



UNIVERSITÉ DE NANTES



IAE NANTES
ÉCONOMIE & MANAGEMENT

Évolution de l'Indice des Prix d'Achat des Moyens de Production Agricole en France

Janvier 2005 – Décembre 2023

Master Économétrie et Statistiques, parcours Économétrie Appliquée

Techniques de prévision et conjoncture

Gloria Isabel PALACIO BARCO

AVRIL 2024

Sommaire

- 1 INTRODUCTION
- 2 ANALYSE PRELIMINAIRE
- 3 ESTIMATION DES MODELES LINEAIRES
- 4 PREVISION ET REPRESENTATION GRAPHIQUE
- 5 QUALITÉ DE PREVISION
- 6 PRECISION PREDICTIVE
- 7 PREVISION SUR UNE ANNEE AVEC UN PAS DE 1 MOIS
- 8 CONCLUSION
- 9 TABLE FIGURES ET TABLEAUX
- 10 BIBLIOGRAPHIE
- 11 ANNEXES

1 Introduction

L'Indice des Prix d'Achat des Moyens de Production Agricole (IPAMPA) joue un rôle très important dans l'économie agricole en offrant un aperçu des coûts auxquels les agriculteurs sont confrontés. En 2023, les agriculteurs européens en Pologne, Roumanie, Slovaquie, Hongrie, Bulgarie, ainsi qu'en France, ont été confrontés à des défis marqués par des conditions climatiques extrêmes et une forte concurrence due aux importations à bas prix en provenance de l'Ukraine, considérées comme une « concurrence déloyale »¹. Ces tensions surviennent après une année 2022 difficile, marquée par des réserves en eau basses et des perturbations climatiques qui ont fortement impacté les rendements agricoles.

Dans ce contexte, comprendre l'évolution de l'IPAMPA ne concerne pas uniquement les acteurs du secteur agricole. En effet, les variations de cet indice ont également des répercussions sur les prix à la consommation. Tout changement dans les coûts de production se répercute, après un certain délai, sur les prix finaux, influençant ainsi le coût de la vie générale.²

Face à cette complexité, notre projet s'est concentré sur la prévision de l'IPAMPA à l'aide de plusieurs modèles de prévision statistique. Nous avons exploré diverses approches pour identifier le modèle le plus performant en fonction des spécificités de nos données.

L'objectif principal de notre étude a été de confronter ces modèles aux données réelles afin d'évaluer leur capacité à prédire précisément les fluctuations de l'IPAMPA. Ce processus nous permet non seulement de comprendre les défis associés à la prévision de cet indice, mais aussi d'appréhender les difficultés à développer un outil prédictif fiable.

¹ (VERGEREAU. 2024)

² (Carré 2022)

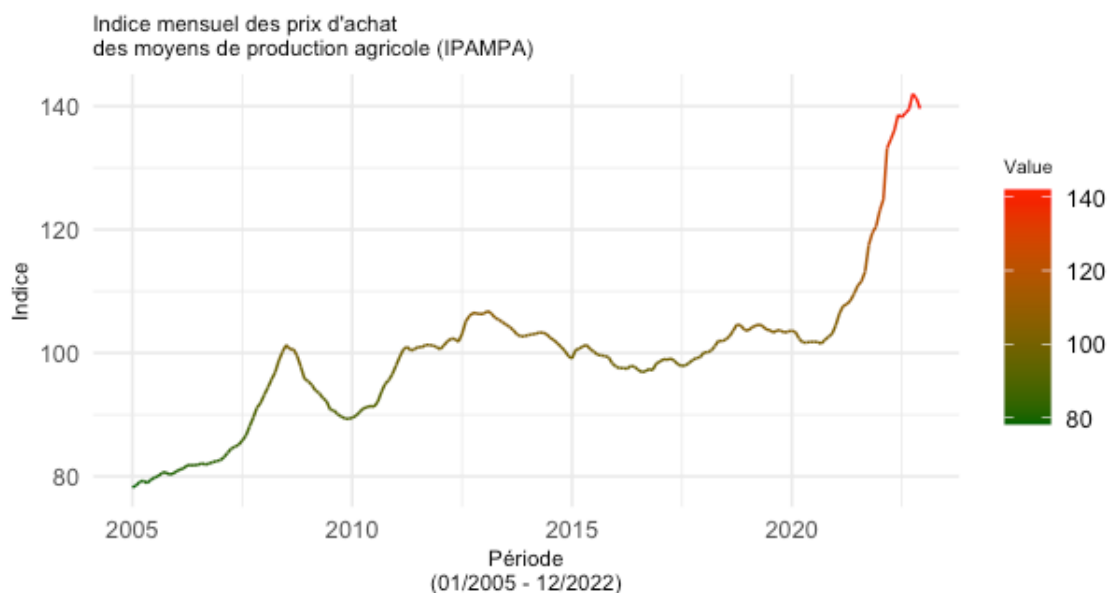
2 Analyse préliminaire

2.1 Présentation et caractérisation de la série étudiée

L'indice mensuel des prix d'achat des moyens de production agricole (IPAMPA) suit les fluctuations des prix des ressources que les agriculteurs achètent, incluant les matières premières et les équipements nécessaires à leur activité. Cet indice est développé conjointement par l'INSEE³ et le ministère de l'Agriculture. Il repose sur une enquête mensuelle effectuée auprès de 250 entreprises fournissant des biens et services aux agriculteurs, ainsi que sur des indices publics non spécifiques à l'agriculture pour les secteurs de l'énergie et de la construction.

La Figure 1 présente l'indice en France, couvrant la période de janvier 2005 à décembre 2022, avec 2015 comme année de base. Cela signifie que toute valeur supérieure à 100 indique une augmentation des coûts, tandis qu'une valeur inférieure à 100 indique une baisse des coûts par rapport à cette année de base.

Figure 1- Série initial



³ (Institut national de la statistique et des études économiques)

Sur le graphique, nous pouvons observer que l'indice semble relativement stable durant la période de 2011 à 2020, suivie par une forte tendance à la hausse. La baisse des coûts en 2008-2009 correspond à une période marquée par la mise en œuvre de politiques de relance économique à la suite de la crise de 2008. À partir de 2010, l'indice enregistre de légères hausses, maintenant cette tendance modérée jusqu'en 2020.

À compter de 2021, la France est confrontée à un épisode d'inflation exceptionnel, affectant particulièrement le secteur agricole. Initialement perçue comme une conséquence temporaire de la reprise économique post-COVID et de la flambée des prix du pétrole, l'inflation s'est révélée être plus complexe et durable. Elle s'est étendue à une large gamme de biens et services, avec d'importantes répercussions sur les coûts. L'escalade du conflit en Ukraine en 2022 a exacerbé la situation, entraînant une hausse significative des prix de l'énergie et des perturbations dans les marchés des matières premières agricoles.

En fournissant un aperçu mensuel, l'IPAMPA contribue à une meilleure compréhension des tendances à court terme, qui sont cruciales pour la gestion opérationnelle et financière des exploitations agricoles. Son utilisation dans le secteur agricole est donc importante, car cet indice non seulement informe les agriculteurs sur les coûts actuels et futurs, mais il sert aussi de base pour les politiques agricoles et les décisions économiques à plus grande échelle, offrant ainsi des données comparables à l'échelle européenne.

2.2 Détection des points atypiques

Nous allons procéder à la détection des points atypiques, car leur présence peut entraîner des problèmes d'estimation non seulement dans les méthodes de désaisonnalisation mais aussi dans les méthodes d'estimation classiques. À cette fin, nous utiliserons la fonction `tso()` de la bibliothèque `tsoutliers`, qui intègre la modélisation avec ARIMA. Cette approche nous aidera à distinguer les effets des outliers des fluctuations normales.

D'après les résultats affichés dans l'0, nous identifions quatre anomalies dans la série temporelle. Ces points atypiques se catégorisent en deux types : 'LS' pour Level Shift, indiquant un changement de niveau, et 'TC' pour Transitory Change, signalant un

changement transitoire. Ces anomalies sont détaillées dans la Tableau 1 :

Tableau 1 Points atypiques

Type	Date	Coefhat	Tstat
LS	2015/02	1.539	3.423
LS	2021/10	2.742	6.067
LS	2022/03	6.675	14.817
TC	2022/10	2.408	6.119

Ces points correspondent à des changements significatifs dans l'indice en raison de changements, que nous pouvons expliquer par :

LS 122 (Février 2015) : Conséquence de la réforme de la Politique Agricole Commune (PAC) de l'Union européenne, mise en œuvre en 2014.

LS 202 (Octobre 2021) : Cette anomalie pourrait être associée à la reprise économique après la pandémie de COVID-19. De nombreuses économies ont connu des pressions inflationnistes dues aux perturbations des chaînes d'approvisionnement et à l'augmentation de la demande à mesure que les confinements étaient levés.

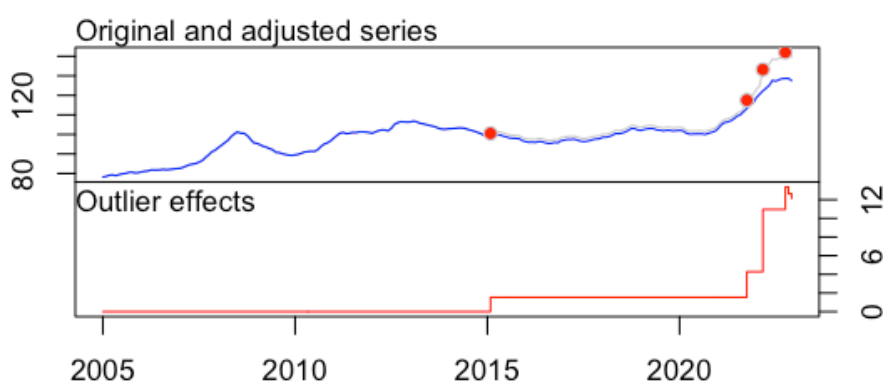
LS 207 (Mars 2022) : L'anomalie significative détectée en mars 2022 est liée à l'invasion russe en Ukraine. Cette année-là, le prix des intrants utilisés par les agriculteurs a augmenté de 25,9 %⁴, représentant la plus forte hausse depuis le début des années 2000. Le conflit a provoqué une flambée immédiate des prix mondiaux de l'énergie et généré une incertitude sur les marchés des produits agricoles, notamment parce que la Russie et l'Ukraine sont des fournisseurs clés d'intrants agricoles tels que les engrais et les céréales.

