

Description et présentation de la série temporelle	3
Analyse préliminaires du graphique de la série chronologiques	3
Décomposition et Lissage exponentiel	4
Commentaire 1	5
Lissage exponentielle , Modèles à espaces d'états et test de robustesse	6
Holt winters	6
Commentaire 2	7
Commentaire 3	8
ETS	8
Commentaire 4	9
Test de robustesse ETS	10
Commentaire 5	10
Modèle Box-Jenkins	10
Commentaire 6	11
Commentaire 7	12
Commentaire 8	13
Commentaire 9	15
Prévisions et Intervalles de prévisions	15
Commentaire 10	15
Autres Méthodes de prévisions Et Tests de Robustesse	16
Tbats	16
STL	16
Combinaisons	17
Conclusion et Vulgarisation des résultats	19

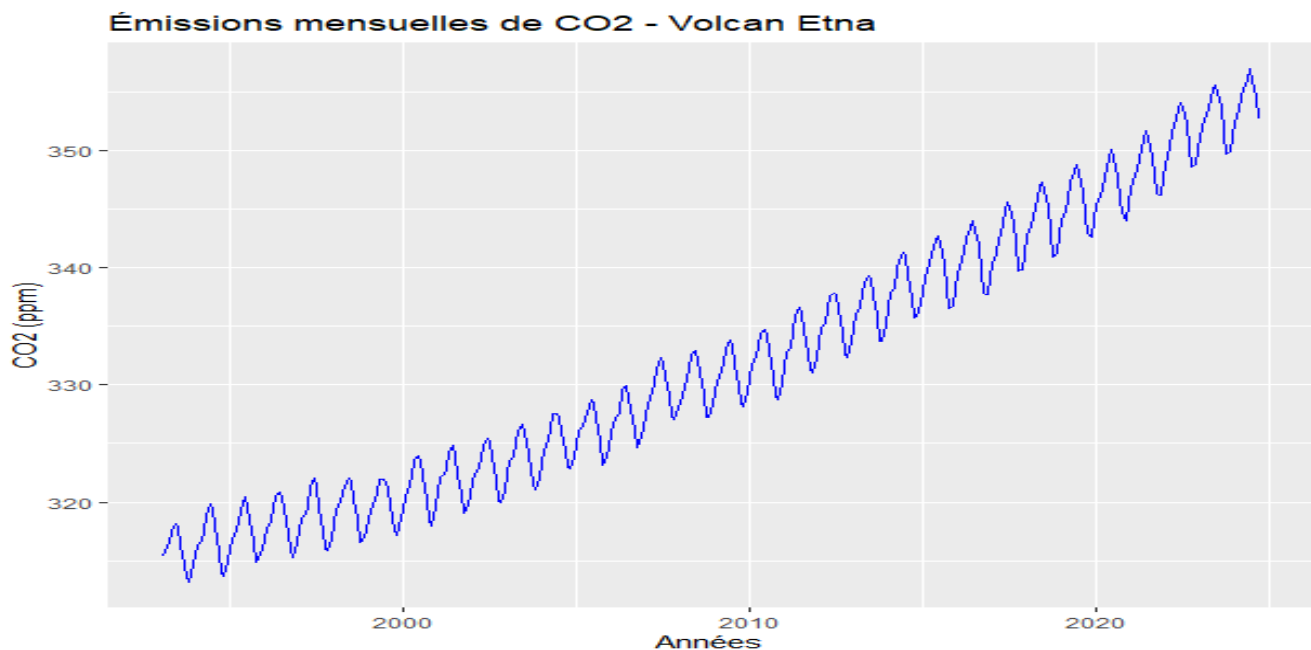
Description et présentation de la série temporelle

Description

L'Etna est un volcan actif situé en Sicile, en Italie. Notre série temporelle analyse les émissions mensuelles de dioxyde de carbone (CO_2) de ce volcan, mesurées en parties par million (ppm), à partir de février 1993. Cette étude est pertinente puisque les variations d'émissions de CO_2 constituent un indicateur précurseur d'une éventuelle éruption volcanique.

Présentation de la série

Les données de la série couvrent une période de 31 ans, de février 1993 à septembre 2024, avec une fréquence mensuelle, pour un total de 380 observations.

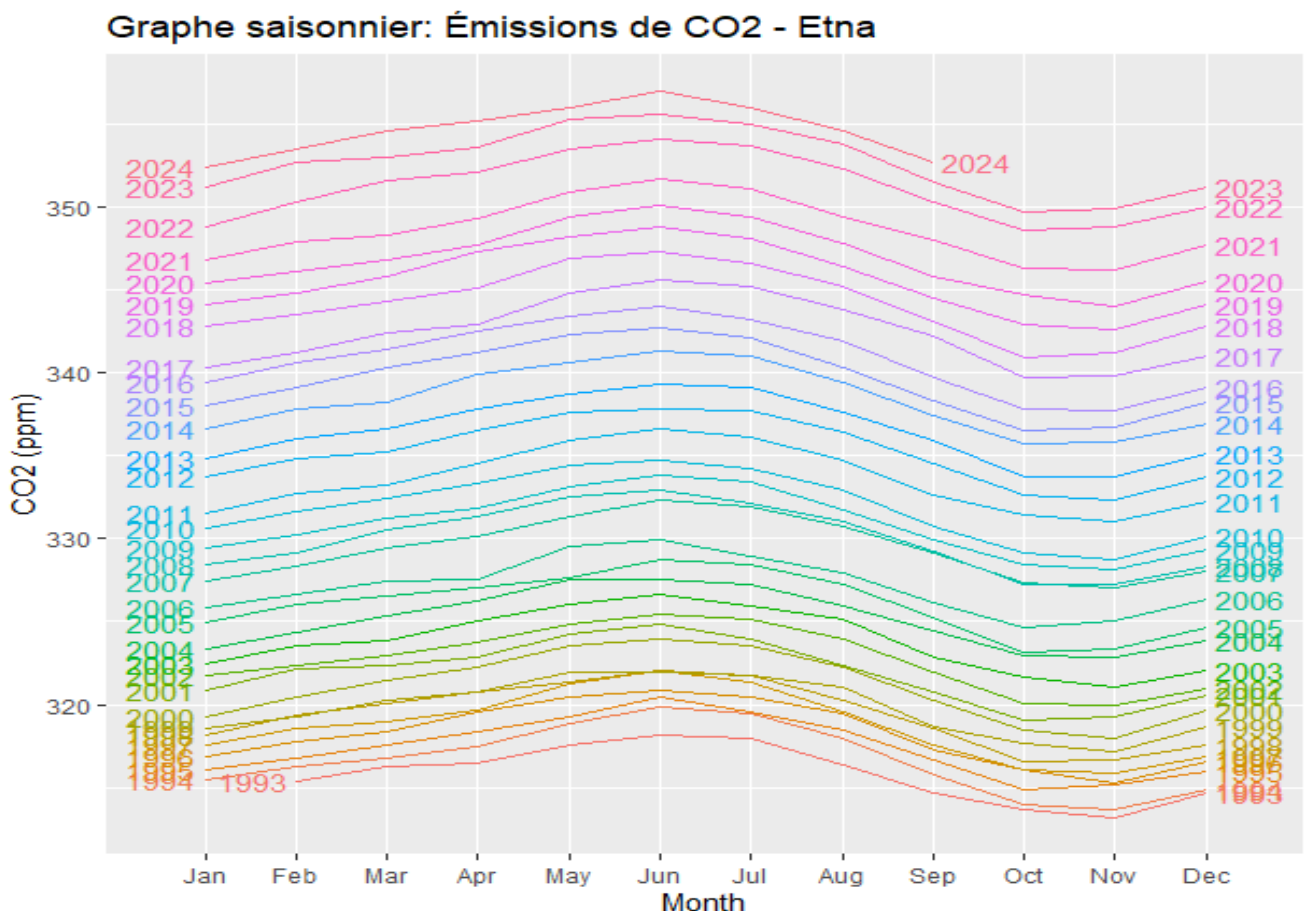


Graphique 1: Graphique de la série dioxcar : émissions mensuelles de CO_2 du volcan Etna

