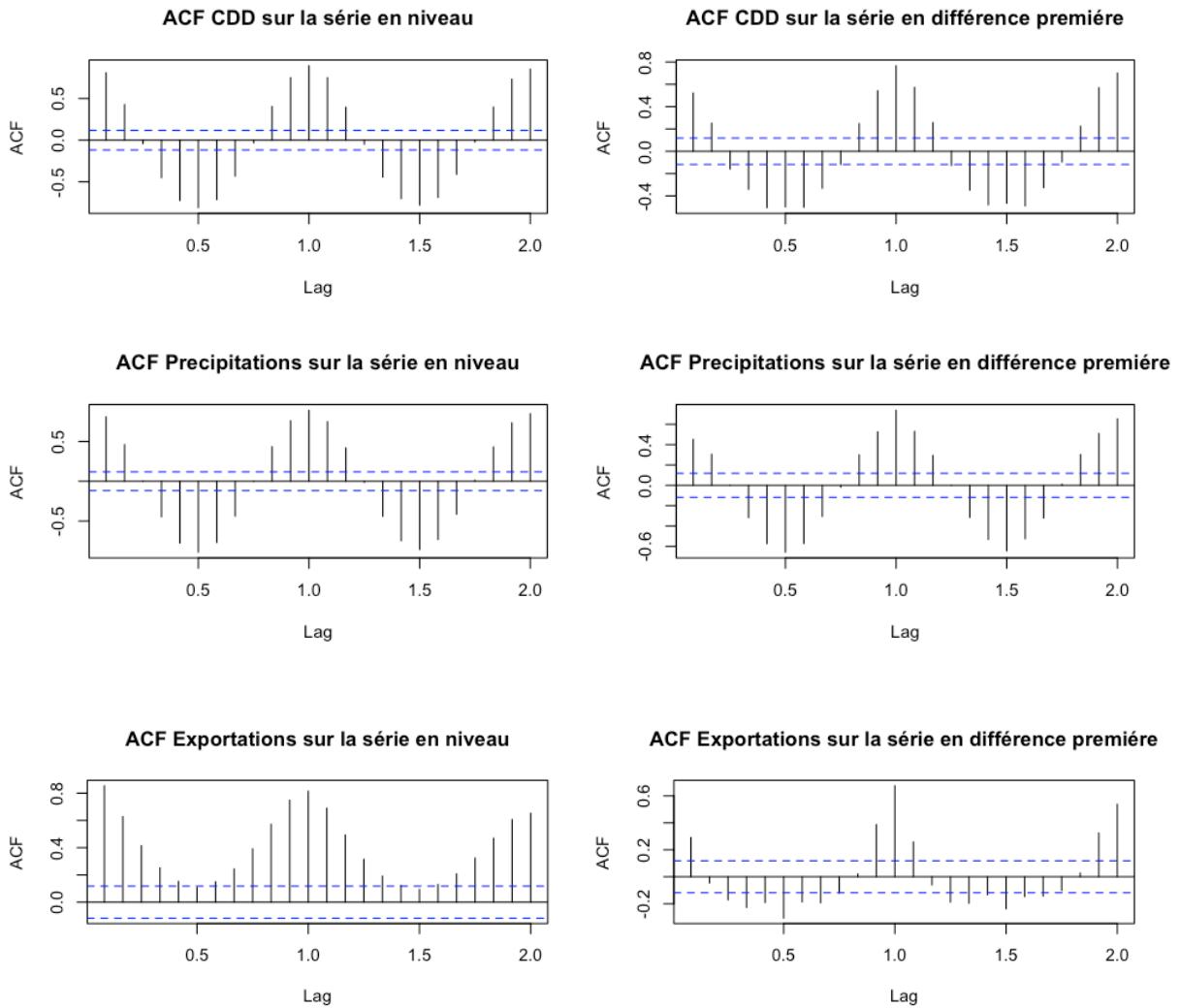
```

Annexe 3. Chart corrélation



Annexe 4. Corrélogrammes





Annexe 5. Test de tendance de Mann-Kendall

```
$CDD
```

```
Mann-Kendall trend test
```

```
data: x
z = 1.4414, n = 276, p-value = 0.07474
alternative hypothesis: true S is greater than 0
sample estimates:
      S      varS      tau
2.210000e+03 2.348661e+06 5.825141e-02
```