La Figure 2 illustre les points atypiques au sein de notre série temporelle et nous permet également de la comparer avec la série corrigée des effets de ces anomalies. La courbe ajustée présente la série temporelle une fois les effets estimés des points atypiques retirés, offrant ainsi un aperçu plus lisse et plus représentatif de la tendance de la série,

⁴ (Agreste 2023)

sans la distorsion provoquée par ces points atypiques.

Figure 2 Détection des points atypique



2.3 Vérification de la stationnarité de la série $I(0)$ ou $I(1)$

Avant de vérifier si notre série est stationnaire, nous examinons la présence éventuelle de saisonnalité et de tendance dans notre série. Il est essentiel de commencer par cette vérification, car la désaisonnalisation constitue une étape préliminaire importante dans l'analyse d'une série temporelle.

Tableau 2 Tests de saisonnalité

Webel-Ollech		Seasonal dummies	
<i>Stat</i>	<i>p-value</i>	<i>Stat</i>	<i>p-value</i>
0	11 0,03	0,86	0,576

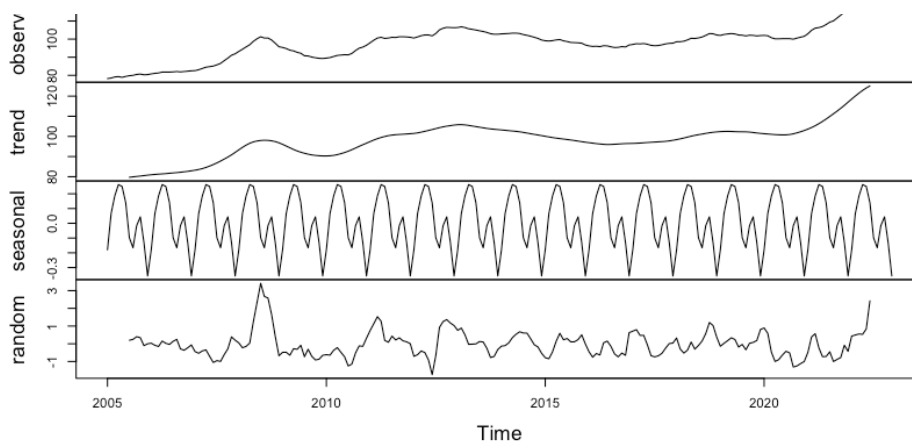
Les tests présentés dans le Tableau 2 sont basés sur l'hypothèse nulle H_0 qui postule l'absence de saisonnalité ; ces tests n'ont pas détecté de saisonnalité dans notre série, ce qui nous permet de ne pas rejeter l'hypothèse nulle et de conclure que les variations observées ne sont pas attribuables à un motif saisonnier.

Tableau 3 Test de tendance

Tau	P value
0,568	$2,22 \times 10^{-16}$

L'analyse de tendance réalisée avec le test de Mann-Kendall révèle, comme indiqué dans le Tableau 3, un tau (τ) de 0.568, suggérant une tendance croissante modérément forte. Avec une p value inférieure à 0.05, cela confirme la présence d'une tendance significative dans la série initiale. Nous pouvons observer ces composants de notre série dans le graphique de décomposition :

Tableau 4 Décomposition



Concernant le test de stationnarité, nous allons utiliser le test Augmented Dickey-Fuller (ADF), qui pose comme hypothèse nulle que la série temporelle possède une racine unitaire, indiquant qu'elle est non-stationnaire.

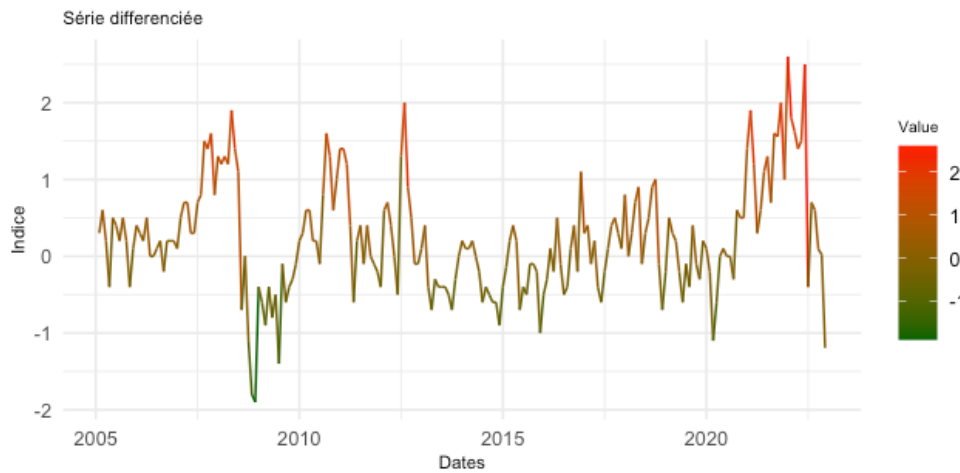
Tableau 5 Test de stationnarité

ADF test	
p-value	0,099

Le Tableau 5 indique une p-value supérieure à 0.05, donc nous ne pouvons pas rejeter l'hypothèse nulle de non-stationnaire .

La stationnarité est cruciale, car elle constitue une condition préalable à l'utilisation de nombreux modèles prédictifs, tels que ARIMA, qui nécessitent une moyenne et une variance constantes au fil du temps. Afin de rendre notre série stationnaire, nous procéderons à sa différenciation Figure 3. Cette technique ajuste la série en calculant les différences entre les observations consécutives, ce qui aide à éliminer les tendances ou les cycles, mais a pour conséquence la perte de la première observation des données.

Figure 3 Série Différenciée



Nous confirmons que notre série différenciée a bien été rendu stationnaire (Tableau 6), avec une p-value de 0,037, nous pouvons rejeter l'hypothèse nulle d'une racine unitaire.

Tableau 6 Test stationnarité série différencié

ADF test - Série différencié	
p-value	0,037

De plus, la visualisation de la stationnarité est renforcée par le corrélogramme de la Figure 4 , qui nous permet de visualiser cette stationnarité. Par ailleurs, le périodogramme de la série, également présenté (Figure 5), illustre l'évolution de la série dans le domaine des fréquences et met en évidence les pics spectraux, permettant ainsi de détecter d'éventuelles fréquences saisonnières.

Figure 4 Corrélogrammes

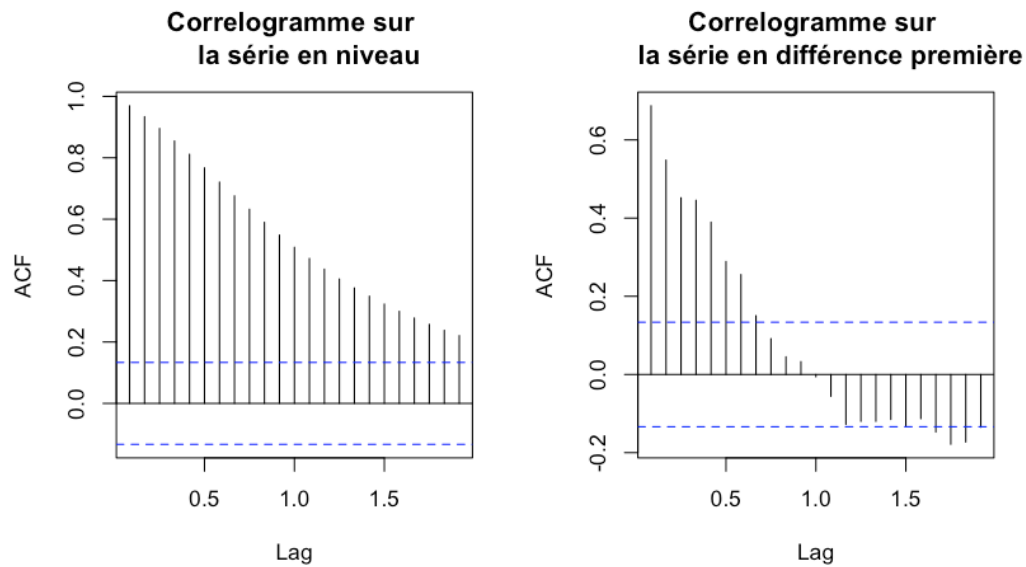
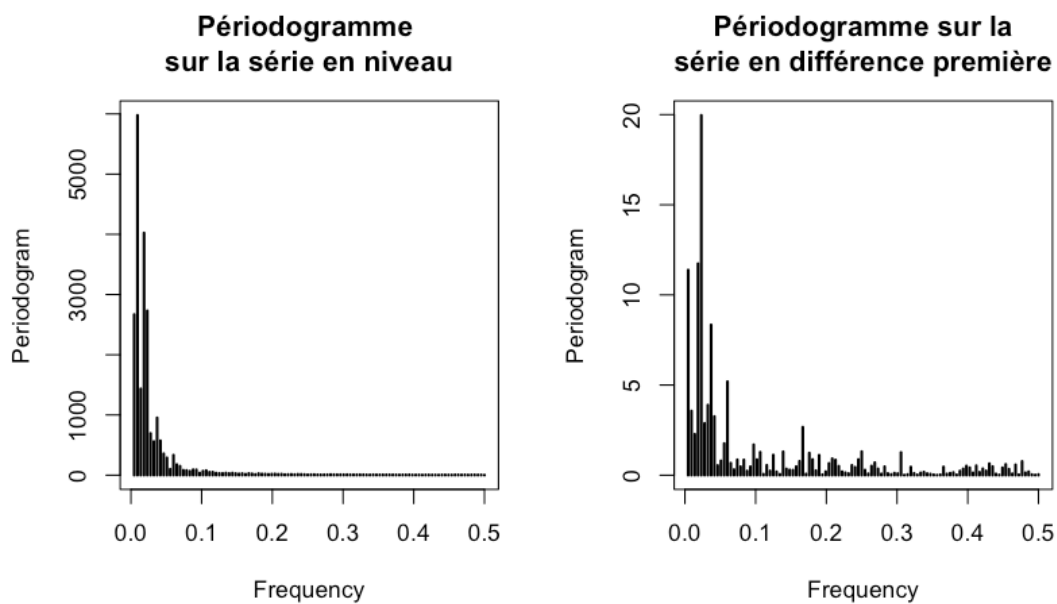


Figure 5 Périodogramme



2.4 Statistiques descriptives

Dans cette dernière partie de la section 2 , le Tableau 7 nous présente les statistiques descriptives de la série, ce qui nous permet, avant de passer à une nouvelle section, de clarifier les différences entre nos séries. La série initiale est celle au niveau, la série ajustée est corrigée des valeurs atypiques, et la série différenciée, qui est rendue stationnaire, qui sera utilisée pour la suite de notre analyse.

