

Méthodes de prévision

Hicham Kaffaf
Usama Sheikh

December 29, 2024

1 Introduction

Ce projet économétrique se concentre sur la prévision de l'Indice des Prix à la Consommation (CPI) en Italie, en utilisant trois méthodes de prévision différentes : les modèles Exponential Smoothing Model (ESM), Singular Spectrum Analysis (SSA), et AutoRegressive Integrated Moving Average (ARIMA). Notre objectif est d'évaluer et de comparer l'efficacité et la précision de chacun de ces modèles pour prédire les futures tendances du CPI italien.

2 Travail préliminaire sous SAS

2.1 Représentation graphique de notre série

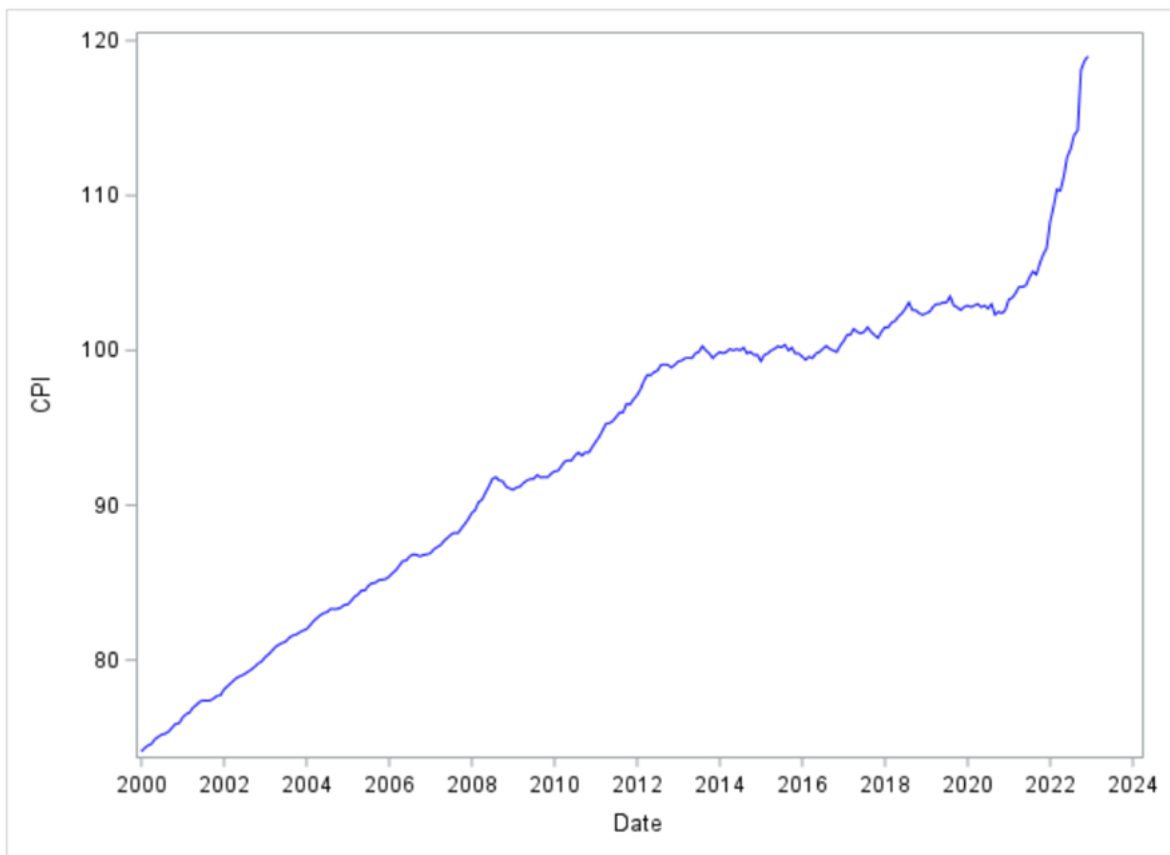


Figure 1: Représentation de la série du CPI (2000-2022).