Analyse préliminaires du graphique de la série chronologiques

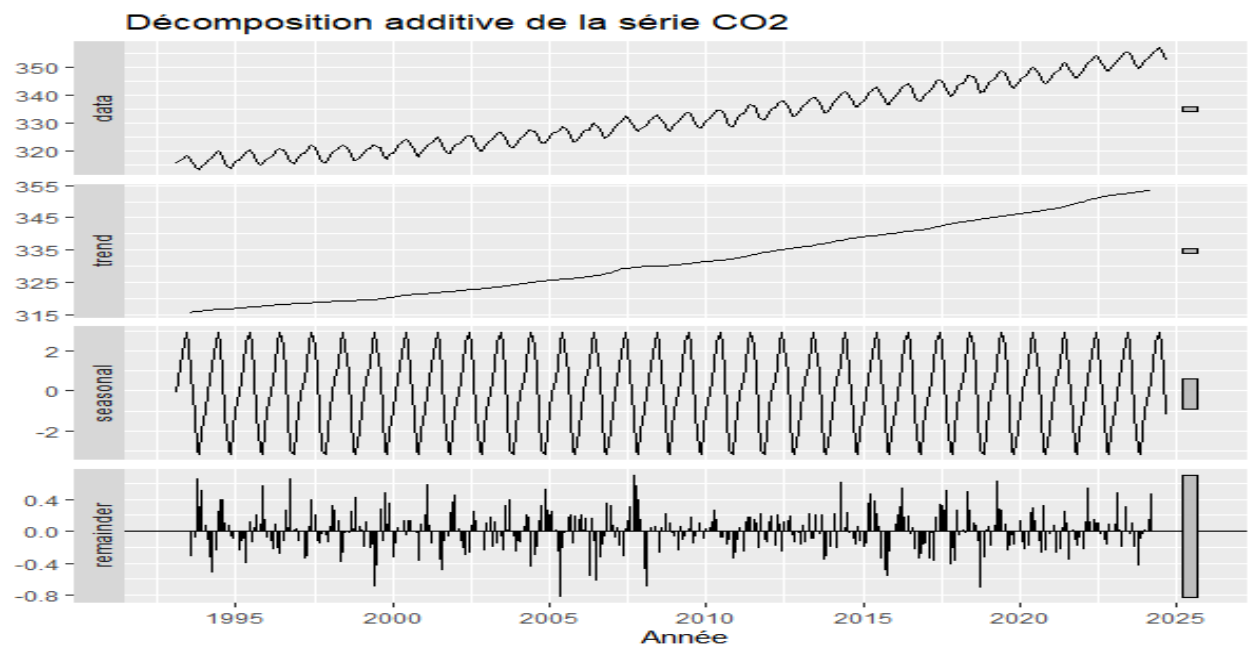
En visualisant la série temporelle, plusieurs composantes importantes apparaissent. On observe une tendance croissante marquée et une saisonnalité annuelle très régulière. La variance reste stable au cours du temps, ce qui indique que la série est homoscédastique : l'amplitude des fluctuations ne s'accroît pas avec le niveau des émissions. La série est donc non stationnaire en moyenne, à cause de la tendance et de la saisonnalité, mais stationnaire en variance, ce qui explique pourquoi aucune transformation Box-Cox n'est nécessaire. Une différenciation saisonnière est indispensable pour rendre la série stationnaire, puisqu'elle contient une composante saisonnière persistante.

Décomposition et Lissage exponentiel

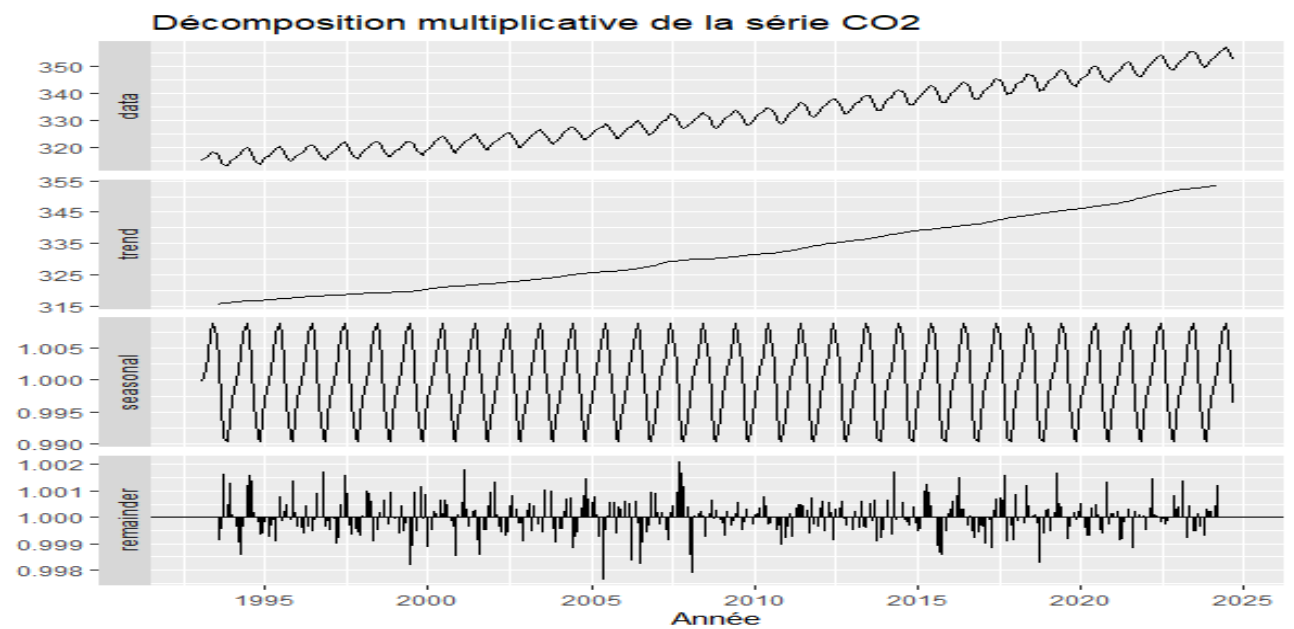


Graphique 2 :Graphes Saisonnier

Les graphes saisonniers mettent en évidence la stabilité structurelle de la série durant l'année, avec des pics au printemps et des creux en automne, sans modification notable au fil des années. Les courbes se décalent verticalement, suggérant une évolution du niveau moyen sans changement de structure saisonnière.



.Graphique: 3: Décomposition Additive



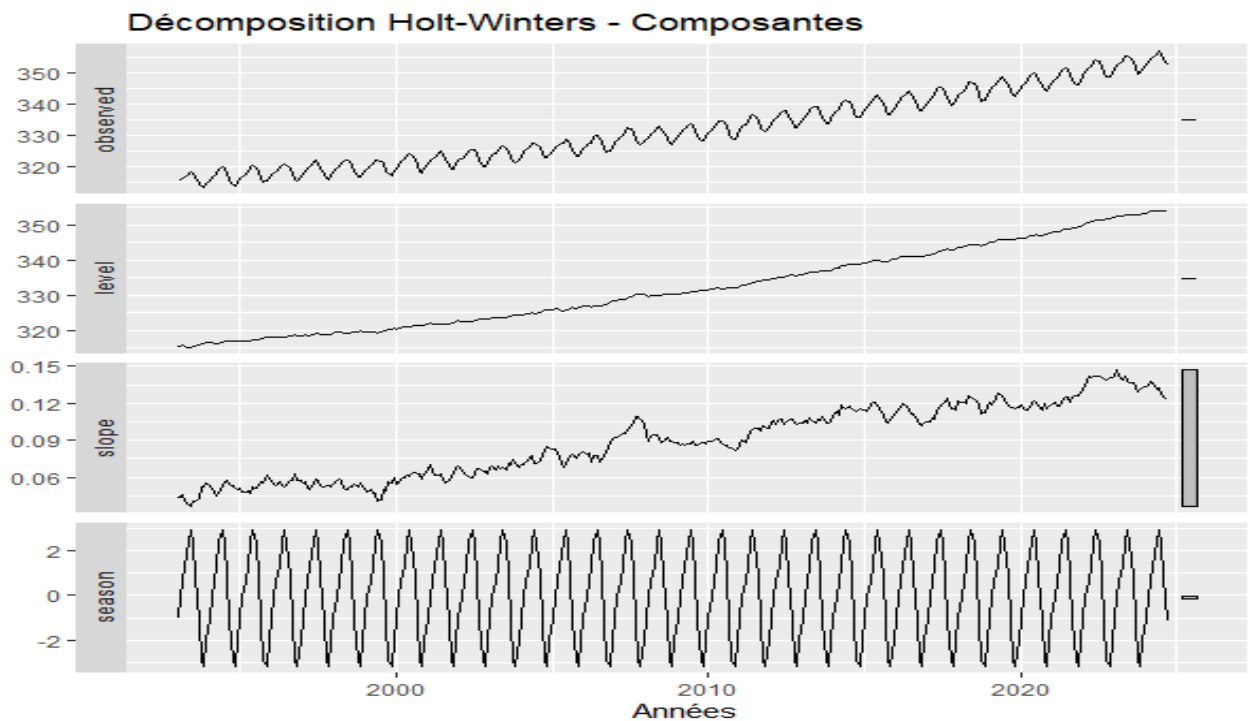
Graphique 4: Décomposition Multiplicative