Tableau 7 Statistiques descriptives comparaison

Série	Différenciée	Ajusté	Initial
Minimum	-1,90	78,20	78,20
1^{er} Quartile	-0,25	93,65	93,65
Médiane	0,20	99,53	100,45
Moyenne	0,22	98,40	99,60
3eme quartile	0,60	102,52	103,33
Maximum	2,60	128,66	142,00

Nous pouvons observer une différence significative entre les valeurs de la série différenciée et la série ajustée, qui s'explique par la nature la différenciation. Lorsqu'on différencie une série, on calcule les écarts entre chaque point consécutif, ce qui reflète les changements d'une période à l'autre plutôt que les valeurs absolues. Ainsi, les valeurs de la série différenciée sont des écarts et non des niveaux de mesure absolus comme dans la série originale.

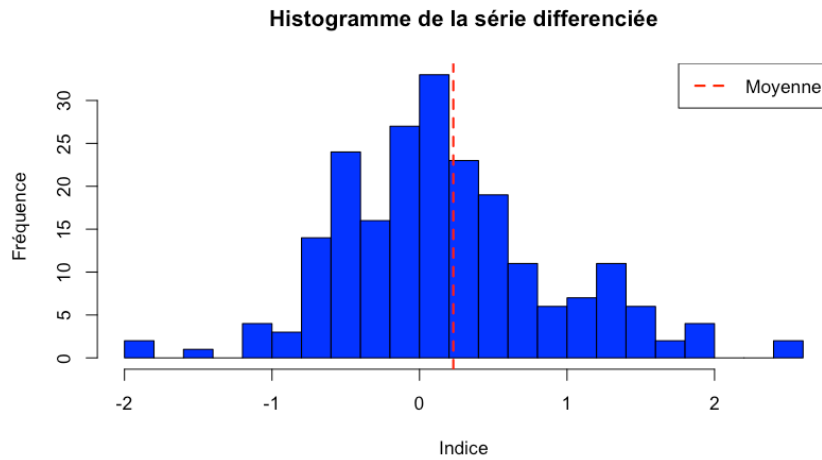
Les statistiques nous montrent une gamme de valeurs allant d'un minimum de -1.9 à un maximum de 2.6, ce qui indique une variation significative des changements d'une période à l'autre. La médiane est positive (0.2), suggérant que les augmentations sont plus fréquentes que les diminutions. Quant à la moyenne elle est aussi légèrement positive 0,22 voir Figure 6 .

Nous pouvons observer une distribution légèrement asymétrique, qui sera confirmé par la valeur de skewness Tableau 8. De plus, une p-value inférieure à 0,05 pour le test de normalité de Shapiro (voir Annexe 5) confirme l'écart de la série par rapport à une distribution normale.

Tableau 8 Skewness et Kurtosis

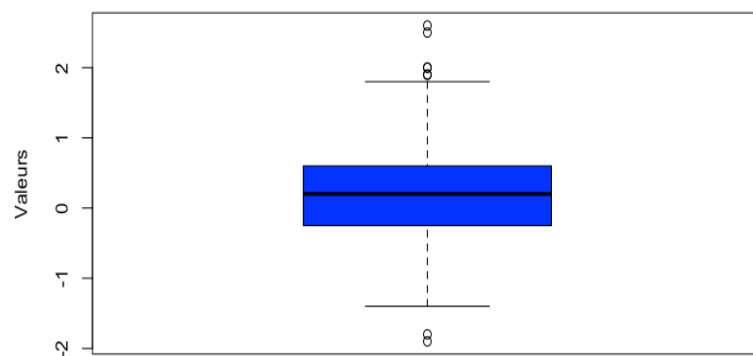
Skewness	Kurtosis
0,45	0,62

Figure 6 Histogramme



La kurtosis de 0,62 indique un aplatissement de la distribution, c'est-à-dire qu'elle présente moins de valeurs extrêmes que ce que l'on attendrait dans une distribution normale, avec une moindre concentration des observations dans les queues. Cela est cohérent avec le traitement préalable des outliers. Cependant, le boxplot de la Figure 7 révèle quelques possibles valeurs atypiques dans la série différenciée. Ces observations, amplifiées par la sensibilité aux variations, ne reflètent pas nécessairement des anomalies significatives dans la série de données originale.

Figure 7 Boxplot



3 Estimation des modèles linéaires

Dans cette section, nous analyserons les paramètres de différents modèles tels que les modèles AR(1), AR(p), et ARIMA(p,d,q), ainsi que des approches plus avancées telles que LED Holt-Winters, ADAM ETS, et d'autres. Nous présenterons un tableau comparatif pour examiner l'efficacité de chaque modèle en fonction de la dynamique de notre série temporelle. L'objectif sera de déterminer le modèle le plus performant en utilisant les critères AIC et AICc, afin de trouver l'équilibre optimal entre la précision de l'ajustement et la complexité du modèle

3.1 Estimation et choix des paramètres

3.1.1 Modèles AR (1), AR(p) et ARIMA(p,d,q)

Le premier modèle est l'AR(1), est un modèle autorégressif d'ordre 1, où la valeur actuelle de la série est principalement influencée par sa valeur immédiatement précédente, ajustée par un terme d'erreur. Nous examinerons également des modèles autorégressifs d'ordre supérieur, notés AR(p), qui étendent le principe de l'AR(1) en incorporant jusqu'à p valeurs antérieures dans la prédiction. Enfin, le modèle ARIMA(p,d,q) combine les techniques d'autorégression et de moyenne mobile, tout en intégrant un processus de différenciation pour stabiliser la série.

Estimation des paramètres

Pour l'estimation des paramètres de ces trois modèles, nous utiliserons la fonction `auto.arima` de la bibliothèque `forecast` en R. Nous préciserons que notre série est déjà stationnaire et non saisonnière, en configurant les options `stationary = TRUE` et `seasonal = FALSE`. Afin de choisir le modèle offrant le meilleur équilibre entre ajustement et complexité, nous opterons pour le critère d'information AIC ou AICc; nous sélectionnerons donc l'option `ic = "aic"` pour optimiser ce critère. Pour accélérer le processus de sélection des modèles, nous utiliserons la méthode `stepwise`, qui élimine efficacement les modèles moins performants.

Nous activerons également l'option `trace = TRUE` pour suivre en détail le processus de

sélection du modèle.

Pour chaque type de modèle spécifique, comme l'AR(1), nous configurerons `auto.arima` pour restreindre la recherche aux paramètres appropriés, par exemple, `max.p=1`, `max.q=0`, et `d=0` pour un modèle strictement AR(1), où `p=1` indique un unique retard autorégressif.

Ces ajustements ont révélé que le meilleur modèle pour la série AR(p) est un ARIMA(2,0,0), intégrant deux termes autorégressifs. Pour le modèle ARIMA(p,q,d), le meilleur modèle identifié est un ARIMA(1,0,1), également avec une moyenne nulle, qui combine un terme autorégressif et un terme de moyenne mobile. Les estimations de ces modèles, ainsi que les informations sur les résidus, sont disponibles dans Annexe 6.

3.1.2 Modèle LED Holt-Winters non saisonnière

La méthode LED (Lissage Exponentiel Double) de Holt-Winters, ajuste deux paramètres essentiels : le niveau de la série (paramètre alpha) et la tendance (paramètre bêta). Un alpha élevé permet au modèle de s'adapter rapidement aux valeurs observées récemment, tandis qu'un bêta faible modère les fluctuations de tendance, favorisant une vision à long terme dans les prévisions. Le réglage optimal de ces paramètres est crucial pour améliorer la précision des prévisions.⁵

Pour mettre en œuvre cette méthode, nous employons la fonction `HoltWinters` de R, configurée pour un lissage exponentiel double sans ajustement saisonnier (`gamma = FALSE`). Dans les résultats du modèle (voir Annexe 7), alpha de 0.653 indique une adaptation active du niveau aux données récentes, réagissant rapidement aux nouvelles informations. En revanche, un bêta de 0.050 signifie que le modèle ajuste peu la tendance face aux variations récentes, privilégiant ainsi une perspective à long terme et une stabilité dans les prévisions de tendance.

Cependant, une limitation de cette approche avec `HoltWinters()` est l'absence de critères d'information tels que l'AIC ou l'AICc, essentiels pour évaluer et comparer la qualité du

⁵ (Darné 2024)

modèle. Pour pallier cette contrainte, nous explorons une alternative avec le package `forecast` : La fonction `holt()`, conçue pour le lissage exponentiel double de Holt sans composante saisonnière. Cette méthode offre l'avantage de fournir automatiquement l'AIC, facilitant ainsi l'évaluation comparative des modèles.

3.1.3 Modèle ADAM ETS

Le modèle ADAM ETS constitue une évolution du modèle ETS classique, permettant d'incorporer une variété de distributions pour le terme d'erreur, afin d'élargir l'éventail des comportements de données que le modèle peut capter.

Nous mettons en œuvre ce modèle via la fonction `auto.adam()` du package `smooth`, initialement, nous avons ajusté un modèle ADAM ETS en spécifiant `model="ZZN"` pour permettre une sélection automatique des composantes du modèle les plus adaptées. Nous avons également inclus `lags=c(1,12)` pour tenir compte de la saisonnalité potentielle dans les données mensuelles et activé `select=TRUE` avec le critère d'information AIC pour optimiser notre choix de modèle.