Ce graphique représente l'indice des prix à la consommation (CPI) sur une période allant de 2000 à 2022. La tendance générale du CPI sur cette période est à la hausse, indiquant une inflation continue

au fil des années. Vers la fin du graphique, on observe une accélération marquée de l'augmentation du CPI, particulièrement à partir de 2020. Cette période coïncide avec la pandémie du COVID-19, qui a entraîné des perturbations économiques majeures. À première vue, le CPI ne semble pas stationnaire car il montre une tendance à la hausse au fil du temps. Il faudrait probablement différencier la série afin d'utiliser le modèle ARIMA.

2.2 Test de stationnarité

Pour évaluer l'ordre d'intégration de la série, on utilise le test KPSS (Kwiatkowski-Phillips-Schmidt-Shin) pour tester la stationnarité de la série. Le test KPSS a pour hypothèse nulle que la série est stationnaire autour d'une tendance ou d'un niveau moyen. Voici les principales hypothèses du test KPSS :

- Hypothèse nulle (H0) : La série est stationnaire (ou stationnaire autour d'une tendance).
- Hypothèse alternative (H1) : La série a une racine unitaire, c'est-à-dire qu'elle n'est pas stationnaire et présente une tendance stochastique.

Table 1: Résultats du test de stationnarité KPSS

Type	Retards	Eta	P-value
Moyenne simple	10	2.5392	0.0001
Tendance	10	0.3518	0.0006

Les p-values sont inférieures au seuil de significativité de 1%, on rejette donc l'hypothèse nulle de stationnarité. Les résultats du test KPSS indiquent donc que la série temporelle n'est pas stationnaire, ni autour d'une moyenne constante, ni autour d'une tendance linéaire. En différenciant notre série, on obtient le résultat suivant :

Table 2: Test de stationnarité KPSS

Type	Retards	Eta	P-value
Moyenne simple	8	0.2750	0.1588
Tendance	8	0.2331	0.0069

La p-value pour la moyenne simple est supérieur à 10%, on ne rejette donc pas l'hypothèse nulle de stationnarité. La série en différence première est stationnaire en moyenne.

3 ESM sous SAS

3.1 Choix et spécification du modèle ESM

À première vue du graphique du CPI, le modèle ESM qui semble le moins adapté est le Single ESM parce que la série présente une tendance. Pour choisir le bon modèle ESM pour nos prévisions, on a dû estimer les 3 autres modèles et choisir celui qui minimise les critères d'informations (MSE, AIC, BIC, ...). Les différents modèles ont été estimés sur tous l'échantillon, on obtient les résultats suivants :

Table 3: Comparaison des critères d'information pour différents modèles

Critères d'info	Single	Double	Holt	DTM
MSE	0,14422108	0,09670746	0,09381888	0,09385998
RMSE	0,37976451	0,31097824	0,30629868	0,30636577
AIC	-532,44857	-642,75385	-649,12341	-647,00252
BIC	-528,82817	-639,13345	-641,88261	-636,14131

Le modèle qui minimise les critères d'information est le ESM Holt, après estimation du modèle sur tous l'échantillon, on obtient le résultat suivant :

Table 4: Lissage exponentiel linéaire Paramètres estimés

Paramètre	Estimation	Erreur type	Valeur du test student	P-value
Pondération du niveau	0.85250	0.04589	18.58	0.0001
Pondération de la tendance	0.19721	0.03406	5.79	0.0001

En utilisant ces paramètres, les équations pour le modèle Holt ESM avec les paramètres estimés sont les suivantes :

1. Équation du Niveau :

$$\hat{X}_t = X_t + (1 - 0.8520)(\hat{X}_{t-1} + b_{t-1})$$

2. Équation de la Pente :

$$b_t = (\hat{X}_t - \hat{X}_{t-1}) + (1 - 0.19721)b_{t-1}$$

Avec un α de 0.8520, cela signifie que $1 - \alpha = 0.1480$, et ainsi le modèle attribue seulement environ 14.8% de poids aux valeurs antérieures du niveau lors de l'ajustement du niveau actuel.

De même, avec un β de 0.19721, cela donne $1 - \beta = 0.80279$, donc environ 80.279% du poids est attribué aux estimations passées de la tendance. Cela indique que le modèle valorise les informations historiques de la tendance et ne réagit pas trop rapidement aux changements récents, ce qui peut aider à lisser les anomalies et à fournir une tendance plus stable.