Commentaire 1

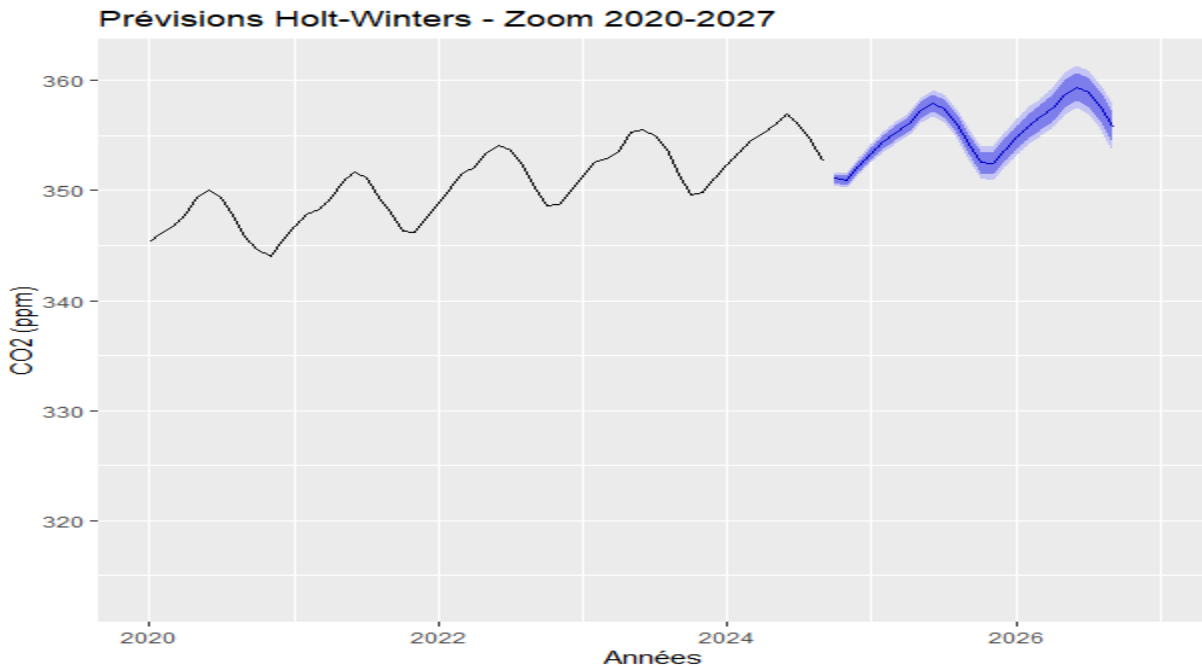
La série présente une tendance haussière claire (de 315 à 355 ppm sur 30 ans) et une saisonnalité régulière d'amplitude constante. Les deux décompositions (additive et multiplicative) donnent des résultats quasi identiques, ce qui confirme que l'amplitude saisonnière ne varie pas avec le niveau de la série — le modèle additif est donc approprié. Les résidus montrent quelques anomalies vers 2007-2008, possiblement liées à une activité volcanique inhabituelle.

Lissage exponentielle , Modèles à espaces d'états et test de robustesse

Holt winters



Graphique 5: Décomposition Holt winters

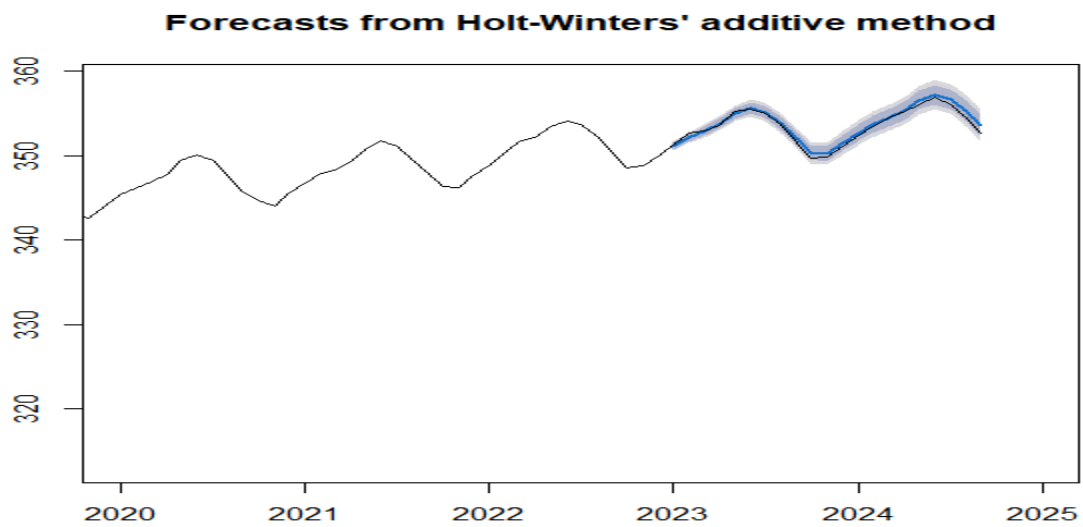


Graphique 6 :Prévisions Holt-winters

Commentaire 2

La méthode de Holt-Winters décompose la série en trois composantes (Graphique 5): le niveau croît de 315 à 355 ppm, la tendance s'accélère de 0.05 à 0.15 ppm/mois, et la saisonnalité présente des cycles réguliers de ± 2 ppm. Les prévisions (Graphique 6) prolongent cette dynamique avec des intervalles de confiance s'élargissant progressivement, atteignant 360 ppm en 2027. Le modèle offre une interprétation simple des composantes mais repose sur l'hypothèse d'une tendance linéaire, ce qui pourrait limiter sa précision à long terme si la croissance change de rythme.

Test de Robustesse Holt winters

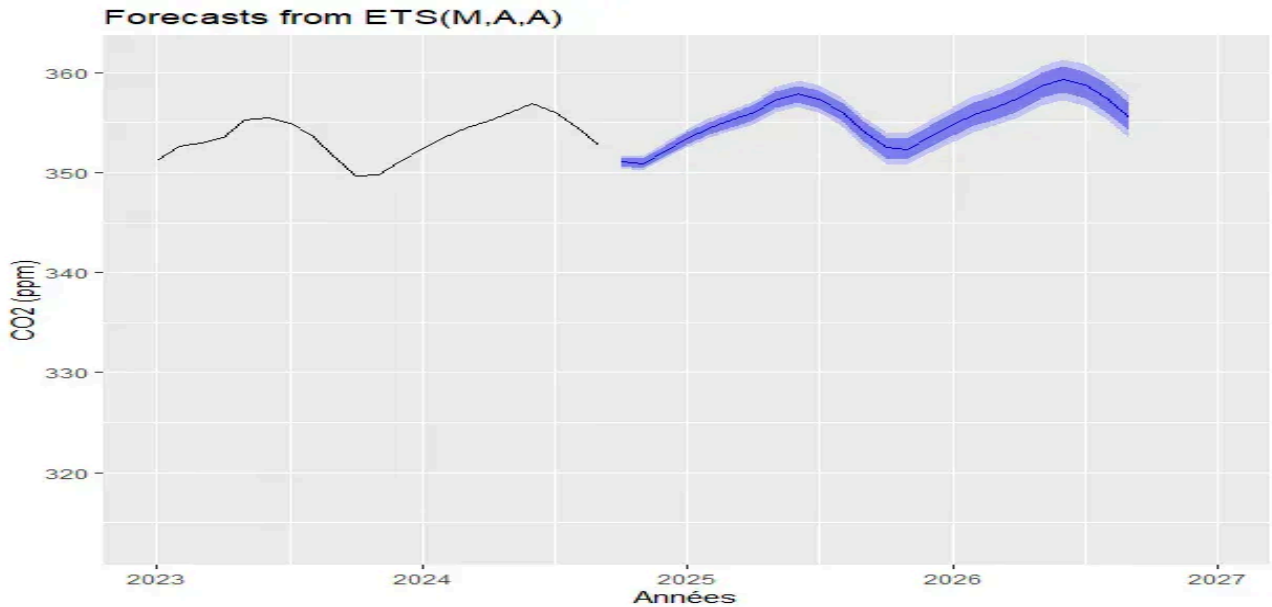


Graphique 7: Pr vision Holt-winters

Commentaire 3

Le mod le Holt-Winters (RMSE test = 0.432) capture la tendance et la saisonnalit  de mani re satisfaisante, comme l'illustre la continuit  visuelle entre les donn es historiques et les pr visions. Cependant, le test de Ljung-Box ($p = 0.039$) r v le une autocorr lation r siduelle significative, violant l'hypoth se de bruit blanc. Cette limitation sous-estime les intervalles de confiance et sugg re que des mod les plus complexes (ETS, ARIMA) capturent mieux la structure temporelle de la s rie.