```
$IPA
```

```
Mann-Kendall trend test
```

```
data: x
z = 22.407, n = 276, p-value < 2.2e-16
alternative hypothesis: true S is greater than 0
sample estimates:
      S      varS      tau
3.434000e+04 2.348683e+06 9.048748e-01
```



```
$exportations
```

```
Mann-Kendall trend test
```

```
data: x
z = 12.289, n = 276, p-value < 2.2e-16
alternative hypothesis: true S is greater than 0
sample estimates:
      S      varS      tau
1.883500e+04 2.348658e+06 4.964745e-01
```



```
$importations
```

```
Mann-Kendall trend test
```

```
data: x
z = 9.3009, n = 276, p-value < 2.2e-16
alternative hypothesis: true S is greater than 0
sample estimates:
      S      varS      tau
1.425500e+04 2.348659e+06 3.757397e-01
```

```

$precipitation                               $petrole
                                              Mann-Kendall trend test

data: x
z = -0.87893, n = 276, p-value = 0.8103      data: x
alternative hypothesis: true S is greater than 0   z = 8.2693, n = 276, p-value < 2.2e-16
sample estimates:                           alternative hypothesis: true S is greater than 0
                                              S       varS       tau
-1.348000e+03  2.348683e+06 -3.552042e-02   sample estimates:
                                              S       varS       tau
$SMIC                                         1.267400e+04 2.348683e+06 3.339657e-01

$taux_change
                                              Mann-Kendall trend test

data: x
z = 10.163, n = 276, p-value < 2.2e-16
alternative hypothesis: true S is greater than 0
sample estimates:
                                              S       varS       tau
1.557600e+04 2.348681e+06 4.104456e-01

$temperature
                                              Mann-Kendall trend test

data: x
z = 6.3914, n = 276, p-value = 8.22e-11
alternative hypothesis: true S is greater than 0
sample estimates:
                                              S       varS       tau
9.796000e+03 2.348657e+06 2.582176e-01

```

Annexe 6. Test saisonnalité

Test Seasonal dummies: IPA :	Test Seasonal dummies: exportations :
Test used: SeasonalDummies	Test used: SeasonalDummies
Test statistic: 1.27	Test statistic: 23.43
P-value: 0.2414814	P-value: 0
Test Seasonal dummies: CDD :	Test Seasonal dummies: petrole :
Test used: SeasonalDummies	Test used: SeasonalDummies
Test statistic: 283.33	Test statistic: 1.06
P-value: 0	P-value: 0.390488
Test Seasonal dummies: importations :	Test Seasonal dummies: taux_change :
Test used: SeasonalDummies	Test used: SeasonalDummies
Test statistic: 6.31	Test statistic: 1.31
P-value: 2.993919e-09	P-value: 0.218587