3.2 Qualité du fit du modèle

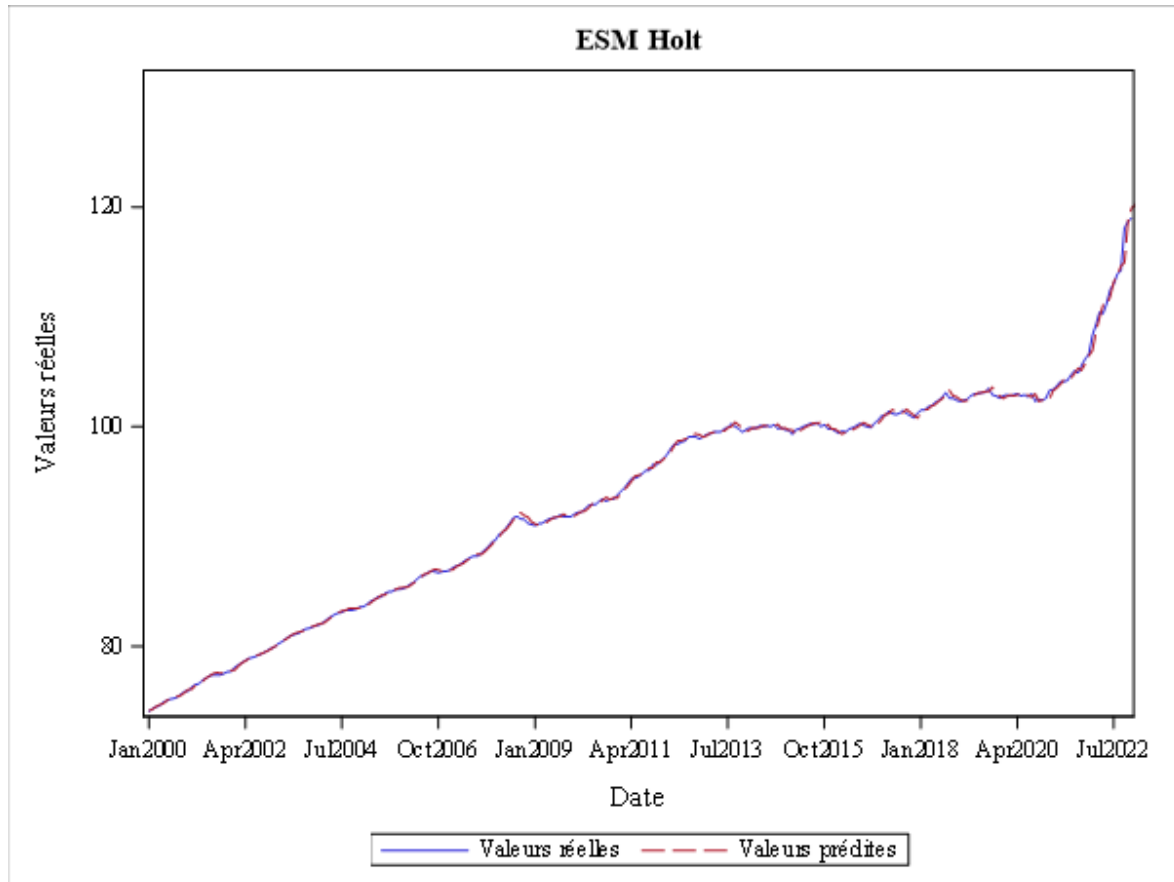


Figure 2: Qualité du fit du modèle ESM Holt.

Les valeurs prédites semblent suivre de près la tendance générale des valeurs réelles, ce qui indique que le modèle de lissage exponentiel de Holt capte bien la tendance de la série temporelle. Le

modèle réagit assez rapidement aux changements dans les données réelles. Cela suggère que le modèle est capable d'ajuster ses prédictions pour s'aligner avec des variations récentes dans les données. Dans l'ensemble, le modèle semble fournir une bonne précision de prédiction, comme en témoigne l'ajustement étroit entre les valeurs réelles et les valeurs prédites sur la majeure partie de la série temporelle.