Cependant, lors de cette première estimation, des messages d'avertissement ont été générés. Le premier indiquait que les données n'étant pas strictement positives, certaines distributions, à savoir la log-normale, la gaussienne inverse et la gamma, n'étaient pas pertinentes et ont donc été écartées. Le deuxième avertissement nous a informés que seuls les modèles additifs étaient appropriés pour notre série, ce qui a automatiquement ajusté notre processus de sélection de modèle.

En réponse à ces avertissements, nous avons réestimé le modèle, cette fois en spécifiant explicitement `model="AAN"`, et en sélectionnant `distribution="dnorm"`. L'ajout de l'option `bootstrap=TRUE` avait pour but de renforcer la robustesse de l'estimation par bootstrap, et le `control=list(maxit=1000, trace=TRUE)` de suivre le processus d'optimisation et d'augmenter le nombre d'itérations maximales pour assurer la convergence.

Malgré ces ajustements, un nouvel avertissement a été émis concernant la maximisation de la vraisemblance. Cela peut indiquer des problèmes de convergence dans l'estimation des paramètres du modèle.

Nous avons donc testé un troisième modèle, en utilisant comme option `model = ANN`, si cet ajustement a résolu le problème de convergence, il a résulté en des prévisions identiques pour chaque mois, reflétant une incapacité à détecter des tendances ou des saisons dans nos données. En conséquence, nous avons décidé de retenir le deuxième modèle. Cela implique une certaine prudence lors de l'interprétation des résultats. Annexe 8.

3.1.4 Modèle ADAM ETS SARIMA

Le modèle ADAM ETS SARIMA fusionne les éléments des modèles ETS et SARIMA, adapté aux séries temporelles avec tendances non linéaires, cycles saisonniers et dépendances autorégressives.

Initialement, nous avons employé la fonction `auto.adam()` avec `model="ZZN"` pour une sélection automatique des composantes, incluant des retards spécifiques et des ordres ARIMA. Cependant, des avertissements ont indiqué l'inadaptation de certaines distributions et la nécessité d'utiliser uniquement des modèles additifs en raison de données non strictement positives.

En réponse, nous avons ajusté notre modèle à `model="AAN"` avec des distributions `"dnorm"` et `"dlaplace"`, et intégré le bootstrap pour renforcer la robustesse de l'estimation. Malgré ces ajustements, un avertissement concernant la maximisation de la vraisemblance a persisté, indiquant des problèmes de convergence. Nous avons finalement opté pour un modèle simplifié `model="ANN"`, toujours avec bootstrap et contrôle accru des itérations, mais avons continué à rencontrer des difficultés, notamment des prévisions homogènes indiquant un manque de capture des tendances et saisons. Cela nous a mené à retenir le modèle "AAN", prenant en compte ces limitations dans notre analyse. Annexe 9

3.1.5 Modèle SSARIMA

Le modèle SSARIMA (State-Space ARIMA ou Several Seasonalities ARIMA) est formulé dans un cadre d'espace d'état, il permet une modélisation qui peut capturer plusieurs saisonnalités et d'autres dynamiques complexes en transformant un modèle ARIMA traditionnel en un système d'équations d'espace d'état, rendant le modèle plus adaptable.

Pour l'implémentation, nous utilisons la fonction `auto.ssaarima` du package `smooth`, sélectionnant l'AICc pour optimiser le modèle, avec un maximum de trois ordres pour les composants autorégressifs (AR) et les moyennes mobiles (MA).

Le modèle estimé par cette approche est un ARIMA(3,0,3). Cela signifie que le modèle inclut trois retards tant pour les termes AR que MA, sans différenciation ($I=0$), car la série était déjà stationnaire. Annexe 10.

3.1.6 Modèle CES

Le modèle CES (Composite Exponential Smoothing) est conçu pour les séries temporelles présentant des dynamiques complexes, intégrant des composantes telles que la tendance et la saisonnalité de manière non linéaire. Nous avons utilisé la fonction `auto.ces` du package `smooth` pour explorer diverses configurations de modèle : "n" pour aucune saisonnalité, "s" pour une saisonnalité simple, "p" pour une saisonnalité partielle équivalente à l'additive, et "f" pour une saisonnalité complète qui peut être à la fois additive et multiplicative.

Pour notre série, le modèle estimé est "CES(f)", signifiant que le modèle a adopté une configuration avec des composantes saisonnières complexes. Concernant les paramètres A et B d'après la documentation, le modèle CES est très sensible et il est donc conseillé de les laisser telles quelles. ⁶

⁶ (I. K. Svetunkov 2015)

3.2 Comparaison des modèles

Tableau 9 Comparaison de modèles

Modèles	AR(1)	AR(p)= ARIMA(2,0,0)	ARIMA (p,d,q) ARIMA(1,0,1)	Holt Winters	ADAM ETS= ETS(AAN)	ADAM ETS ARIMA= ETS(ANN)+ AIMA(3,0,0)	SSARIMA = ARIMA(3,0,3)	CES= (f)
<i>Alpha</i>				0,57	0,585	0,0009		
<i>Beta</i>				1e-04	0,00			
<i>Gamma</i>								
<i>Phi</i>						0,57/0,07 /0,09		
<i>Lambda</i>								
<i>Ar 1</i>	0,696	0,614	0,843			-0,005	0,17	
<i>AR 2</i>		0,148				-0,002	0,02	
<i>AR 3</i>						-0,14	0,54	
<i>SAR 1</i>								
<i>SAR 2</i>								
<i>SAR 3</i>								
<i>MA (1)</i>			-0,269				0,44	
<i>MA (2)</i>							0,36	
<i>MA 3</i>							-0,36	
<i>SMA 1</i>								
<i>SMA 2</i>								
Distribution	Normale	Normale	Normale	-	Normale	Laplace	Normale	-
AIC	341,3	339,6	338.2	900.04	354,80	338,98	336,03	327,14
AICc	341,41	339,71	338.31	900.33	355,09	339,86	336,57	327,43

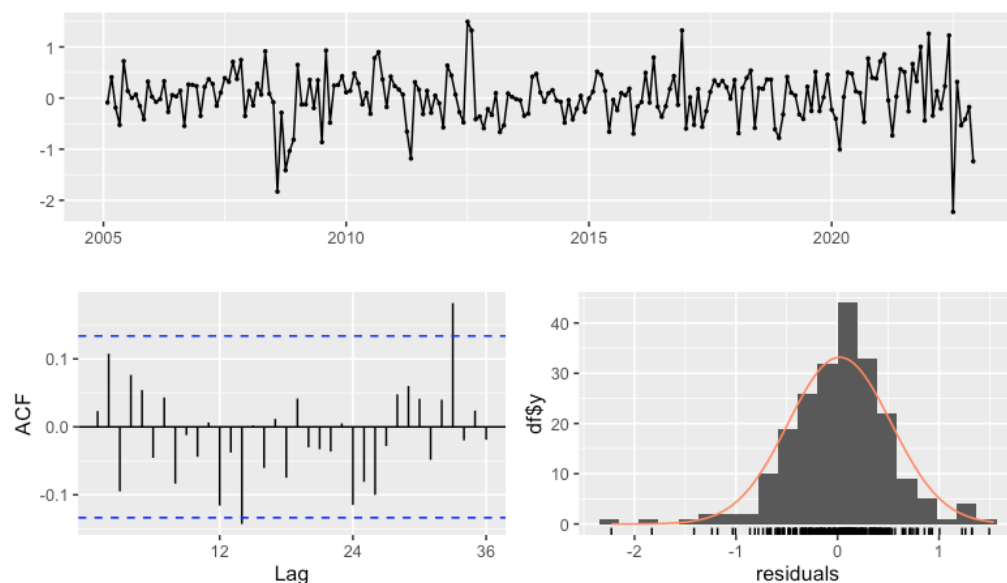
D'après le Tableau 9, le modèle avec la valeur AIC la plus basse est le modèle CES avec un AIC de 327,14. Le critère d'information d'Akaike (AIC) nous permet de comparer les modèles : plus la valeur de l'AIC est faible, meilleur est le modèle en termes de compromis entre la complexité et la capacité à bien ajuster les données. Donc, selon ce critère, le modèle CES ("f") serait le meilleur choix parmi ceux présentés.

Les paramètres du modèle sont :

- a_0 (1.465) qui représente la partie réelle du paramètre de lissage complexe pour la tendance de la série. Une valeur élevée pourrait indiquer une réactivité importante du modèle aux changements de tendance. ia_1 (0.9264i) est la partie imaginaire qui permet de capturer les cycles ou les oscillations dans la tendance.
- b_0 (0.995) est la partie réelle du paramètre de lissage pour la composante saisonnière. ib_1 (0.9989i) est la partie imaginaire.