ETS

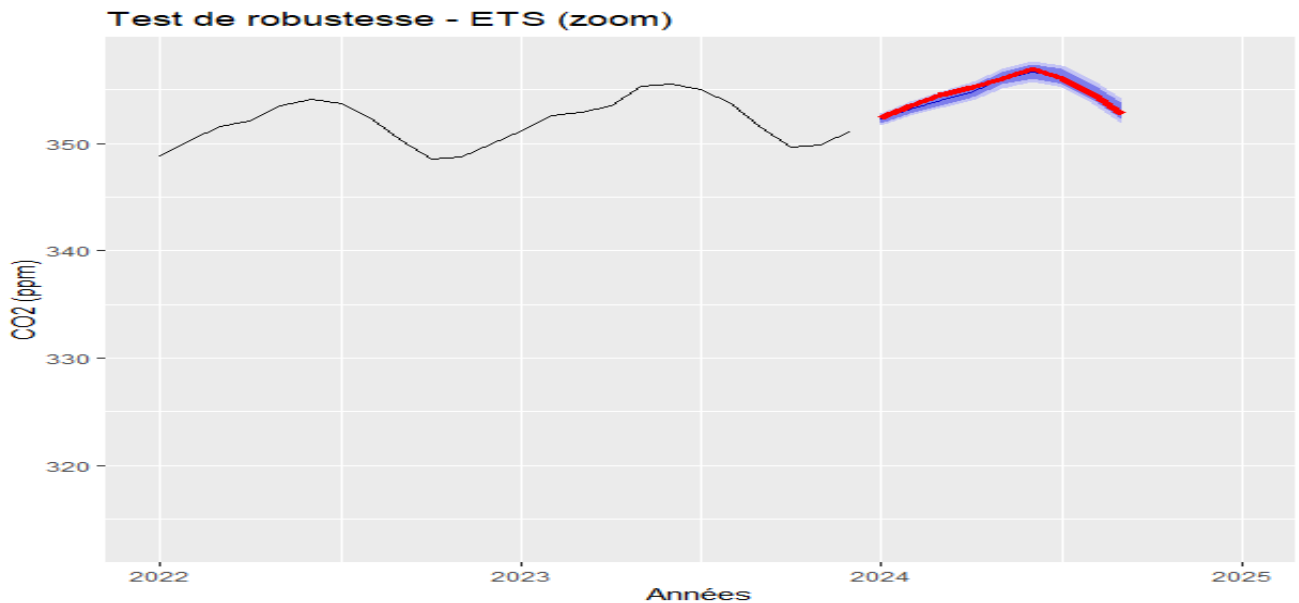


Graphique 8:Prévisions ETS

Commentaire 4

ETS(M,A,A) est un modèle à espaces d'états où l'erreur est multiplicative, la tendance est additive et la saisonnalité est additive. Ce cadre permet d'obtenir des intervalles de prévision fiables, visibles sur le graphique (zones bleues qui s'élargissent avec l'horizon). Le modèle ETS(M,A,A) sélectionné automatiquement est nettement supérieur aux méthodes basiques pour cette série. La méthode moyenne ignorait la tendance haussière, la méthode naïve ignorerait la saisonnalité, et la naïve saisonnière sous-estimerait les valeurs futures en ignorant la tendance. ETS intègre ces trois composantes et fournit des intervalles de prévision probabilistes qui s'élargissent avec l'horizon, reflétant l'incertitude croissante.

Test de robustesse ETS



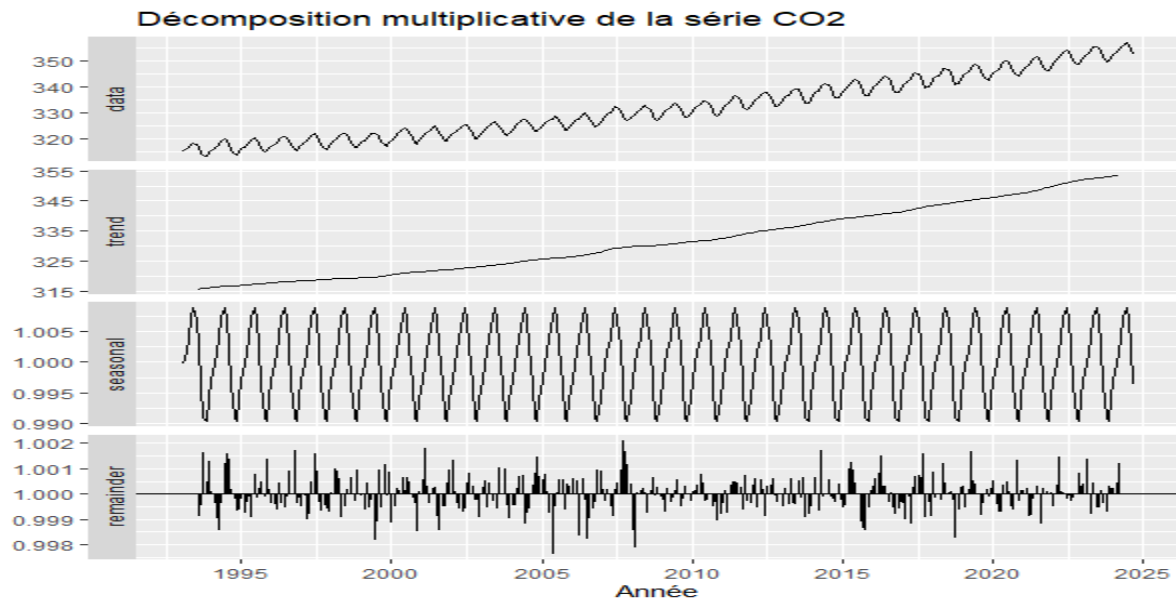
Graphique 9: Test de robustesse de ETS

Commentaire 5

Le test de robustesse utilise le modèle ETS(M,A,A) — erreur multiplicative, tendance et saisonnalité additives. Le RMSE test (0.272) est comparable au RMSE train (0.277), confirmant que le modèle ne surajuste pas. Le MAPE inférieur à 0.1% indique une précision élevée. Le faible écart entre les métriques train/test confirme la stabilité et la robustesse du modèle pour des prévisions à court terme. (combien d'années pour le modèle de ETS)