3.3 Erreur de prévision à 1 pas in sample

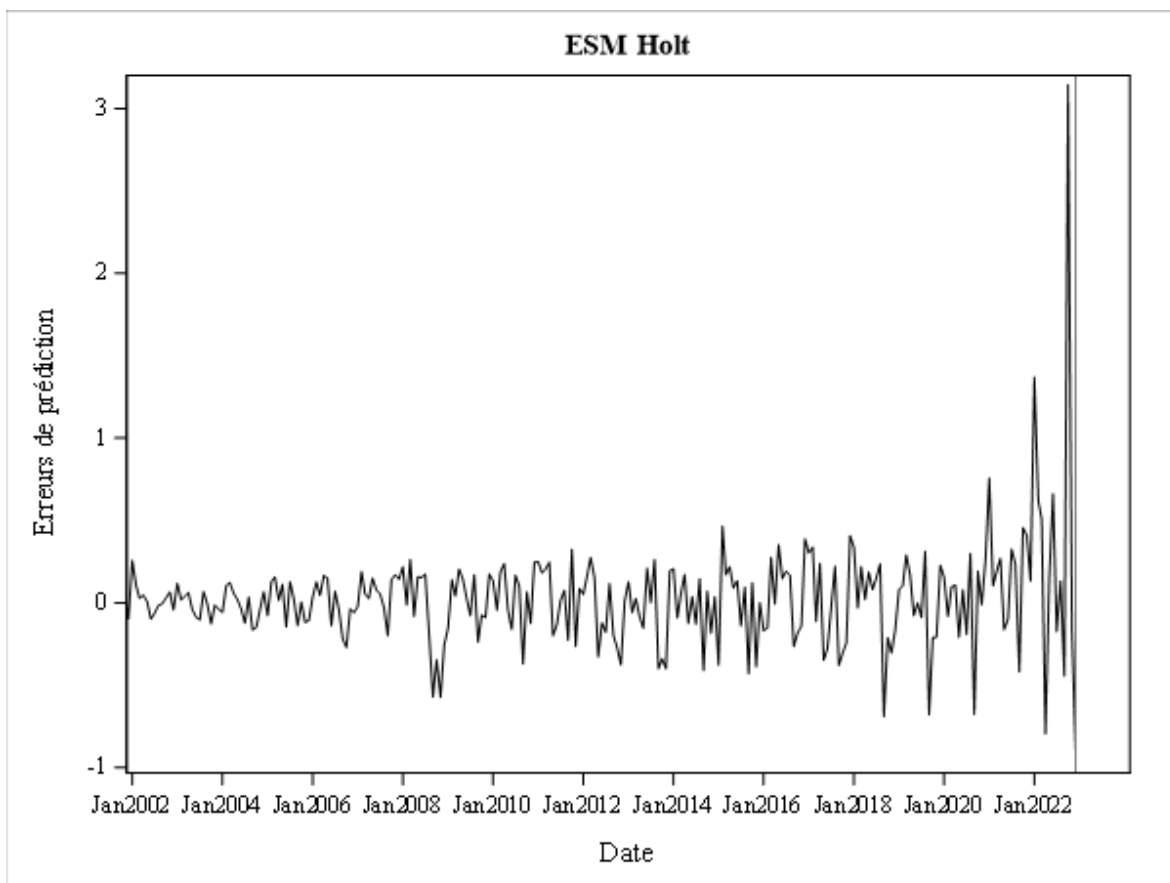


Figure 3: Représentation de l'erreur à 1 pas.

La majorité des erreurs se situent dans une fourchette relativement étroite autour de zéro, ce qui indique que les prévisions sont pour la plupart proches des valeurs réelles. Les fluctuations dans les erreurs semblent relativement stables au fil du temps, sauf pour les données à partir de mars 2020 et pour le pic à la fin. Ce pic est significatif dans l'erreur de prévision. Il peut indiquer une période d'instabilité (notamment la hausse de prix due à la crise des produits énergétique causé par la guerre entre la Russie et l'Ukraine) qui n'a pas été captée par le modèle.

4 ARIMA sous SAS

4.1 Spécification du modèle ARIMA

Le test de stationnarité KPSS nous a bien montré que la série CPI n'est pas stationnaire en niveau. Ce constat nous causait aucun problème durant la partie ESM, mais durant cette partie, on sera obligé de travailler avec la série en différence première afin de purger la tendance et régler le problème de stationnarité.