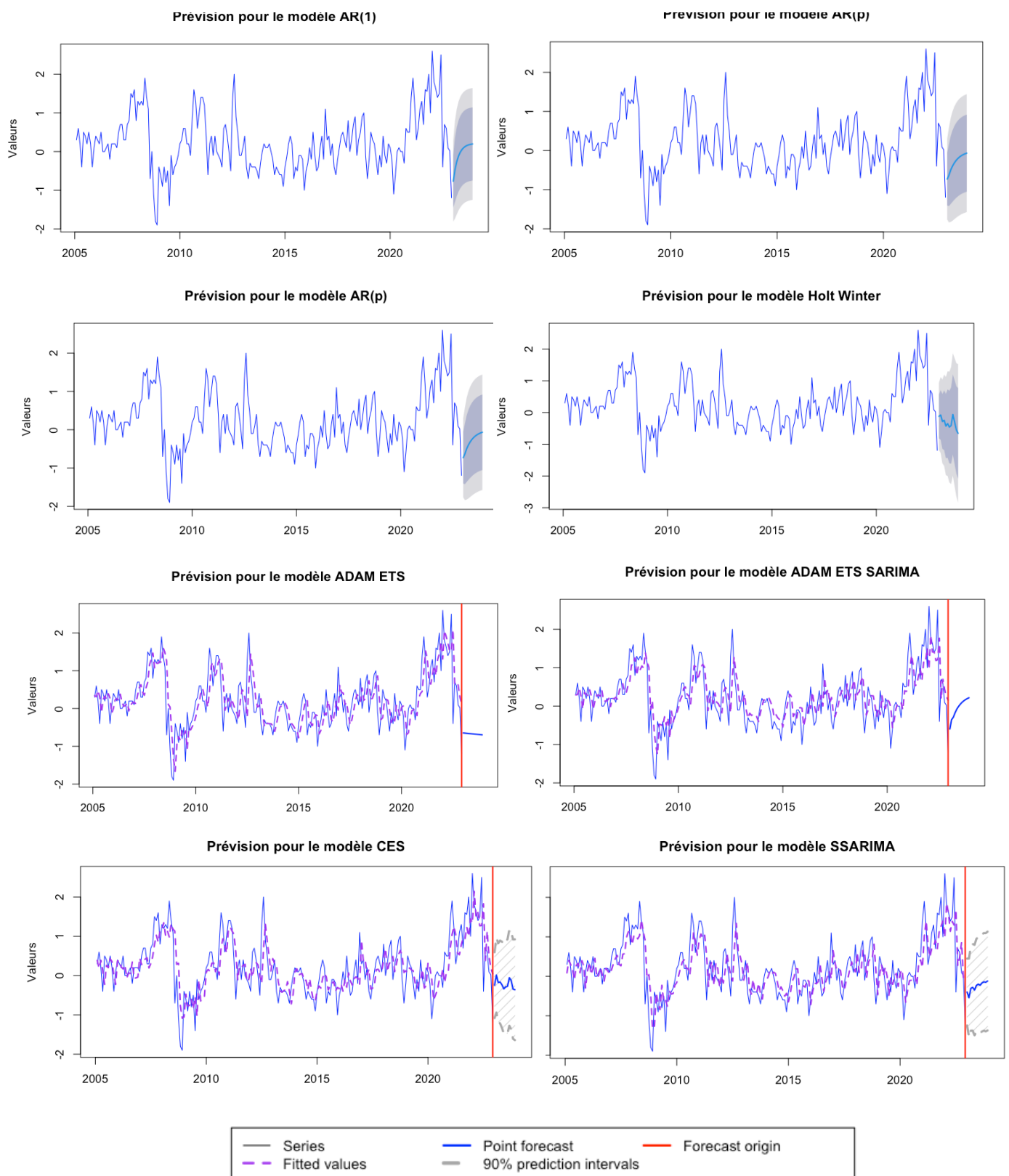
L'analyse des résidus montre que le modèle CES semble avoir bien capturé l'information des données. Les résidus se comportent comme du bruit blanc, ce qui est confirmé par le test de Ljung-Box qui donne une p-value de 0,440, indiquant l'absence d'autocorrélation significative (Annexe 12).

Figure 8 Analyse de résidus



4 Prédiction et représentation graphique

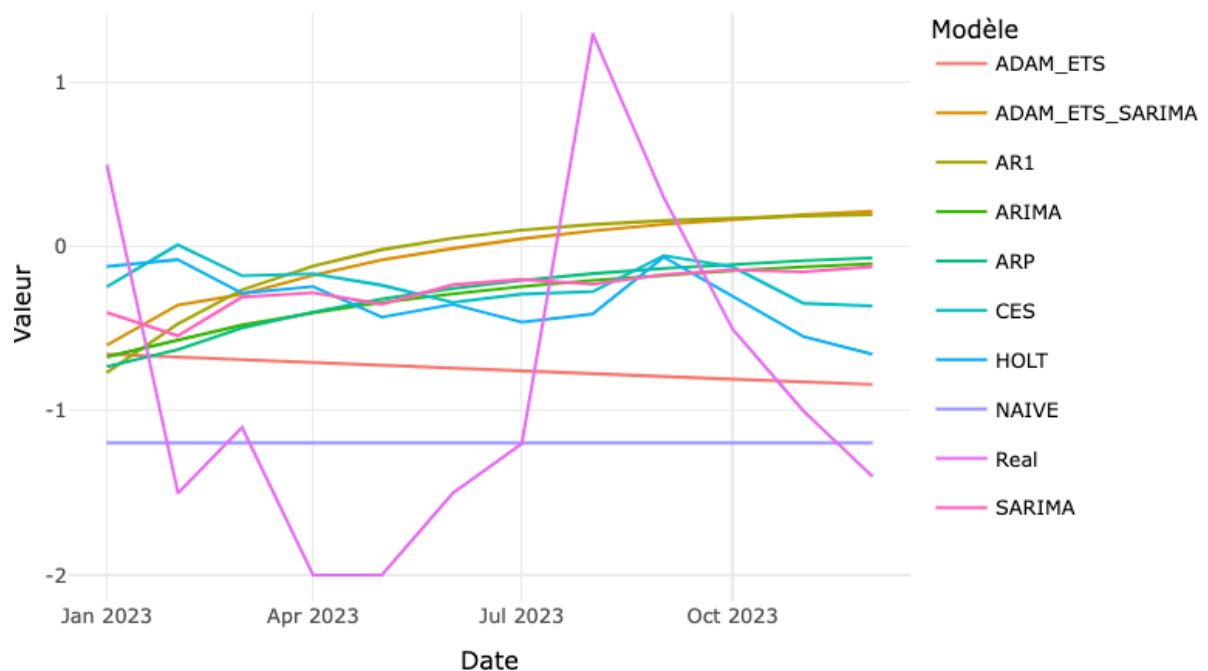
Figure 9 Prévisions



La Figure 9 Prévisions, représente les prévisions produites par les différents modèles estimés dans la section précédente, ces projections sont le résultat de l'utilisation de la fonction `forecast()` pour une durée prévisionnelle de 12 mois. Les intervalles de confiance à 90% mettent en évidence le degré d'incertitude entourant chaque estimation ponctuelle produite par les modèles.

La Figure 10 nous permet de comparer l'ensemble de prévisions pour l'année 2023, de janvier à décembre. Cette comparaison graphique vise à évaluer la justesse des modèles et leur capacité à se rapprocher des données effectives. Bien que globalement, nos modèles ne montrent pas une bonne concordance avec les données réelles, cette analyse reste utile pour apprécier la relative efficacité de chaque modèle et déterminer ceux qui correspondent le mieux aux observations effectives.

Figure 10 Comparaison des prévisions – Données différenciées
01/2023 à 12/2023



5 Qualité de prévision

Nous utilisons trois indicateurs principaux pour évaluer la qualité des prévisions précédentes : Le Mean Squared Error (MSE) qui permet de quantifier la précision des prédictions en calculant la moyenne des carrés des différences entre les valeurs prédites et les valeurs réelles. Un MSE plus faible indique une meilleure précision du modèle.

Parallèlement, le R-carré hors échantillon (R2OOS) fournit un aperçu de la capacité du modèle à généraliser aux nouvelles données. Un R2OOS élevé indique que les prédictions du modèle surpassent celles d'un modèle de référence, dans notre cas le modèle naïf, démontrant ainsi sa supériorité pour les prévisions hors échantillon.

Enfin, la Somme Cumulative des Erreurs de Prévision (CSPE) évalue l'accumulation des erreurs de prévision en sommant séquentiellement les différences entre les prédictions et les observations réelles. Un CSPE proche de zéro indique que le modèle maintient une bonne précision sur le long terme, tandis qu'une valeur élevée peut signaler une dégradation des prédictions au fil du temps

Tableau 10 MSE - R2OOS

Modèle	MSE	R2OOS
AR(1)	1,729	-0.551
AR((p)	1,308	-0.172
ARIMA	1,283	-0.151
Holt Winter	1,197	-0.073
ADAM ETS	1,017	0.087
ADAM ETS SARIMA	1,662	-0.490
SSARIMA	1,308	-0.173
CES	1,386	-0.242

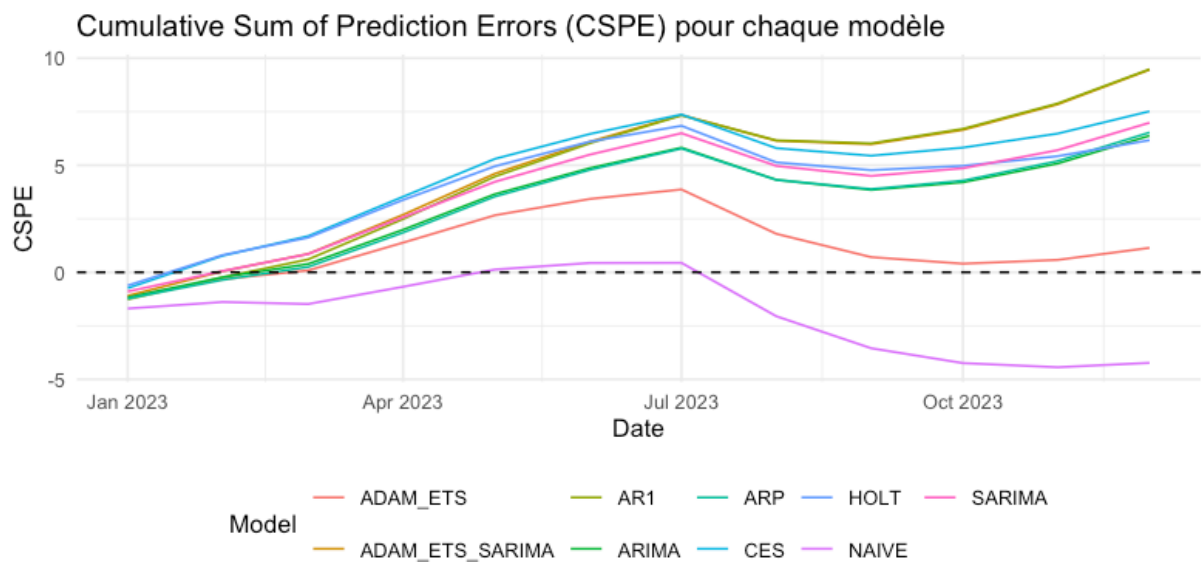
D'après Tableau 10 le modèle ADAM ETS s'est distingué par sa performance supérieure, affichant le MSE le plus bas à 1.017. En revanche, le modèle AR1, avec le MSE le plus élevé, produit des prévisions moins précises que les autres modèles. Malgré la combinaison de lissage exponentiel et d'ARIMA dans le modèle ADAM ETS SARIMA, celui-ci n'a pas performé aussi bien que le ADAM ETS, suggérant que l'ajout de la composante SARIMA n'a pas amélioré les prévisions pour ces données.