```

Test Seasonal dummies: SMIC :
Test used: SeasonalDummies

Test statistic: 0.06
P-value: 0.999995

Test Seasonal dummies: precipitation :
Test used: SeasonalDummies

Test statistic: 351.69
P-value: 0

Test Seasonal dummies: temperature :
Test used: SeasonalDummies

Test statistic: 4.35
P-value: 5.416069e-06

```

Annexe 7. Test ADF et KPSS

KPSS Test for Trend Stationarity	Augmented Dickey-Fuller Test
<pre> data: ipa_diff KPSS Trend = 0.12657, Truncation lag parameter = 5, p-value = 0.08599 </pre>	<pre> data: ipa_diff Dickey-Fuller = -3.9389, Lag order = 6, p-value = 0.01251 alternative hypothesis: stationary </pre>
KPSS Test for Trend Stationarity	Augmented Dickey-Fuller Test
<pre> data: petrole_diff KPSS Trend = 0.038075, Truncation lag parameter = 5, p-value = 0.1 </pre>	<pre> data: petrole_diff Dickey-Fuller = -6.3498, Lag order = 6, p-value = 0.01 alternative hypothesis: stationary </pre>
KPSS Test for Trend Stationarity	Augmented Dickey-Fuller Test
<pre> data: ts_SMIC KPSS Trend = 0.018996, Truncation lag parameter = 5, p-value = 0.1 </pre>	<pre> data: ts_SMIC Dickey-Fuller = -6.0839, Lag order = 6, p-value = 0.01 alternative hypothesis: stationary </pre>
KPSS Test for Trend Stationarity	Augmented Dickey-Fuller Test
<pre> data: CDD_deseason KPSS Trend = 0.053447, Truncation lag parameter = 5, p-value = 0.1 </pre>	<pre> data: CDD_deseason Dickey-Fuller = -6.1831, Lag order = 6, p-value = 0.01 alternative hypothesis: stationary </pre>
KPSS Test for Trend Stationarity	Augmented Dickey-Fuller Test
<pre> data: importation_diff KPSS Trend = 0.014578, Truncation lag parameter = 5, p-value = 0.1 </pre>	<pre> data: importation_diff Dickey-Fuller = -10.305, Lag order = 6, p-value = 0.01 alternative hypothesis: stationary </pre>

<p>KPSS Test for Trend Stationarity</p> <pre>data: precipitations_deseason KPSS Trend = 0.078174, Truncation lag parameter = 5, p-value = 0.1</pre>	<p>Augmented Dickey-Fuller Test</p> <pre>data: precipitations_deseason Dickey-Fuller = -5.8362, Lag order = 6, p-value = 0.01 alternative hypothesis: stationary</pre>
<p>KPSS Test for Trend Stationarity</p> <pre>data: temperature_deseason KPSS Trend = 0.14896, Truncation lag parameter = 5, p-value = 0.04753</pre>	<p>Augmented Dickey-Fuller Test</p> <pre>data: temperature_deseason Dickey-Fuller = -3.9825, Lag order = 6, p-value = 0.01032 alternative hypothesis: stationary</pre>
<p>KPSS Test for Trend Stationarity</p> <pre>data: exportations_deseason KPSS Trend = 0.067186, Truncation lag parameter = 5, p-value = 0.1</pre>	<p>Augmented Dickey-Fuller Test</p> <pre>data: exportations_deseason Dickey-Fuller = -6.2104, Lag order = 6, p-value = 0.01 alternative hypothesis: stationary</pre>
<p>\$taux_change</p> <p>KPSS Test for Trend Stationarity</p> <pre>data: x KPSS Trend = 0.070678, Truncation lag parameter = 5, p-value = 0.1</pre>	<p>\$taux_change</p> <p>Augmented Dickey-Fuller Test</p> <pre>data: x Dickey-Fuller = -6.4556, Lag order = 6, p-value = 0.01 alternative hypothesis: stationary</pre>

Annexe 8. Modèle I

Call:

```
lm(formula = IPA ~ exportations + CDD + petrole + taux_change +
precipitation, data = stationnaire_data)
```

Residuals:

Min	1Q	Median	3Q	Max
-62.049	-9.874	-1.524	7.272	114.365

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.78407	18.61192	0.150	0.881204
exportations	0.22733	0.05030	4.520	9.29e-06 ***
CDD	1.57990	0.93540	1.689	0.092376 .
petrole	0.77049	0.20997	3.670	0.000293 ***
taux_change	20.96618	10.19045	2.057	0.040608 *
precipitation	-0.16106	0.08034	-2.005	0.045992 *

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 20.01 on 269 degrees of freedom
Multiple R-squared: 0.1691, Adjusted R-squared: 0.1537
F-statistic: 10.95 on 5 and 269 DF, p-value: 1.316e-09

```

> # Test de normalité des résidus
> ks.test(residus, "pnorm", mean = mean(residus), sd = sd(residus))

Asymptotic one-sample Kolmogorov-Smirnov test

data: residus
D = 0.11327, p-value = 0.001723
alternative hypothesis: two-sided

>
> # Test de Breusch-Pagan pour l'hétéroscléasticité
> bptest(lm_model1)

studentized Breusch-Pagan test

data: lm_model1
BP = 26.02, df = 5, p-value = 8.844e-05

> # Test RESET
> reset(lm_model1)

RESET test

data: lm_model1
RESET = 6.6669, df1 = 2, df2 = 267, p-value = 0.001495

> vif(lm_model1)
  exportations          CDD          petrole    taux_change precipitation
     1.048992        1.531173       1.115762      1.107182       1.482521
> checkresiduals(lm_model1)