Modèle Box-Jenkins

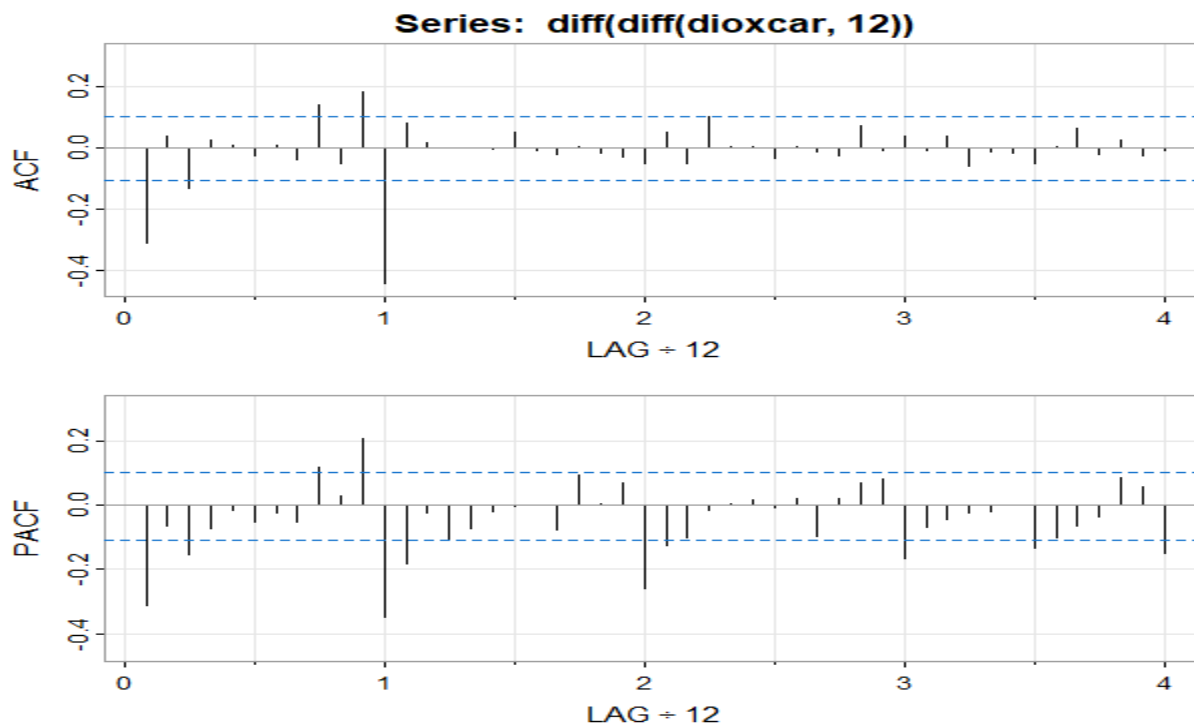
1. Identification du modèle



Graphique 10: décomposition multiplicative de la série CO2

Commentaire 6

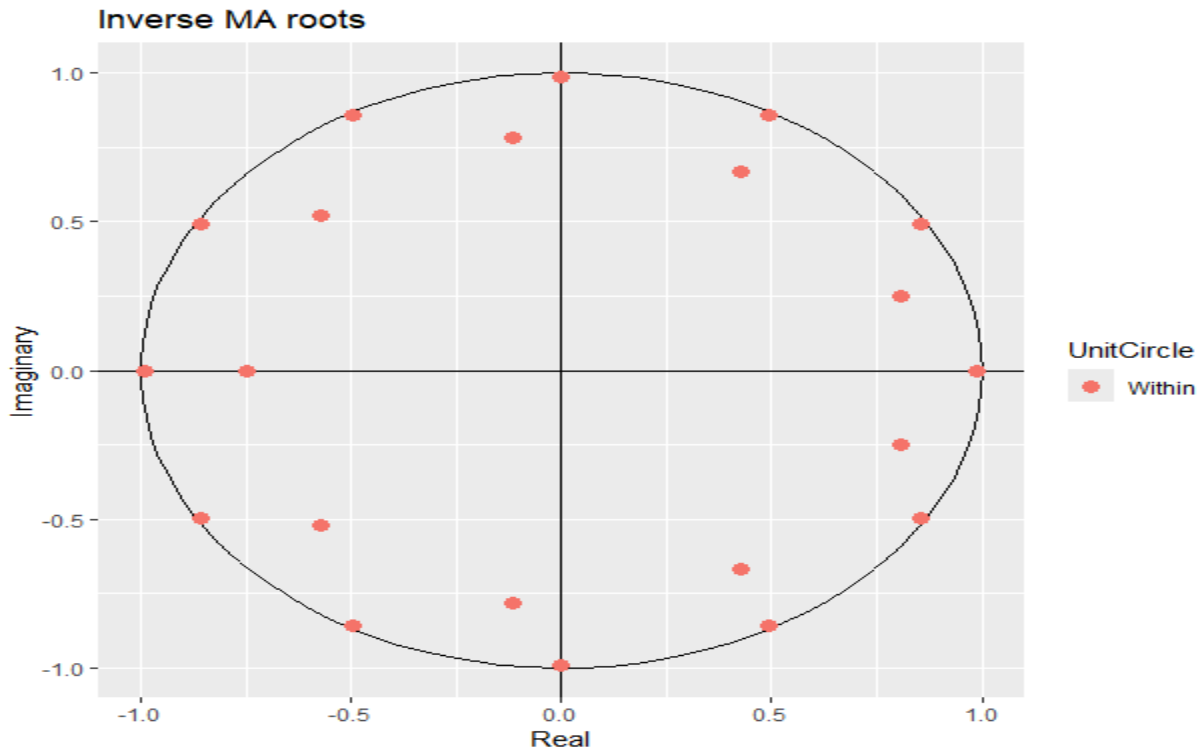
La série présente une tendance haussière ainsi qu'une saisonnalité annuelle marquée. La série a ensuite été différenciée saisonnièrement ($D = 1$) pour éliminer la composante saisonnière. Les tests ADF (Augmented Dickey–Fuller) et KPSS réalisés sur cette série montrent que celle-ci demeure non stationnaire. Une différenciation supplémentaire d'ordre $d = 1$ est donc nécessaire afin d'obtenir une série stationnaire.



Graphique 11: Graphiques des corrélations partielles et totales

Commentaire 7

Dans le cadre de l'approche Box–Jenkins, seize modèles SARIMA ont été estimés après application des différenciations $d = 1$ et $D = 1$. Les modèles testés incluent diverses combinaisons de composantes autorégressives et moyennes mobiles, tant non saisonnières que saisonnières : $\text{SARIMA}(0,1,0)(1,1,1)_{12}$, $\text{SARIMA}(1,1,1)(0,1,1)_{12}$, $\text{SARIMA}(1,1,0)(0,1,1)_{12}$, $\text{SARIMA}(2,1,0)(0,1,1)_{12}$, $\text{SARIMA}(2,1,1)(0,1,1)_{12}$, $\text{SARIMA}(1,1,2)(0,1,1)_{12}$, $\text{SARIMA}(1,1,0)(0,1,2)_{12}$, $\text{SARIMA}(0,1,3)(0,1,1)_{12}$ avec différentes contraintes sur les coefficients, $\text{SARIMA}(1,1,3)(0,1,1)_{12}$, $\text{SARIMA}(0,1,1)(0,1,1)_{12}$, et finalement $\text{SARIMA}(0,1,9)(0,1,1)_{12}$ avec structure parcimonieuse. L'ensemble de ces modèles a été comparé selon les critères AIC, BIC et AICc, les tests diagnostiques (Ljung-Box, Shapiro-Wilk) et la significativité des coefficients. Le modèle retenu est le $\text{SARIMA}(0,1,9)(0,1,1)_{12}$ avec coefficients MA aux retards 1, 3 et 9 uniquement. Il présente le meilleur AIC (132.74), tous les coefficients sont hautement significatifs ($|t| > 1.82$), et les résidus sont compatibles avec un bruit blanc gaussien (Ljung-Box : $p = 0.91$, Shapiro-Wilk : $p = 0.47$), tout en offrant une interprétation potentielle des dynamiques d'émission volcanique à différentes échelles temporelles.