Quant au modèle CES, son MSE de 1.385 le place au-dessus de la moyenne en termes d'erreur parmi les modèles évalués. Du côté du R2OOS, le modèle ADAM ETS se distingue encore une fois comme le seul avec un R2OOS positif, ce qui démontre sa capacité à surpasser le modèle naïf et à s'adapter aux données pour généraliser efficacement ses prévisions hors échantillon.

En revanche, les modèles AR1, ARP, ARIMA, ADAM ETS SARIMA, SARIMA et CES, avec leurs R2OOS négatifs, révèlent une capacité prédictive inférieure à celle du modèle naïf.

Concernant le CSPE Figure 11 , le ADAM ETS montre une tendance à se rapprocher de zéro, ce qui indique une précision prédictive supérieure avec des valeurs atteignant un 1.14 en décembre. En revanche, les modèles AR1 et ADAM ETS SARIMA montrent des écarts plus significatifs, en particulier, le modèle AR1 atteint une erreur cumulative de 9.48 en décembre, la plus élevée parmi tous les modèles évalués, soulignant des déficiences notables dans sa capacité prédictive.

Figure 11 CSPE



6 Précision prédictive

6.1 Test de Diebold-Mariano (DM)

Le test DM examine les différences entre les erreurs de prévision de deux modèles. Il teste l'hypothèse nulle selon laquelle il n'y a pas de différence significative dans les performances de prévision entre les modèles sur l'ensemble des données testées.⁷

Le test utilise une statistique standardisée qui suit, sous l'hypothèse nulle, une distribution normale. Une valeur négative indique que le premier modèle (le modèle de référence, ici le modèle naïf) a de meilleures performances que le second modèle. Inversement, une valeur positive indique que le second modèle surpasse le premier.

Une p value faible (inférieure à 0.05) permet de rejeter l'hypothèse nulle, indiquant que les performances de prévision des deux modèles sont significativement différentes.

Tableau 11 DM Test

Model	DM_statistic	P_value
AR(1)	-0,891	0,195
AR(p)	-0,358	0,363
ARIMA	-0,319	0,377
HOLT	-0,154	0,440
ADAM_ETS	0,326	0,624
ADAM_ETS_SARIMA	-0,806	0,218
SSARIMA	-0,350	0,366
CES	-0,460	0,327

Nous obtenons les résultats du Tableau 11 en précisant l'option « less » qui test l'hypothèse que les prédictions du modèle testé sont moins précises que celles du modèle

⁷ (G. Colletaz 2007)

naïf.

Les modèles AR(1), AR(p), ARIMA, Holt Winter, ADAM_ETS_SARIMA, SSARIMA, CES, ont tous des statistiques DM négatives, indiquant qu'ils pourraient être moins précis que le modèle naïf. Cependant, les p values sont toutes supérieures à 0.05, ce qui signifie que les différences de précision ne sont pas statistiquement significatives. En d'autres termes, il n'y a pas de preuve suffisante pour affirmer que ces modèles sont moins précis que le modèle naïf de manière significative.

Le modèle ADAM_ETS est le seul à avoir une statistique DM positive, suggérant qu'il pourrait être plus précis que le modèle naïf. Néanmoins, la valeur p de 0.624 indique que cette différence n'est pas statistiquement significative. Donc, bien qu'il semble avoir de meilleures performances, on ne peut pas conclure définitivement qu'il est supérieur au modèle naïf en termes de précision des prévisions.

6.2 Accuracy

Au vu des résultats peu concluants, nous allons recourir à diverses mesures de précision des prédictions via la fonction `accuracy()`. Cette approche nous permet d'évaluer les performances des modèles de manière plus détaillée. Nous analyserons :

ME (Mean Error) : La moyenne des erreurs résiduelles qui nous renseigne sur la tendance générale des prédictions du modèle. Des valeurs proches de zéro reflètent une bonne précision, où le modèle prédit les valeurs avec peu de biais.

RMSE (Root Mean Squared Error) : Cette mesure, qui est la racine carrée de la moyenne des carrés des erreurs, pénalise plus fortement les grandes erreurs.

MAE (Mean Absolute Error) : Le MAE mesure la moyenne des valeurs absolues des erreurs, fournissant une perspective directe de l'erreur moyenne sans considération de sa direction.

Tableau 12 Accuracy

Model	ME	RMSE	MAE
AR1	-0.790	1.315	1.218
ARP	-0.544	1.143	1.064
ARIMA	-0.531	1.132	1.055
HOLT Winter	-0.514	1.094	0.962
ADAM_ETTS	-0.132	1.009	0.875
ADAM_ETTS_SARIMA	-0.7880	1.289	1.199
SSARIMA	-0.582	1.143	1.064
CES	-0.626	1.177	1.071

En examinant les valeurs des métriques du Tableau 12, nous pouvons observer que:

Le modèle ADAM ETS affiche la meilleure performance globale avec le RMSE le plus bas à 1.009 et le MAE le plus bas à 0.875. Le ME proche de zéro indique une précision et un biais minimisé dans les prévisions.

Le modèle Holt Winter suit en performance, avec un RMSE de 1.094 et le ME le à 0.962.

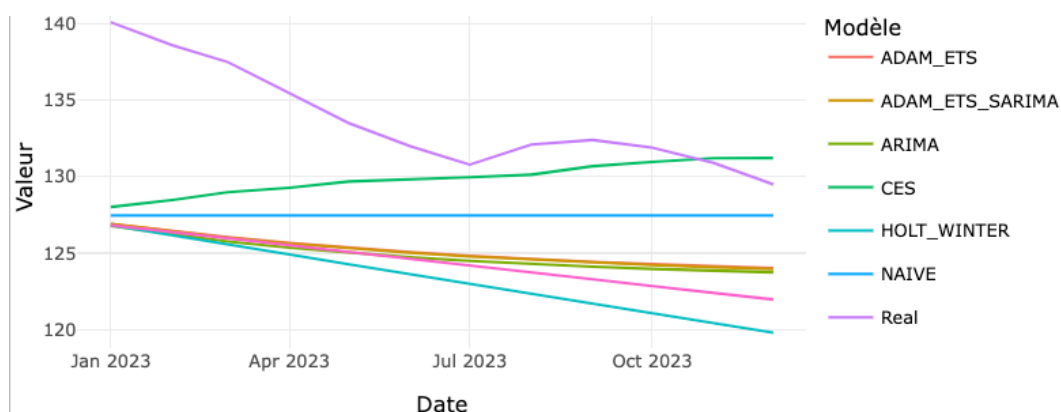
Ces résultats nous amènent à penser que le modèle ADAM_ETTS s'adapte le mieux à nos données, car il présente de meilleures performances. Cependant, cela ne signifie pas nécessairement qu'il capture les dynamiques des données. En observant le graphique, on remarque que son comportement est similaire à celui d'un modèle naïf, c'est-à-dire qu'il tend à produire une ligne droite et horizontale.

Il se pourrait que l'utilisation systématique de la série différenciée pour tous les modèles ne soit pas l'approche la plus judicieuse. En effet, pour des modèles tels que ADAM ETS et ADAM ETS ARIMA, qui sont conçus pour s'ajuster à diverses tendances et saisonnalités, obtenir une prévision linéaire peut indiquer une incapacité du modèle à saisir la dynamique des données, limitant ainsi sa capacité à anticiper les variations futures de tendance ou de niveau. Par ailleurs, il est envisageable que certains modèles, notamment ceux basés sur le lissage exponentiel, offrent de meilleures prévisions lorsqu'ils sont appliqués aux données brutes, particulièrement lorsque la non-stationnarité des données est attribuable à une tendance déterministe plutôt qu'à une racine unitaire.

6.3 Réestimation des modèles

Nous allons procéder à une réestimation des modèles Holt-Winter, ADAM ETS, ADAM ETS, SSARIMA et CES en utilisant des données non différenciées, mais corrigées des valeurs atypiques, afin de vérifier si les modèles s'adaptent mieux aux données initiales.

Figure 12 Comparaison des prévisions - Données réelles
01/2023 à 12/2023

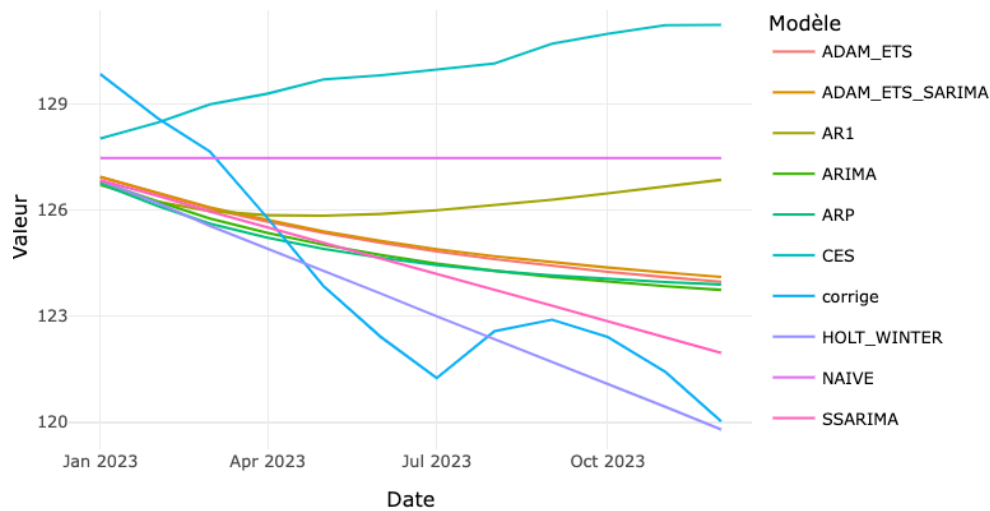


Après cette nouvelle estimation, l'analyse de la Figure 12, montre que toutes les prévisions des différents modèles tendent à sous-estimer les valeurs réelles.