Breusch-Godfrey test for serial correlation of order up to 10

data: Residuals
LM test = 109.93, df = 10, p-value < 2.2e-16

```

Annexe 9. Modèle II différencié-logarithmique

```

Call:
lm(formula = IPA ~ taux_change + petrole + precipitation + exportations,
  data = st2_data)

Residuals:
    Min      1Q   Median      3Q      Max 
-0.053072 -0.013313 -0.002177  0.009744  0.072178 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 3.342e-02 1.095e-02  3.053 0.002494 ** 
taux_change 4.133e-02 1.092e-02  3.786 0.000189 *** 
petrole      8.381e-04 2.247e-04  3.730 0.000234 *** 
precipitation -2.054e-04 7.110e-05 -2.889 0.004173 ** 
exportations 9.054e-05 5.277e-05  1.716 0.087389 .  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02144 on 270 degrees of freedom
Multiple R-squared:  0.1048,    Adjusted R-squared:  0.09151 
F-statistic: 7.9 on 4 and 270 DF,  p-value: 4.911e-06

```

```

> ## TESTS
> residus2 <- residuals(lm_model2)
> # Test de normalité des résidus
> ks.test(residus2, "pnorm", mean = mean(residus2), sd = sd(residus2))

Asymptotic one-sample Kolmogorov-Smirnov test

data: residus2
D = 0.075036, p-value = 0.09039
alternative hypothesis: two-sided

> bptest(lm_model2)

studentized Breusch-Pagan test

data: lm_model2
BP = 3.7142, df = 4, p-value = 0.4461

> reset(lm_model2)

RESET test

data: lm_model2
RESET = 0.21788, df1 = 2, df2 = 268, p-value = 0.8044

> vif(lm_model2)
taux_change      petrole precipitation  exportations
    1.106656     1.113121     1.011154     1.005594
> checkresiduals(lm_model2)

Breusch-Godfrey test for serial correlation of order up to 10

data: Residuals
LM test = 85.459, df = 10, p-value = 4.237e-14

```

Annexe 10. Modèle III– Index temporel

```

Call:
lm(formula = (IPA) ~ taux_change + petrole + precipitation +
    exportations + time, data = st3_data)

Residuals:
    Min      1Q      Median      3Q      Max 
-0.052038 -0.013781 -0.002281  0.010896  0.074264

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 3.606e-02 1.088e-02  3.313  0.00105 ** 
taux_change 4.359e-02 1.084e-02  4.021 7.53e-05 *** 
petrole      8.601e-04 2.226e-04  3.864  0.00014 *** 
precipitation -2.207e-04 7.061e-05 -3.126  0.00197 ** 
exportations 2.781e-04 8.938e-05  3.111  0.00206 ** 
time        -7.182e-05 2.778e-05 -2.586  0.01025 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02122 on 269 degrees of freedom
Multiple R-squared:  0.1265,   Adjusted R-squared:  0.1102 
F-statistic:  7.79 on 5 and 269 DF,  p-value: 7.282e-07

```

```

> ## TESTS
> residus3 <- residuals(lm_model3)
> # Test de normalité des résidus
> ks.test(residus3, "pnorm", mean = mean(residus3), sd = sd(residus3))

Asymptotic one-sample Kolmogorov-Smirnov test

data: residus3
D = 0.066946, p-value = 0.1699
alternative hypothesis: two-sided

> reset(lm_model3)

RESET test

data: lm_model3
RESET = 0.67732, df1 = 2, df2 = 267, p-value = 0.5088

> bptest(lm_model3)

studentized Breusch-Pagan test

data: lm_model3
BP = 4.4605, df = 5, p-value = 0.4852

> vif(lm_model3)
taux_change      petrole precipitation exportations       time
1.113852        1.114739       1.018312      2.945289      2.970363
> checkresiduals(lm_model3)

Breusch-Godfrey test for serial correlation of order up to 10

data: Residuals
LM test = 80.408, df = 10, p-value = 4.177e-13

```

Annexe 11. Modèle IV avec régressions (tendance déterministe corrigée)

```

Call:
lm(formula = IPA_detrended ~ CDD_detrended + petrole_detrended +
    exportations_detrended + taux_change_detrended + precipitation_detrended,
    data = stationnaire_data)