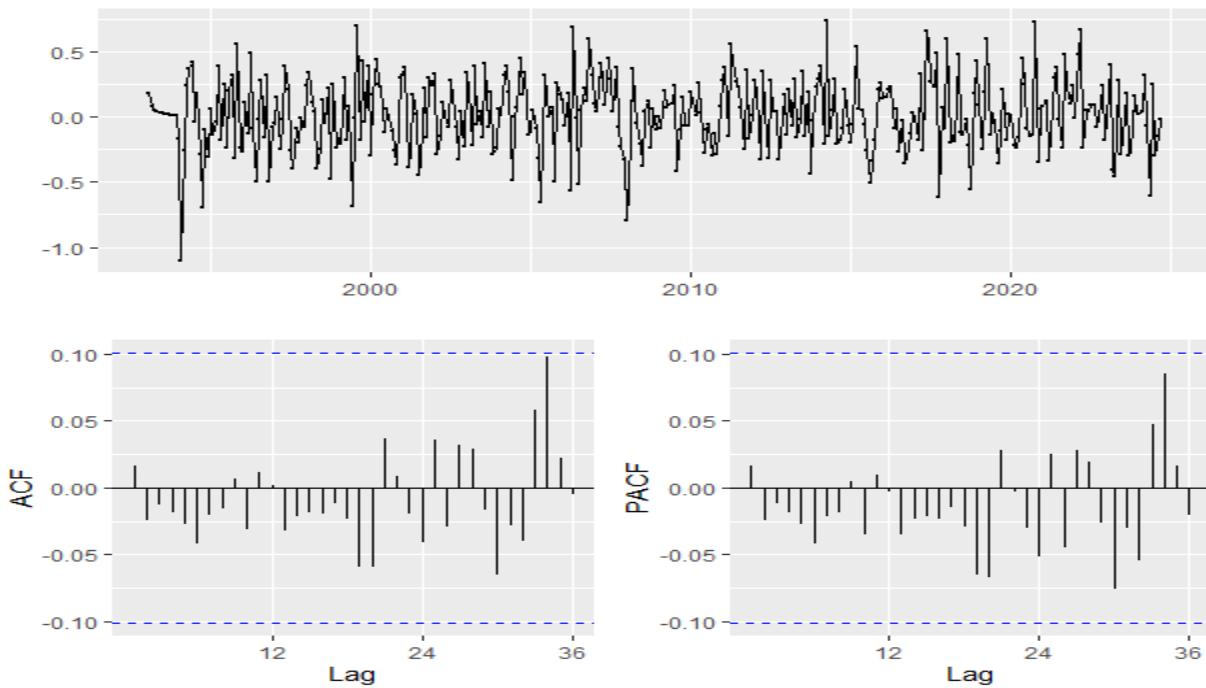


Graphique 12: les racines inverse du polynômes caractéristiques sont à l'intérieur du cercle unité.

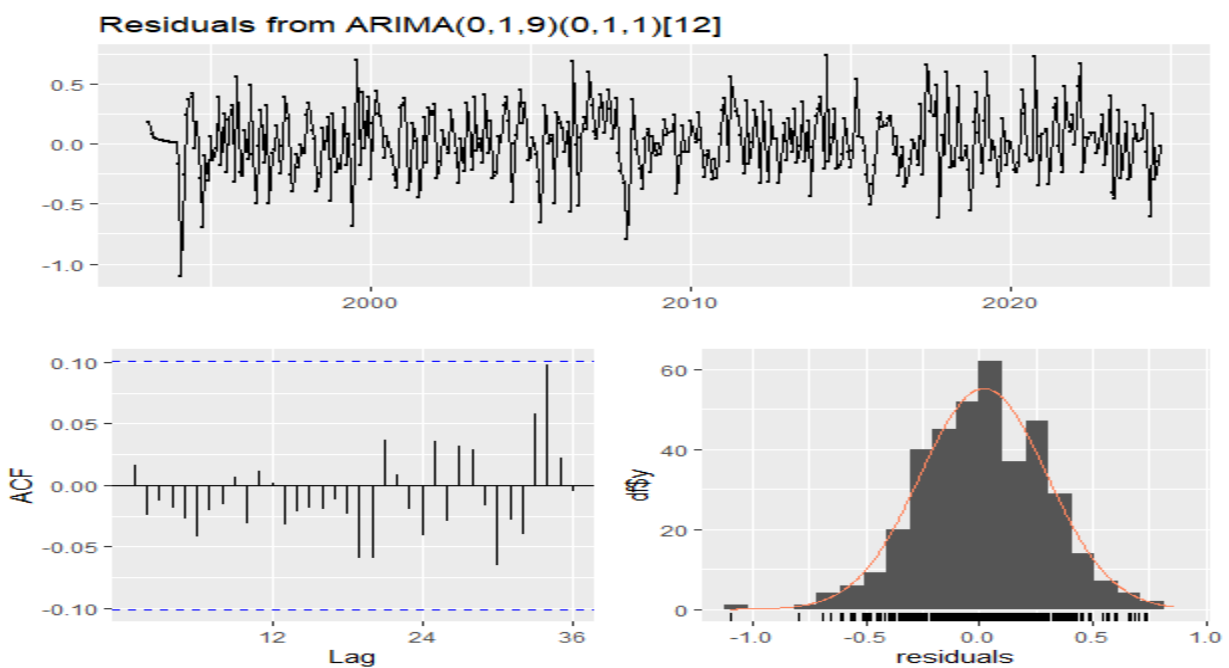
Commentaire 8

Au sujet du modèle : $SARIMA(0,1,9)(0,1,1)_{12}$ avec coefficients MA aux retards 1, 3 et 9 uniquement, L'ensemble des racines du polynôme caractéristique se situent à l'extérieur du cercle unité, ce qui confirme que le modèle estimé est stationnaire pour la partie autorégressive et inversible pour la partie moyenne mobile. Les conditions de validité du modèle Box–Jenkins

sont respectées.



Graphique 13: La nullité des résidus

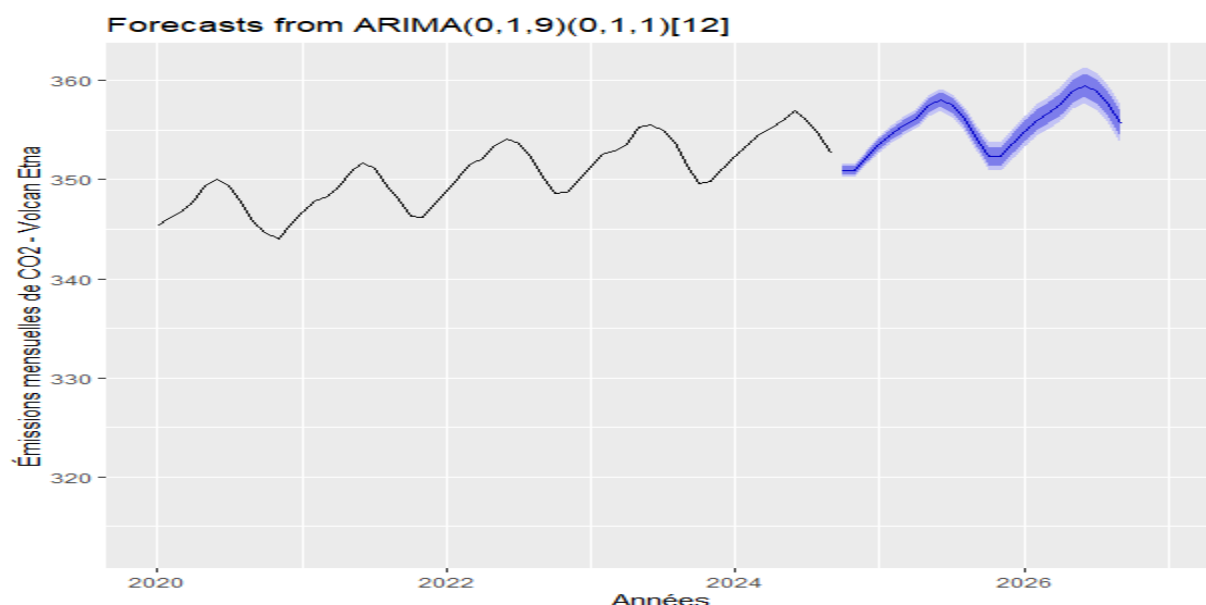


Graphique 14: La normalité des résidus

Commentaire 9

Les différents diagnostics confirment la qualité statistique du modèle $\text{ARIMA}(0,1,9)(0,1,1)[12]$. Le test de Shapiro–Wilk ($p = 0.26$) indique que les résidus ne s'écartent pas significativement de la normalité, tandis que le test de Ljung–Box ($p = 0.91$) montre l'absence d'autocorrélation résiduelle. Combinés à une moyenne de résidus non significativement différente de zéro, ces résultats suggèrent que le modèle capture adéquatement la dynamique de la série et que les hypothèses classiques des modèles ARIMA sont satisfaites.