Cette sous-estimation pourrait être attribuée à l'utilisation des données réelles sans ajustement pour les outliers. En effet, lors de la détection des valeurs atypiques, nous avons identifié trois valeurs particulières vers la fin de la période étudiée. Nos modèles ont ensuite été estimés à partir de cette série ajustée pour tenir compte de ces anomalies.

Il serait donc pertinent de comparer les prévisions avec les valeurs de l'année 2023 après correction, donc sans la présence des outliers. Cette démarche permettrait une meilleure évaluation de la performance des modèles prévisionnels, en les alignant avec des données dépourvues de perturbations exceptionnelles.

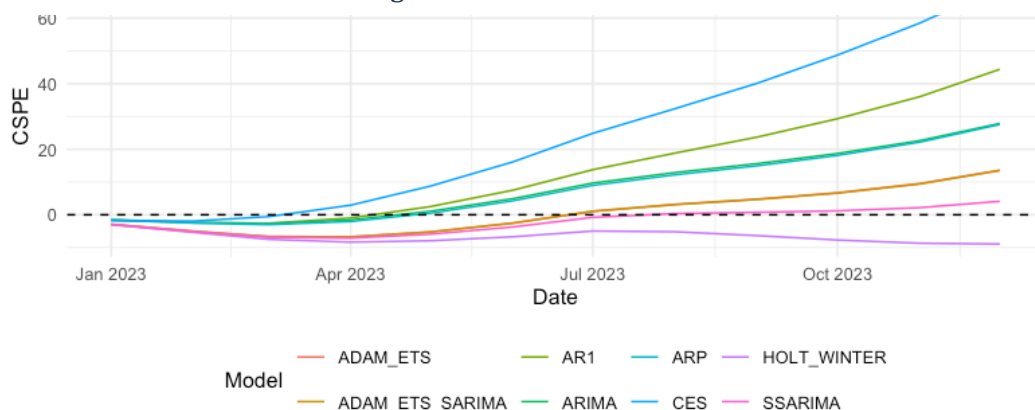
Figure 13 Comparaison des prévisions – Données réelles corrigées
01/2023 à 12/2023



Dans la Figure 13, nous présentons l'ensemble des modèles, y compris AR(1), AR(p) et ARIMA, pour lesquels nous avons conservé les prévisions réalisées à partir des données stationnaires. Nous avons ensuite réintégré ces prévisions aux données originales pour obtenir les résultats au niveau global. Nous pouvons observer que nos modèles reproduisent bien la tendance des valeurs réelles corrigées, et certains modèles se rapprochent davantage des données réelles que d'autres.

Grâce à la Figure 14, il est possible d'observer que des modèles tels que Holt-Winter ou SSARIMA affichent un CSPE, proche de 0, mais aussi principalement négatif, indiquant une surestimation initiale des valeurs réelles qui diminue progressivement au fil du temps. Par ailleurs, le modèle CES montre des limitations dans sa capacité à s'adapter aux tendances, comme en témoigne la variation significative de son CSPE.

Figure 14 - CSPE -Réestimation



Afin de sélectionner le meilleur modèle, nous allons à nouveau vérifier la qualité des prévisions (Tableau 13).

Tableau 13 Qualité de prévisions - Série corrigée

Modèle	MSE	ROOS	DM
AR(1)	23,193	-0,118	0,057
AR(p)	10,084	0,513	0,999
ARIMA	9,929	0,521	0,999
Holt Winter	2,406	0,883	0,997
ADAM ETS	5,784	0,713	0,998
ADAM ETS SARIMA	5,784	0,721	0,998
SSARIMA	3,170	0,847	0,997
CES	49,494	-1,387	0,000

D'après le Tableau 13, le modèle Holt-Winter se révèle être le plus performant. Il enregistre le plus faible MSE, ce qui implique une meilleure précision des prévisions, et le ROOS le plus élevé, indiquant de meilleures performances par rapport au modèle de référence, qui est le modèle naïf. La statistique DM étant positive et la p-value très proche de 1, cela nous dit qu'il n'y aucune preuve statistique que le modèle Holt-Winter soit moins performant que le modèle naïf.

À l'inverse, le modèle CES apparaît comme le moins performant. Il a le MSE le plus élevé, signalant des erreurs de prévision plus grandes, et le ROOS le plus bas, indiquant des performances largement inférieures au modèle naïf. La statistique DM très négative et la p-value de 0,00 montrent que les performances du modèle CES sont significativement moins bonnes que celles du modèle naïf.

7 Prédiction sur une année avec un pas de 1 mois

Tableau 14 Prédiction sur 12 mois Holt Winter

	Prédiction Holt Winter	Valeurs réelles corrigées
Janvier	126.8253	129.8342
Février	126.1869	128.5782
Mars	125.5485	127.6490
Avril	124.9101	125.7685
Mai	124.2717	123.8522
Juin	123.6333	122.4108
Juillet	122.9950	121.2518
Aout	122.3566	122.5805
Septembre	121.7182	122.9006
Octobre	121.0798	122.4146
Novembre	120.4414	121.4245
Décembre	119.8020	120.0314

Nous pouvons évaluer les performances du modèle Holt-Winter, qui s'est avéré être le plus efficace parmi les modèles testés. Le Tableau 14 nous permet de comparer les prévisions mensuelles sur une année et nous constatons que les écarts les plus marqués apparaissent au premier trimestre de 2023. En janvier, l'écart le plus significatif est observé, suivi de près par février et mars, ce qui révèle que le modèle a tendance à systématiquement sous-estimer les valeurs réelles en début d'année. À l'inverse, décembre et août montrent les écarts les plus faibles, indiquant que les prévisions sont très proches des valeurs réelles, décembre affichant la meilleure prédiction de l'année.

Ces prédictions résultent de l'application du modèle Holt-Winter avec lissage exponentiel, intégrant une composante de tendance mais dépourvue de composante saisonnière, optimisée automatiquement. En ce qui concerne les paramètres de lissage, la valeur d'alpha est fixée à 1 (voir Annexe 13), signifiant que le modèle privilégie fortement les observations les plus récentes en ajustant complètement vers la dernière valeur observée à chaque étape. La valeur de beta, quant à elle, est d'environ 0.584, ce qui témoigne du

degré de lissage appliqué à la composante de tendance. Une β inférieure à 1 mais supérieure à 0 permet une modération significative de l'impact des fluctuations à court terme sur la tendance estimée. Le dernier niveau lissé, noté « a », est d'environ 127.463, représentant l'estimation la plus récente de la valeur de la série, et le coefficient de tendance « b » est de -0.638, signalant une tendance descendante. Cela indique que le modèle prévoit une baisse continue à chaque période suivante, ajustée par le facteur de lissage de la tendance.

8 Conclusion

Au cours de ce projet, l'examen des différents modèles prévisionnels a révélé la complexité à la sélection d'un modèle optimal pour notre série. Nous avons commencé par traiter les données d'entraînement s'étalant de janvier 2005 à décembre 2022, période durant laquelle nous avons identifié et corrigé des valeurs atypiques et constaté que la série était non stationnaire. Pour remédier à cette non-stationnarité, nous avons appliqué une différenciation, ce qui a permis une comparaison plus homogène des différents modèles en utilisant une série corrigée et rendue stationnaire.

Malgré l'absence de saisonnalité initiale, la différenciation a introduit des fluctuations qui ont influencé la configuration des modèles. Nous avons distingué les approches en appliquant la série différenciée spécifiquement pour les modèles AR et ARIMA, tandis que pour les autres modèles, nous avons procédé à une réestimation en utilisant les données corrigées. Ces modèles ont été confrontés aux données réelles, puis aux données réelles ajustées des outliers.

Le modèle Holt-Winter, paramétré sans composante saisonnière, s'est révélé être le plus performant. Cependant, il est clair qu'il n'est pas sans défauts. La crise économique de 2022, ayant provoqué une hausse des indices et entraîné des valeurs atypiques cruciales avant notre période de prévision, a été un facteur clé mis en évidence par notre étude. La correction de ces valeurs atypiques a permis d'affiner notre compréhension des défis liés à la modélisation.

Il est évident que le modèle Holt-Winter pourrait être amélioré en intégrant des paramètres additionnels pour capturer les aspects spécifiques de la série, tels les variations discrètes ou les ruptures structurelles, qui n'ont peut-être pas été pleinement prises en compte. Ces ajustements permettraient d'augmenter la précision des prévisions en s'alignant davantage sur les comportements intrinsèques des données

9 Table des figures et tableaux

Figures

FIGURE 1- SERIE INITIAL	2
FIGURE 2 DETECTION DES POINTS ATYPIQUE.....	5
FIGURE 3 SERIE DIFFERENCIEE	7
FIGURE 4 CORRELOGRAMMES	8
FIGURE 5 PERIODOGRAMME.....	8
FIGURE 6 HISTOGRAMME	10
FIGURE 7 BOXPLOT	10
FIGURE 8 ANALYSE DE RESIDUS.....	17
FIGURE 9 PREVISIONS	18
FIGURE 10 COMPARAISON DES PREVISIONS – DONNEES DIFFERENCIEES	19
FIGURE 11 CSPE.....	21
FIGURE 12 COMPARAISON DES PREVISIONS - DONNEES REELLES	25
FIGURE 13 COMPARAISON DES PREVISIONS – DONNEES REELLES CORRIGEEES.....	26
FIGURE 14 - CSPE -REESTIMATION	26

Tableaux

TABEAU 1 POINTS ATYPIQUES	4
TABEAU 2 TESTS DE SAISONNALITE.....	5
TABEAU 3 TEST DE TENDANCE	5
TABEAU 4 DECOMPOSITION	6

TABLEAU 5 TEST DE STATIONNARITE	6
TABLEAU 6 TEST STATIONNARITE SERIE DIFFERENCIE	7
TABLEAU 7 STATISTIQUES DESCRIPTIVES COMPARAISON	9
TABLEAU 8 SKEWNESS ET KURTOSIS	9
TABLEAU 9 COMPARAISON DE MODELES.....	16
TABLEAU 10 MSE - R2OOS	20
TABLEAU 11 DM TEST	22
TABLEAU 12 ACCURACY.....	24
TABLEAU 13 QUALITE DE PREVISIONS - SERIE CORRIGEE	27