Grâce à la proc arima sur logiciel Sas, on obtient le résultat suivant :

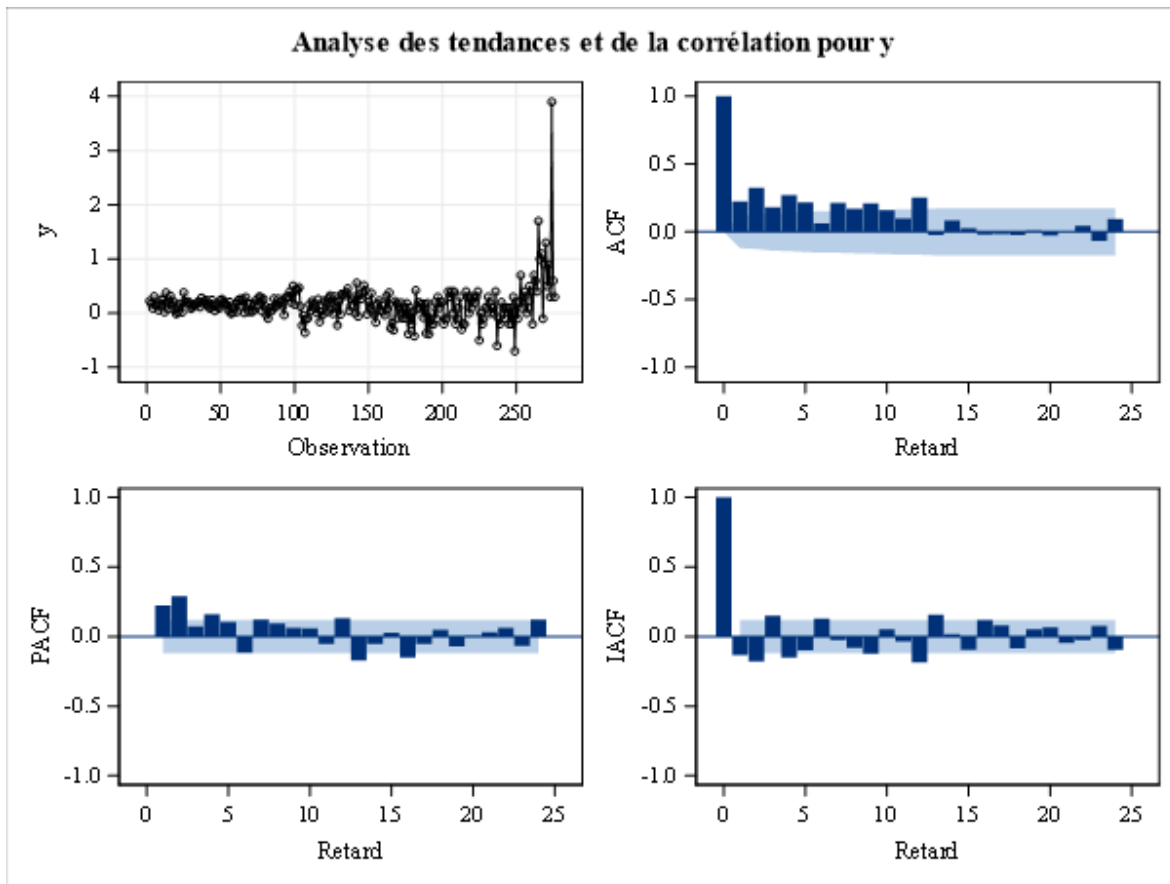


Figure 4: ACF et PACF de la variable CPI (en différence première).

On remarque qu'il existe 6 retards significativement différents de 0 pour les ACF et 2 retards significativement différents de 0 pour les PACF. Cela nous fait un modèle ARIMA(2,1,6), après estimation de notre modèle sur logiciel SAS on obtient le résultat suivant :

Table 5: Estimation du maximum de vraisemblance					
Paramètre	Estimation	Erreur type	Valeur du test t	Approx P-value	Retard
MU	0.25229	0.21460	1.18	0.2397	0
MA1,1	0.15638	0.48232	0.32	0.7458	1
MA1,2	0.52548	0.41149	1.28	0.2016	2
MA1,3	0.05471	0.10128	0.54	0.5891	3
MA1,4	-0.18176	0.10729	-1.69	0.0902	4
MA1,5	-0.05717	0.13448	-0.43	0.6707	5
MA1,6	0.20147	0.09916	2.03	0.0422	6
AR1,1	0.25817	0.48209	0.54	0.5923	1
AR1,2	0.70437	0.46022	1.53	0.1259	2

On remarque que les paramètres estimés ne sont pas significatifs au seuil de 10%. Cela est peut-être dû à des problèmes de normalité ou d'hétéroscédasticité des résidus.