Residuals:
    Min      1Q  Median      3Q     Max 
-61.876 -9.716 -1.679   7.230 114.837 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -1.146e-15  1.206e+00  0.000 1.0000000  
CDD_detrended 1.636e+00  9.472e-01  1.727 0.085330 .  
petrole_detrended 7.729e-01  2.100e-01  3.680 0.000282 *** 
exportations_detrended 2.524e-01  8.427e-02  2.995 0.002998 ** 
taux_change_detrended 2.129e+01  1.022e+01  2.082 0.038288 *  
precipitation_detrended -1.605e-01  8.034e-02 -1.997 0.046809 * 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 20 on 269 degrees of freedom
Multiple R-squared:  0.1173,    Adjusted R-squared:  0.1009 
F-statistic: 7.149 on 5 and 269 DF,  p-value: 2.673e-06

```

```

Asymptotic one-sample Kolmogorov-Smirnov test

data: residus4
D = 0.10967, p-value = 0.002679
alternative hypothesis: two-sided

> # Test de Breusch-Pagan pour l'hétéroscléasticité
> bptest(lm_model4)

studentized Breusch-Pagan test

data: lm_model4
BP = 8.9646, df = 5, p-value = 0.1105

> # Test RESET
> reset(lm_model4)

RESET test

data: lm_model4
RESET = 0.58521, df1 = 2, df2 = 267, p-value = 0.5577

> ##### multicollinéarité
> vif(lm_model4)
      CDD_detrended    petrole_detrended  exportations_detrended
           1.461501          1.116101            1.004650
taux_change_detrended precipitation_detrended
           1.110756          1.468620

> ##### Autocorrelation de résidus
> checkresiduals(lm_model4)

Breusch-Godfrey test for serial correlation of order up to 10

data: Residuals
LM test = 109.32, df = 10, p-value < 2.2e-16

```

Annexe 12. Modèle VI avec un lag sur la variable IPA

```

Call:
lm(formula = (IPA) ~ IPA_lag1 + taux_change + petrole + precipitation,
  data = st3_data)

Residuals:
      Min        1Q        Median        3Q       Max
-3.690e-16  5.900e-19  1.390e-18  2.150e-18  2.202e-17

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -8.369e-18  1.130e-17 -7.410e-01  0.460
IPA_lag1     1.000e+00  6.392e-17  1.564e+16 <2e-16 ***
taux_change  1.465e-18  1.183e-17  1.240e-01  0.902
petrole      -1.174e-21  2.431e-19 -5.000e-03  0.996
precipitation -1.530e-20  7.617e-20 -2.010e-01  0.841
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.264e-17 on 270 degrees of freedom
Multiple R-squared:      1, Adjusted R-squared:      1
F-statistic: 6.76e+31 on 4 and 270 DF,  p-value: < 2.2e-16

```

```

Asymptotic one-sample Kolmogorov-Smirnov test

data: residus7
D = 0.41462, p-value < 2.2e-16
alternative hypothesis: two-sided

> bptest(model5)

studentized Breusch-Pagan test

data: model5
BP = 1.407, df = 4, p-value = 0.843

> reset(model5)

RESET test

data: model5
RESET = 34.971, df1 = 2, df2 = 268, p-value = 3.2e-14

> vif(model5)
  IPA_lag1   taux_change      petrole precipitation
    1.104992     1.165378     1.167885     1.040446
Warning message:
In summary.lm(object, ...) :
  essentially perfect fit: summary may be unreliable
> checkresiduals(model5)

Breusch-Godfrey test for serial correlation of order up to 10

data: Residuals
LM test = 1.1638, df = 10, p-value = 0.9997

```

Annexe 13. Modèle V mix - GLS - AR

```

Ljung-Box test

data: Residuals from REML
Q* = 128.74, df = 10, p-value < 2.2e-16

Model df: 0.  Total lags used: 10

Generalized least squares fit by REML
  Model: IPA ~ taux_change + petrole + precipitation + exportations
  Data: st2_data
        AIC      BIC  logLik
-1358.525 -1333.336 686.2625

Correlation Structure: AR(1)
  Formula: ~1
  Parameter estimate(s):
    Phi
0.5772593