10 Bibliographie

- Agreste . *Intrants Agricoles*. No 403 04 2023. <https://agreste.agriculture.gouv.fr/agreste-web/download/publication/publie/SynCpr23403/consyn403202304-Intrants.pdf>.
- Carré, Camille. *Insee*. 05 2022. <https://blog.insee.fr/hausse-prix-production-tensions-inflationnistes/>.
- G. Colletaz, C. Hurlin. «HAL open science.» *Modèles Non Linéaires et Prévisions - pag, 55*. 06 2007. <https://shs.hal.science/halshs-00156692/document> (accès le 04 2024).
- Institut national de la statistique et des études économiques*. 05 02 2024. <https://www.insee.fr/fr/statistiques/serie/010538987#Tableau>.
- Institut National de la statistique et des études économiques*. 05 02 2024. <https://www.insee.fr/fr/statistiques/serie/010538987#Tableau>.
- La méthode de Holt et Winters (non saisonnière)*. Méthodes de Lissage Exponentiel (Techniques de Prévision et Conjoncture, O.Danréné, 2024).
- Svetunkov, I., Kourentzes. *ces: Complex Exponential Smoothing*. 2015. <https://www.rdocumentation.org/packages/smooth/versions/1.4.4/topics/ces>.
- Svetunkov, Ivan. *ADAM ETS Modèle adaptatif dynamique augmenté* . 04 2024. <https://cran.r-project.org/web/packages/smooth/vignettes/adam.html>.
- VERGEREAU., Alexis. *Ouest-France* . 01 2024. <https://www.ouest-france.fr/economie/agriculture/colere-des-agriculteurs-anatomie-dune-crise-qui-couve-depuis-longtemps-54e61b72-b9c4-11ee-9ea4-b02fbeb9c343>.

11 Annexes

Annexe 1 Détection de points atypiques

```
Series: ipampa
Regression with ARIMA(1,1,1) errors

Coefficients:
      ar1      ma1  LS122  LS202  LS207  TC214
      0.8434 -0.2691 1.5395 2.7421 6.6747 2.4083
s.e.  0.0531  0.1062 0.4497 0.4520 0.4505 0.3936

sigma^2 = 0.2814: log likelihood = -166.1
AIC=346.2  AICc=346.74  BIC=369.79

Outliers:
  type ind   time coefhat  tstat
1  LS 122 2015:02  1.539  3.423
2  LS 202 2021:10  2.742  6.067
3  LS 207 2022:03  6.675 14.817
4  TC 214 2022:10  2.408  6.119
```

Annexe 2 Tests de saisonnalité

```
> ft <- fried(ipampa)
> show(ft)
Test used: Friedman rank

Test statistic: 15.81
P-value: 0.1485093

>
> # Testing the seasonality of series
> # a boolean value is returned : TRUE or FALSE
> is <- isSeasonal(ipampa, test="wo")
> show(is)
[1] FALSE

> sd <- seasdum(ipampa)
> show(sd)
Test used: SeasonalDummies

Test statistic: 0.86
P-value: 0.5759353

>
> # Webel-Ollech test
> # Webel-Ollech test - new version of seastests (2021-09)
> # important
>
> wot <- combined_test(ipampa)
> show(wot)
Test used: WO

Test statistic: 0
P-value: 1 1 0.03397038
```

Annexe 3 Test de stationnarité ADF (Augmented Dickey-Fuller) sur la série ajusté 'ip

```
> adf.test(ipa_adj)

Augmented Dickey-Fuller Test

data: ipa_adj
Dickey-Fuller = -3.1388, Lag order = 5, p-value =
0.09931
alternative hypothesis: stationary
```

Annexe 4 Test de stationnarité ADF série différenciée

```
> adf.test(d_ipampa)

Augmented Dickey-Fuller Test

data: d_ipampa
Dickey-Fuller = -3.5546, Lag order = 5, p-value = 0.03868
alternative hypothesis: stationary
```

Annexe 5 Test de normalité, Shapiro test

```
> #normalité
> stats::shapiro.test(d_ipampa)

Shapiro-Wilk normality test

data: d_ipampa
W = 0.97467, p-value = 0.0006498
```

Annexe 6 Modèle AR(1), AR(p) et ARIMA(p,q,d)

```
> summary(ar1)
Series: d_ipampa
ARIMA(1,0,0) with non-zero mean

Coefficients:
      ar1      mean
    0.6966  0.2150
s.e.  0.0492  0.1172

sigma^2 = 0.2802: log likelihood = -167.65
AIC=341.3 AICc=341.41 BIC=351.41

Training set error measures:
              ME      RMSE      MAE MPE MAPE
Training set -0.000387161 0.526914 0.3958932 NaN  Inf
              MASE      ACF1
Training set 0.5050639 -0.08971407
```

```
Best model: ARIMA(2,0,0) with zero mean

> # Estimation
> summary(ar)
Series: d_ipampa
ARIMA(2,0,0) with zero mean

Coefficients:
      ar1      ar2
    0.6144  0.1489
s.e.  0.0680  0.0680

sigma^2 = 0.2779: log likelihood = -166.8
AIC=339.6 AICc=339.71 BIC=349.71

Training set error measures:
              ME      RMSE      MAE MPE MAPE
Training set 0.04941427 0.5246927 0.3970709 NaN  Inf
              MASE      ACF1
Training set 0.5065663 -0.01668794
```

```

Best model: ARIMA(1,0,1)          with zero mean

> # Estimation
> summary(arima)
Series: d_ipampa
ARIMA(1,0,1) with zero mean

Coefficients:
      ar1      ma1
    0.8434 -0.2691
s.e.  0.0527  0.1045

sigma^2 = 0.2761:  log likelihood = -166.1
AIC=338.2   AICc=338.31   BIC=348.31

Training set error measures:
              ME      RMSE      MAE MPE MAPE
Training set 0.04405366 0.5229599 0.3932828 NaN  Inf
              MASE      ACF1
Training set 0.5017336 0.01465994

```

Annexe 7 Modèle LED Holt-Winter

```

> show(hw)
Holt-Winters exponential smoothing with trend and without seasonal
component.

Call:
HoltWinters(x = d_ipampa, gamma = FALSE)

Smoothing parameters:
alpha: 0.6535009
beta : 0.05000744
gamma: FALSE

Coefficients:
      [,1]
a -0.76012345
b -0.07822051

```

Annexe 8 Modèle ADAM ETS

```

Best model: ARIMA(1,0,1)           with zero mean

> # Estimation
> summary(arima)
Series: d_ipampa
ARIMA(1,0,1) with zero mean

Coefficients:
      ar1      ma1
    0.8434  -0.2691
s.e.  0.0527  0.1045

sigma^2 = 0.2761:  log likelihood = -166.1
AIC=338.2  AICc=338.31  BIC=348.31

Training set error measures:
      ME      RMSE      MAE MPE MAPE
Training set 0.04405366 0.5229599 0.3932828 NaN Inf
      MASE      ACF1
Training set 0.5017336 0.01465994

```

Annexe 9 Modèle ADAM ETS ARIMA

```

> summary(adam_ets_sa)

Model estimated using auto.adam() function: ETS(ANN)+ARIMA(3,0,0)
Response variable: y
Distribution used in the estimation: Laplace
Loss function type: likelihood; Loss function value: 160.492
Coefficients:
      Estimate Std. Error Lower 2.5% Upper 97.5%
alpha      0.0009    0.0340    0.0000    0.0679
phi1[1]     0.5742    0.2028    0.1743    0.9739 *
phi2[1]     0.0759    0.0143    0.0477    0.1041 *
phi3[1]     0.0993    0.0968   -0.0917    0.2901
level       0.3170    1.0759   -1.8042    2.4373
ARIMASate1 -0.0050    1.0634   -2.1015    2.0905
ARIMASate2 -0.0020    0.3725   -0.7364    0.7319
ARIMASate3 -0.0146   18.1383  -35.7752   35.7290

Error standard deviation: 0.5488
Sample size: 215
Number of estimated parameters: 9
Number of degrees of freedom: 206
Information criteria:
      AIC      AICc      BIC      BICc
338.9841 339.8621 369.3198 371.6776

```

Annexe 10 Modèle SSARIMA

```
Time elapsed: 1.55 seconds
Model estimated: ARIMA(3,0,3)
Matrix of AR terms:
      Lag 1
AR(1) 0.1719
AR(2) 0.0180
AR(3) 0.5445
Matrix of MA terms:
      Lag 1
MA(1) 0.4367
MA(2) 0.3575
MA(3) -0.3599
Initial values were produced using backcasting.

Loss function type: likelihood; Loss function value: 161.0136
Error standard deviation: 0.5202
Sample size: 215
Number of estimated parameters: 7
Number of degrees of freedom: 208
Information criteria:
      AIC      AICc      BIC      BICc
336.0272 336.5682 359.6216 361.0746
```

Annexe 11 Modèle CES

```
> fitces1 <- auto.ces(d_ipampa, models=c("n", "s", "p", "f"), ic="AIC") # tester
plusieurs types de modèles
> summary(fitces1)
Time elapsed: 1.94 seconds
Model estimated: CES(f)
a0 + ia1: 1.4655+0.9264i
b0 + ib1: 0.995+0.9989i
Initial values were produced using backcasting.

Loss function type: likelihood; Loss function value: 158.5716
Error standard deviation: 0.5119
Sample size: 215
Number of estimated parameters: 5
Number of degrees of freedom: 210
Information criteria:
      AIC      AICc      BIC      BICc
327.1431 327.4302 343.9963 344.7672
```

Annexe 12 Résidus modèle CES

```
      Ljung-Box test

data:  Residuals
Q* = 24.373, df = 24, p-value = 0.4405
```

Annexe 13 Holt Winter, deuxième estimation avec les données corrigées

```
Call:
HoltWinters(x = ipampa, gamma = FALSE)

Smoothing parameters:
alpha: 1
beta : 0.5846717
gamma: FALSE

Coefficients:
      [,1]
a 127.4636507
b -0.6383839
```