4.2 Test de normalité et d'hétéroscédasticité

En utilisant la proc univariate sur logiciel SAS, on effectue le test de normalité, et à l'aide de la proc autoreg on effectue le test ARCH sur les résidus. On obtient les résultats suivants :

Table 6: Tests de normalité

Test	Statistique	p-value
Shapiro-Wilk	0.73545	0.0001
Kolmogorov-Smirnov	0.134665	0.0100
Cramer-von Mises	1.430609	0.0050
Anderson-Darling	9.133287	0.0050

On remarque que les p-values de tous les tests sont très faibles, prenant par exemple le test de Shapiro-Wilk (Le test le plus adapté pour nos données $T < 2000$), Le test de Shapiro-Wilk est utilisé en statistique pour tester l'hypothèse de normalité d'un échantillon. Voici les hypothèses impliquées dans ce test :

- **Hypothèse nulle (H_0):** L'échantillon provient d'une population normalement distribuée.
- **Hypothèse alternative (H_1):** L'échantillon ne provient pas d'une population normalement distribuée.

La p-value associée au test de shapiro-wilk est très faible (< 0.05), on rejette donc l'hypothèse nulle et on conclue que les données de notre échantillon ne sont pas normalement distribuées.

On va aussi tester l'hétéroscédasticité dans nos résidus à l'aide du test ARCH, Voici les hypothèses de base du test ARCH :

- **Hypothèse nulle (H_0):** Il n'y a pas d'hétéroscédasticité conditionnelle dans les erreurs du modèle.
- **Hypothèse alternative (H_1):** Il existe de l'hétéroscédasticité conditionnelle dans les erreurs du modèle.

Table 7: Tests pour perturbations ARCH basés sur les résidus OLS

Ordre	Q	$Pr > Q$	LM	$Pr > LM$
1	0.0005	0.9823	0.0004	0.9835
2	2.4727	0.2904	2.4635	0.2918
3	2.5169	0.4722	3.0791	0.3796
4	3.0650	0.5470	10.2008	0.0372
5	3.0675	0.6896	10.2016	0.0697
6	3.5233	0.7409	14.3665	0.0258
7	3.5339	0.8316	14.4454	0.0438
8	3.7134	0.8820	15.7036	0.0468
9	17.6645	0.0393	203.9362	$< .0001$
10	17.6681	0.0608	203.9609	$< .0001$
11	17.8235	0.0858	204.0604	$< .0001$
12	18.5245	0.1007	214.6309	$< .0001$

1. **Ordre 1 à 3 :** Pour les retards de 1 à 3, les valeurs p sont supérieures à 0.05, indiquant qu'il n'y a pas d'effet ARCH à ces niveaux de retard.
2. **Ordre 4 à 12 :** À partir du retard 4, les valeurs Pr sont inférieures à 0.05, suggérant que l'hypothèse nulle d'absence d'hétéroscédasticité conditionnelle peut être rejetée. Cela indique une présence potentielle d'effet ARCH dans les résidus de notre modèle.

4.3 Modèle ARIMA-X

Notre modèle rencontre des problèmes de normalité et d'hétéroscédasticité. Afin de régler ceci, on fait recours un modèle ARIMA-X. On repère les outliers à l'aide de la proc univarité, puis on les intègre dans notre modèle.