Prévisions et Intervalles de prévisions



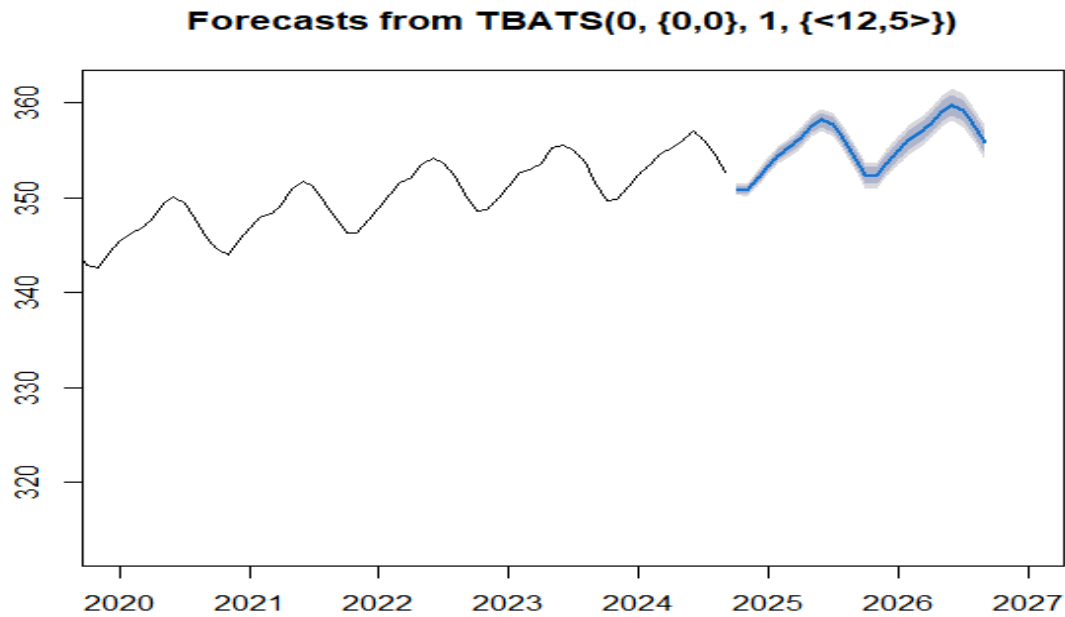
Graphique 15: Prevision avec le modèle $\text{SARIMA}(0,1,9)(0,1,1)_{12}$ sur 24 mois

Commentaire 10

Le graphique présente les prévisions du modèle $\text{SARIMA}(0,1,9)(0,1,1)_{12}$ sur 24 mois (2024-2026). Le modèle capture fidèlement la tendance croissante et les cycles saisonniers des émissions de CO₂, avec des prévisions oscillant entre 350 et 360 ppm. Les performances sont excellentes : RMSE = 0.281 ppm et MAPE = 6.72%. Les intervalles de confiance à 80% et 95% s'élargissent normalement avec l'horizon de prévision et sont fiables grâce à la validation des résidus (Ljung-Box : $p = 0.91$, Shapiro-Wilk : $p = 0.47$). La structure parcimonieuse avec coefficients MA aux retards 1, 3 et 9 offre une interprétation des dynamiques d'émission à court, moyen et long terme, tout en présentant le meilleur AIC (132.74) parmi les modèles testés.

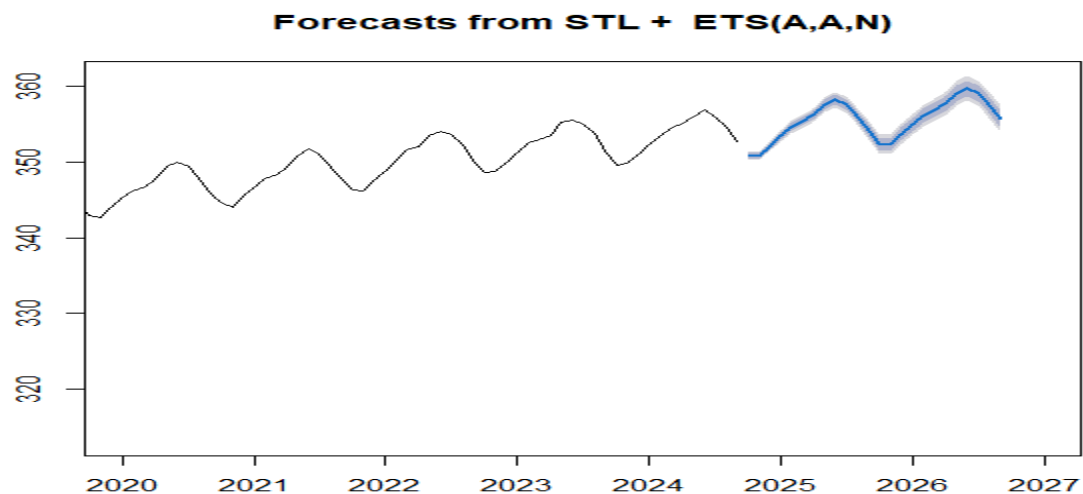
Autres Méthodes de prévisions Et Tests de Robustesse