Coefficients:
            Value Std.Error t-value p-value
(Intercept) 0.014436953 0.009895293 1.458972 0.1457
taux_change  0.023704728 0.008770127 2.702895 0.0073
petrole       0.000515314 0.000186171 2.767966 0.0060
precipitation -0.000085580 0.000059949 -1.427533 0.1546
exportations  0.000123323 0.000080308 1.535635 0.1258

```

Annexe 14. Modèle ARMAX avec variables sélectionnées et IPA diff log

```

> summary(armax_model4)
Series: y4
Regression with ARIMA(1,0,0) errors

Coefficients:
            ar1  intercept  exportations  petrole  precipitation  taux_change
            0.5706     0.0146      1e-04     5e-04      -1e-04      0.0238
s.e.    0.0501     0.0108      1e-04     2e-04      1e-04      0.0088

sigma^2 = 0.0003175: log likelihood = 720.18
AIC=-1426.35 AICc=-1425.94 BIC=-1401.04

Training set error measures:
          ME        RMSE       MAE       MPE       MAPE       MASE
Training set 4.259534e-05 0.01762421 0.01372311 13.16189 288.5038 0.5377013
          ACF1
Training set 0.03113072
> coeftest(armax_model4)

z test of coefficients:

            Estimate Std. Error z value Pr(>|z|)
ar1      5.7057e-01 5.0115e-02 11.3853 < 2.2e-16 ***
intercept 1.4609e-02 1.0759e-02 1.3578 0.174539
exportations 1.2270e-04 1.0064e-04 1.2193 0.222748
petrole    5.1753e-04 1.9494e-04 2.6549 0.007934 **
precipitation -8.6525e-05 7.1846e-05 -1.2043 0.228467
taux_change 2.3816e-02 8.7529e-03 2.7209 0.006510 **
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Asymptotic one-sample Kolmogorov-Smirnov test

data: residuals_arimax
D = 0.059652, p-value = 0.2818
alternative hypothesis: two-sided

> # Test de Ljung-Box pour l'autocorrélation des résidus
>
> checkresiduals(armax_model4)

Ljung-Box test

data: Residuals from Regression with ARIMA(1,0,0) errors
Q* = 14.994, df = 23, p-value = 0.8948

Model df: 1. Total lags used: 24

> # Test de Breusch-Pagan pour l'hétéroscédasticité
> bptest(residuals_arimax ~ vars3)

studentized Breusch-Pagan test

data: residuals_arimax ~ vars3
BP = 12.819, df = 8, p-value = 0.1182

```

Annexe 15.Modèle avec les résidus ajustés

```
Call:
lm(formula = IPA ~ taux_change + petrole + precipitation + exportations +
    ar1_fitted, data = st2_data)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.048087 -0.011929 -0.001883  0.012298  0.046143 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 2.689e-02 9.236e-03  2.912  0.00390 ** 
taux_change 3.888e-02 9.192e-03  4.229 3.22e-05 *** 
petrole      5.884e-04 1.906e-04  3.087  0.00224 ** 
precipitation -1.603e-04 6.000e-05 -2.671  0.00802 ** 
exportations 9.157e-05 4.442e-05  2.062  0.04020 *  
ar1_fitted    1.022e+00 9.654e-02 10.589 < 2e-16 *** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01805 on 269 degrees of freedom
Multiple R-squared:  0.3681,   Adjusted R-squared:  0.3564 
F-statistic: 31.35 on 5 and 269 DF,  p-value: < 2.2e-16

Ljung-Box test

data: Residuals
Q* = 6.4074, df = 10, p-value = 0.78

Model df: 0.  Total lags used: 10

Asymptotic one-sample Kolmogorov-Smirnov test

data: residuals_corrected
D = 0.060107, p-value = 0.2735
alternative hypothesis: two-sided

> bptest(lm_corrected)

studentized Breusch-Pagan test

data: lm_corrected
BP = 17.395, df = 5, p-value = 0.003808

> reset(lm_corrected)