Table 8: Observations extrêmes			
La plus petite		La plus grande	
Valeur	Obs	Valeur	Obs
-0.962360	276	0.591707	262
-0.698020	225	0.659032	270
-0.683995	237	0.796704	253
-0.656002	249	1.502914	265
-0.618973	268	3.077857	274

On a pu intégrer 6 dummy pour notre modèle ARIMA-X. On refait les mêmes tests sur les nouveaux résidus, on obtient les résultats suivant :

Table 9: Tests de normalité		
Test	Statistique	p-value
Shapiro-Wilk W	0.942586	0.0001
Kolmogorov-Smirnov D	0.04342	0.1500
Cramer-von Mises W-Sq	0.08185	0.2029
Anderson-Darling A-Sq	0.73397	0.0566

On remarque que l'intégration des ces dummy a amélioré les p-values des différents tests de normalité sauf pour celui du Shapiro-Wilk.

Table 10: Tests pour perturbations ARCH basés sur les résidus OLS

Ordre	Q	Pr >Q	LM	Pr >LM
1	0.2438	0.6215	0.2473	0.6190
2	6.2970	0.0429	6.2069	0.0449
3	6.4723	0.0908	6.2957	0.0981
4	10.4630	0.0333	8.9672	0.0619
5	15.3902	0.0088	13.2169	0.0214
6	16.9421	0.0095	13.6928	0.0333
7	17.6741	0.0135	13.7490	0.0558
8	17.8012	0.0228	13.7651	0.0881
9	17.8187	0.0373	13.9785	0.1231
10	18.0525	0.0541	13.9785	0.1740
11	18.1376	0.0784	14.3206	0.2158
12	18.2641	0.1079	14.7441	0.2557

Les dummies ont globalement amélioré le test d'hétéroscédasticité (en améliorant les p-values associées au test).

Précision sur nos données : Après avoir examiné et testé de manière exhaustive plusieurs méthodes (combinaison de dummy/test d'autres "p" et "q", approche minic de SAS) pour corriger la normalité des résidus de notre modèle ARIMA, nous sommes parvenus à la conclusion que la solution actuelle, bien que n'étant pas parfaite, est la meilleure parmi toutes celles que nous avons envisagées. Nous avons exploré la transformation logarithmique, qui est souvent préconisée pour stabiliser la variance et atténuer les problèmes liés à la non-normalité. Cependant, cette dernière technique n'a pas abouti à des améliorations significatives dans notre cas. Les résultats obtenus avec notre modèle actuel surpassent ceux des modèles antérieurs.

Table 11: Estimation ARIMA-X

Paramètre Changement de niveau	Estimation	Erreur type	Valeur du test t	Approx $Pr > t $	Retard	Variable
MU	0.16170	0.03409	4.74	< .0001	0	Yt
MA1,1	-0.69649	0.17162	-4.06	< .0001	1	Yt
MA1,2	0.08610	0.21351	0.40	0.6867	2	Yt
MA1,3	-0.27958	0.11581	-2.41	0.0158	3	Yt
MA1,4	-0.16702	0.11038	-1.51	0.1302	4	Yt
MA1,5	-0.20687	0.10900	-1.90	0.0577	5	Yt
MA1,6	-0.14899	0.07752	-1.92	0.0546	6	Yt
AR1,1	-0.43250	0.16492	-2.62	0.0087	1	Yt
AR1,2	0.55446	0.17122	3.24	0.0012	2	Yt
NUM1	-0.06039	0.22445	-0.27	0.7879	0	dum1
NUM2	-0.57279	0.13900	-4.12	< .0001	0	dum2
NUM3	-0.91480	0.14068	-6.50	< .0001	0	dum3
NUM4	0.53578	0.19318	2.77	0.0055	0	dum4
NUM5	3.41214	0.21154	16.13	< .0001	0	dum6
NUM6	0.22744	0.20022	1.14	0.2560	0	dum7

On remarque que la majorité des paramètres sont significatifs, sauf pour quelques uns notamment le MA1,2, le MA1,4 et le NUM1 et 6. L'estimation de ce modèle (ARIMA-X) nous a donné un meilleur résultat que celui du ARIMA (tous les coeffs étaient non significatifs avec le modèle ARIMA).

- Les termes MA significatifs indiquent que les erreurs passées ont un effet (dans ce cas, négatif) sur les valeurs actuelles de la série temporelle.

- Les termes AR significatifs suggèrent que les valeurs passées de la série ont un effet direct sur les valeurs actuelles, dans ce cas, un effet négatif pour le premier retard et positif pour le second.