Tbats



Graphique 16: Prédiction avec Tbats

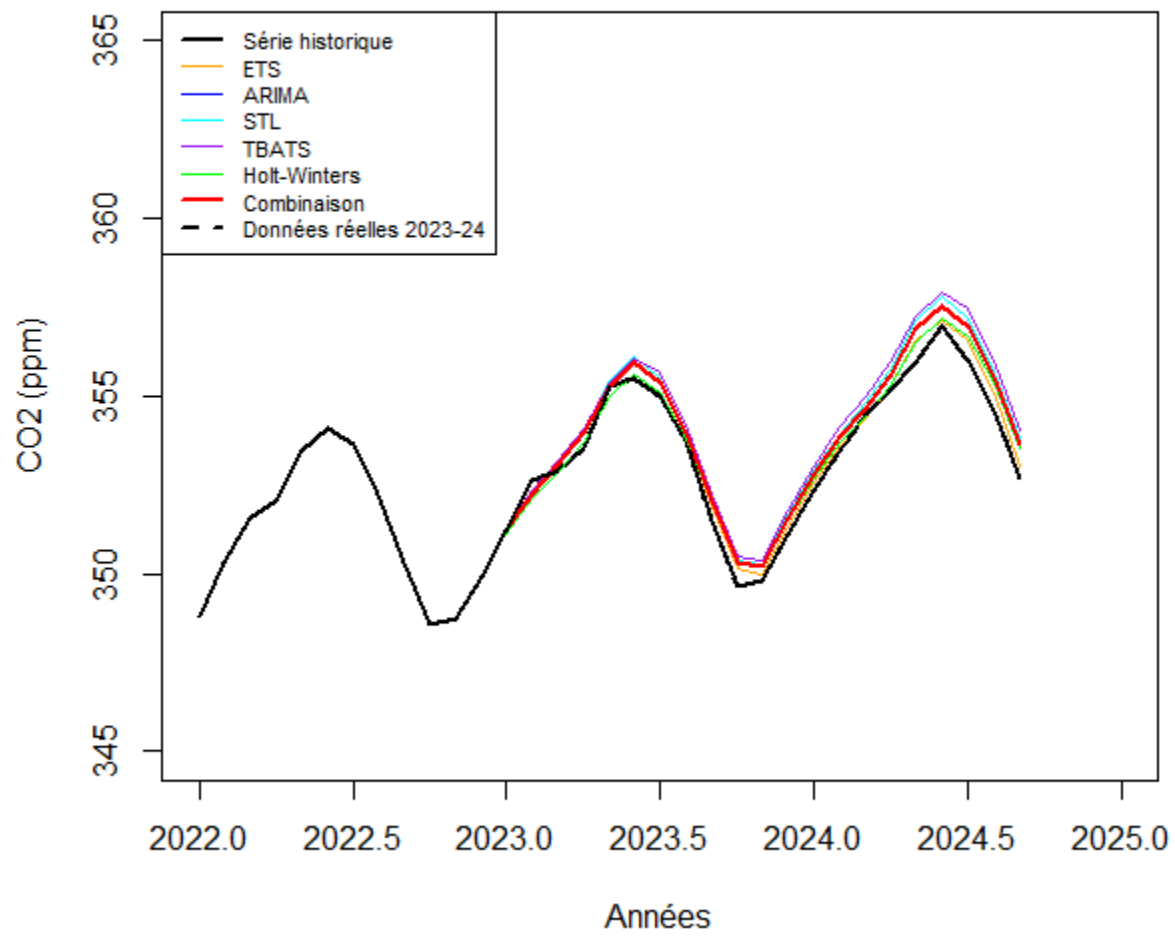
STL



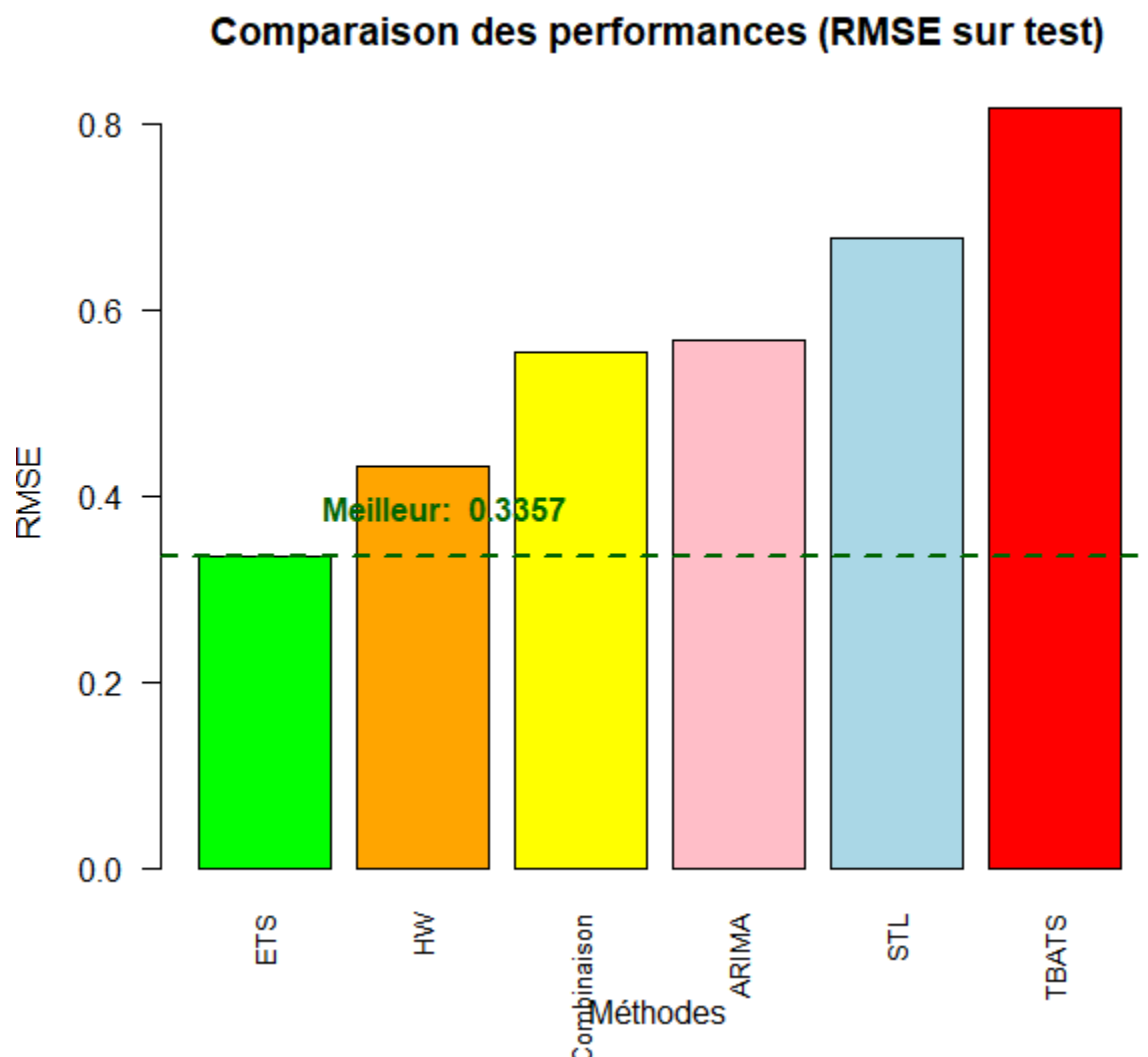
Graphique 17: Prédiction avec STL

Combinaisons

Comparaison des prévisions avec Combinaison - ZOOM 2022-2025



Graphique 18: Comparaison des prévisions avec combinaison



Graphique 19 : Comparaison de L'efficacité de chaque modèle sur la base du RMSE

Bilan

Le test révèle que ETS (RMSE test = 0.336) surpasse toutes les méthodes en prédiction, notamment ARIMA (0.568) et STL (0.678), malgré leur meilleur ajustement historique. STL s'effondre sur le test malgré son excellent ajustement d'entraînement (RMSE = 0.242) : cas classique de surapprentissage. Pour les prévisions 2024-2026, ETS devrait être privilégié en raison de sa capacité prédictive supérieure, bien que ARIMA conserve sa pertinence théorique (meilleur AIC, diagnostics validés).

Conclusion et Vulgarisation des résultats

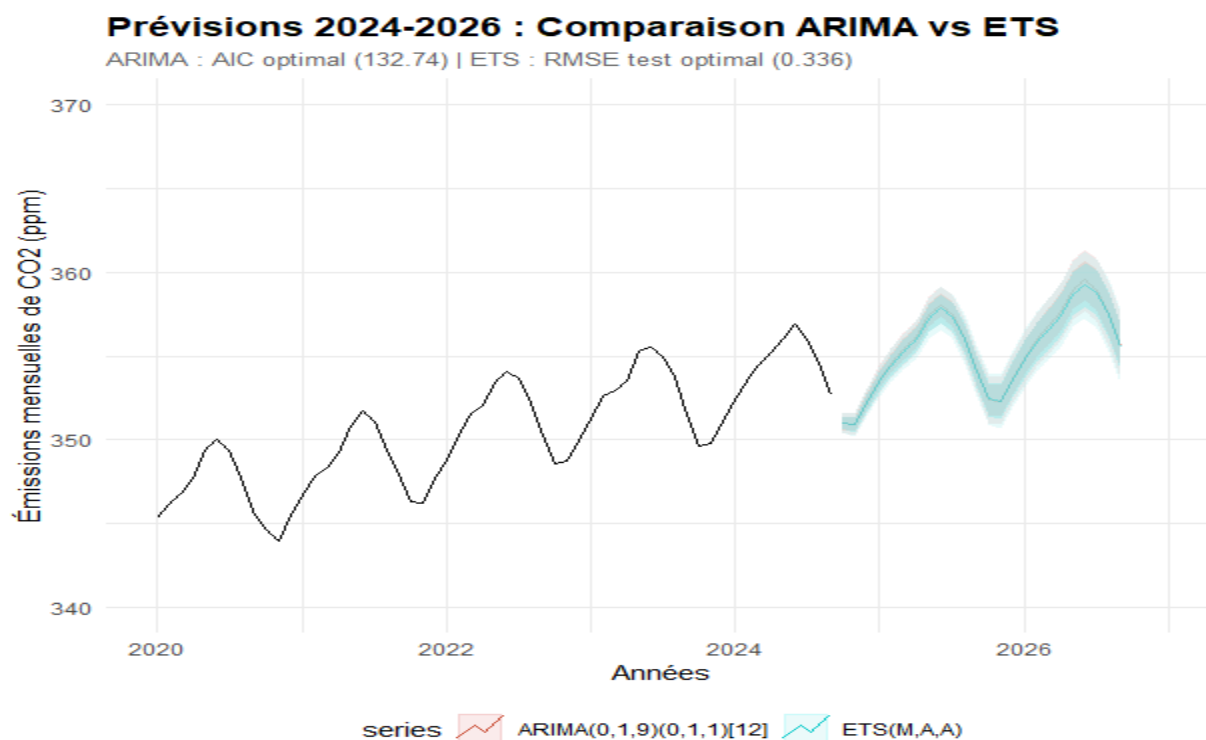
L'analyse statistique a identifié un modèle ARIMA optimal pour prédire les émissions futures de CO₂ du volcan Etna. Ce modèle, sélectionné selon une méthodologie rigoureuse (Box-Jenkins), prédit une poursuite de l'augmentation des émissions jusqu'en 2026.

Cependant, une validation croisée révèle une nuance importante : bien que ce modèle soit théoriquement optimal, une méthode alternative plus simple (ETS) s'avère 41% plus précise lorsqu'on teste les prévisions sur des données réelles récentes.

En pratique :

- Le modèle ARIMA reste valable pour comprendre les mécanismes d'émission
- Mais pour des prévisions opérationnelles concrètes, la méthode ETS serait plus fiable

C'est pourquoi les deux sont à prendre en compte et les deux sont présentées ici dans le graphique finale :



Conclusion brève

Les émissions de CO₂ de l'Etna montent régulièrement et devraient atteindre environ 360 ppm d'ici 2026, une prévision solide confirmée par deux méthodes différentes