RESET test

data: lm_corrected
RESET = 5.2333, df1 = 2, df2 = 267, p-value = 0.005897

> vif(lm_corrected)
taux_change      petrole precipitation  exportations
    1.107362      1.130431      1.016294      1.005598
ar1_fitted
    1.019534
```

Annexe 16. Package

Ce projet a été réalisé en utilisant Quarto avec le logiciel et langage R pour l'analyse des données. Les packages utilisés sont :

BeSS - Méthodes pour la sélection de meilleurs sous-ensembles

car - Fournit des fonctions pour les modèles linéaires, y compris les diagnostics et les tests

dplyr - Utilisé pour la manipulation de données

EnvStats - Offre des outils pour l'analyse statistique

forecast - Méthodes pour la prévision de séries temporelles

ggcorrplot - Utilisé pour visualiser des matrices de corrélation

ggplot2 - Utilisé pour créer des visualisations de données complexes

leaps - Implémente des méthodes pour la sélection de modèle de régression

lmtest - Offre des tests statistiques pour les modèles de régression linéaire

lubridate - Facilite le travail avec des dates et des heures

MASS - Contient des fonctions et des ensembles de données

nlme - Utilisé pour l'ajustement des modèles linéaires et non linéaires à effets mixtes

PerformanceAnalytics - Fournit des outils pour l'analyse de performance

plotly - Permet de rendre des graphiques ggplot2 interactifs

RJDemetra - Fournit une interface pour les méthodes de désaisonnalisation

readxl - Utilisé pour lire des fichiers Excel

seastests - Pour effectuer des tests statistiques sur des données saisonnières

tidyverse - Fournit des outils pour transformer les données

TSA - Fournit des outils pour l'analyse de séries temporelles

trend - Utilisé pour détecter et gérer les tendances dans les données temporelles

tsoutliers - Déetecte et gère les valeurs aberrantes dans les séries temporelles

tseries - Contient des fonctions pour l'analyse de séries temporelle

Source des données

Tableau 12 Sources

Variable	Source	Mesure
IPA-DI - origine - prod. Agricole	IPEA DATA http://www.ipeadata.gov.br/Default.aspx	Indices (ago. 1994 = 100)
Température	FAO STAT https://www.fao.org/faostat/fr/#data/ET	Degrés Celsius
Précipitation	Climate change knowledge portal https://climateknowledgeportal.worldbank.org/download-data#htab-1499	Millimètres (mm)
CDD : Nombre de jours secs consécutifs	Climate change knowledge portal https://climateknowledgeportal.worldbank.org	Jours
Prix du pétrole	IPEA http://www.ipeadata.gov.br/Default.aspx Preço - petróleo bruto - Brent (FOB)	Dollar US\$
SMIC (Salaire minimum nominal)	IPEA DATA : Salário mínimo nominal	Real brésilien (R\$)
Exportations	IPEA DATA Exportações - agricultura e pecuária - quantum	Indice (2018 = 100)
Importations	IPEA DATA Importações - agricultura e pecuária - quantum	Indice (2018 = 100)
Taux de change	IPEA DATA Taxa de câmbio - R\$ / US\$ - commercial - venda - média	Real brésilien (R\$)

Tableau de figures

FIGURE 1 IPA-DI - ORIGINE - PROD. AGRICOLE - INDICE (AOUT 1994 = 100) JAN 2000 A DEC 2022	15
FIGURE 2 VARIATION DE TEMPERATURES EN °C	18
FIGURE 3 PRECIPITATIONS EN MM	21
FIGURE 4 NOMBRE DE JOURS SECS CONSECUTIFS	22
FIGURE 5 PRIX DU PETROLE EN DOLLAR US	23
FIGURE 6 SALAIRE MINIMUM NOMINAL ACTUEL EN R\$	25
FIGURE 7 EXPORTATIONS DES PRODUITS AGRICOLES	27
FIGURE 8 IMPORTATION DES PRODUITS AGRICOLES	28
FIGURE 9 TAUX DE CHANGE - R\$ / US\$	29
FIGURE 10 DETECTION DES POINTS ATYPIQUES SUR LA SERIE NIVEAU SERIE IPA	34
FIGURE 11 MATRICE DE CORRELATION SERIES AJUSTEES	37
FIGURE 12 PERIODOGRAMME VARIABLES SAISONNIERES	41
FIGURE 13 BOXPLOT VARIABLES PRECIPITATION ET CDD	42
FIGURE 14 TENDANCES - SERIES NON STATIONNAIRES	45
FIGURE 15 DECOMPOSITION ET CORRECTION DE LA SERIE CDD	48
FIGURE 16 BEST SUBSET	50
FIGURE 17 RESIDUS MODELE ARMAX	58
FIGURE 18 QUALITE D'AJUSTEMENT ARMAX	60
TABLEAU 1 STATISTIQUES DESCRIPTIVES	31
TABLEAU 2 OUTLIERS SERIE IPA	33
TABLEAU 3 SKEWNESS – KURTOSIS - NORMALITE	35
TABLEAU 4 TESTS DE SAISONNALITE	40
TABLEAU 5 TEST DE TENDANCE MANN-KENDALL	43
TABLEAU 6 RESULTATS DU TEST DE STATIONNARITE ADF	44
TABLEAU 7 KPSS TEST	46
TABLEAU 8 TEST LOG LEVEL	47
TABLEAU 9 RESUME STATIONNARITE	49
TABLEAU 10 RESULTAT MODELES MCO	57
TABLEAU 11 COEFFICIENTS DE REGRESSION	59
TABLEAU 12 SOURCES	87

Table des matières

RESUME	3
ABSTRACT	3
Glossaire des sigles.....	5
1 Introduction.....	1
2 Environnement économique et politique.....	4
3 Méthodologie économétrique.....	7
3.1 STATIONNARITE	7
3.2 SAISONNALITE	9
3.3 MODELES DE REGRESSION	10
3.3.1 MOINDRES CARRES ORDINAIRES (MCO).....	10
3.3.2 REGRESSION LINEAIRE ET SERIES TEMPORELLES.....	11
3.3.3 MODELEAR.....	12
3.3.4 MODELE MOYENNE MOBILE MA.....	12
3.3.5 MODELE ARMA(P, Q).....	13
3.3.6 MODELE ARMAX.....	13
4 Présentation des données.....	14
4.1 VARIABLE D'INTERET : INDICE DES PRIX A LA PRODUCTION AGRICOLE	14
4.2 VARIABLES CLIMATIQUES.....	17
4.2.1 VARIATION DE TEMPERATURE.....	18
4.2.2 PRECIPITATIONS	20
4.2.3 LE NOMBRE DE JOURS SECS CONSECUTIFS (CDD)	21
4.3 VARIABLES ECONOMIQUES.....	22
4.3.1 PRIX DU PETROLE.....	22
4.3.2 SMIC.....	24
4.3.3 EXPORTATIONS.....	26
4.3.4 IMPORTATIONS POUR L'AGRICULTURE ET L'ELEVAGE	28
4.3.5 TAUX DE CHANGE	29
5 Analyse de Données.....	30
5.1 STATISTIQUES DESCRIPTIVES	30
5.2 DETECTION ET CORRECTIONS DES VALEURS ATYPIQUES	32
5.2.1 SKEWNESS, KURTOSIS ET TEST DE NORMALITÉ DE LA SERIE AJUSTEE	34
5.3 CORRELATION	37
6 Estimation économétriques et interprétation des résultats	39

6.1	IDENTIFICATION DES COMPOSANTS	39
6.1.1	<i>TEST DE SAISONNALITE</i>	39
6.1.2	<i>ANALYSE DE LA TENDANCE</i>	43
6.2	DONNEES CORRIGEES DES VARIATIONS SAISONNIERES ET TENDANCES	47
6.3	SELECTION DES VARIABLES	49
6.4	ESTIMATIONS ET VALIDATION DES MODELES MCO	50
6.4.1	<i>MODELE I</i>	51
6.4.2	<i>MODELE II- FORME SEMI LOGARITHMIQUE</i>	52
6.4.3	<i>MODELE III – AVEC VARIABLE DE TEMPS</i>	53
6.4.4	<i>MODELE IV - AVEC CORRECTION DE LA TENDANCE DETERMINISTE</i>	55
6.4.5	<i>MODELE V DE MOINDRES CARRES GENERALISES ET AR(1)</i>	56
6.4.6	<i>MODELE VI – MCO AVEC « LAG »</i>	56
6.5	ESTIMATIONS ET VALIDATION DES MODELES ARMAX	57
6.6	INTERPRETATION DU MODELE SELECTIONNE	58
7	Conclusion	61
8	Discussion.....	63
9	Bibliographie	66
10	Annexes.....	70
	Source des données	87
	Tableau de figures	88
	Table des matières	89