4.4 Qualité du fit du modèle

La ligne de prévision suit de près la série réelle, ce qui suggère que le modèle ARIMA-X avec les termes sélectionnés fait un bon travail pour capturer la dynamique de la série temporelle. Le modèle semble bien capturer la tendance générale de la série temporelle.

En fin de compte, l'efficacité des prévisions et l'ajustement du modèle dépendent non seulement des termes AR et MA mais aussi de la correction de la non-stationnarité et de la normalisation des résidus, ce que les variables dummy significatives aident à réaliser.

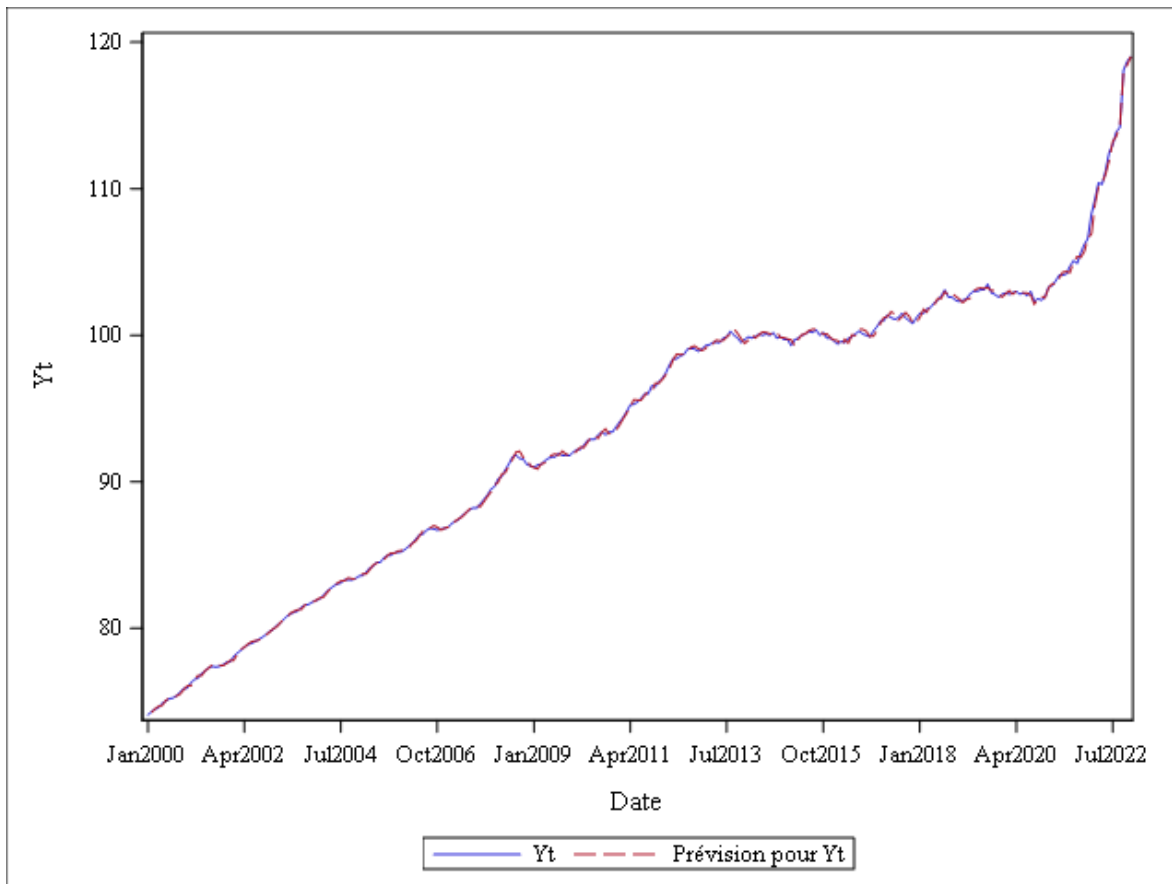


Figure 5: Qualité du fit du modèle ARIMA-X.

4.5 Erreur de prévision à un pas in sample

La variance des résidus semble rester assez constante au fil du temps, ce qui est une bonne indication que le modèle a une variance homogène (homoscédasticité). Cependant, il y a une augmentation notable de la volatilité vers la fin de la série, autour de juillet 2022.

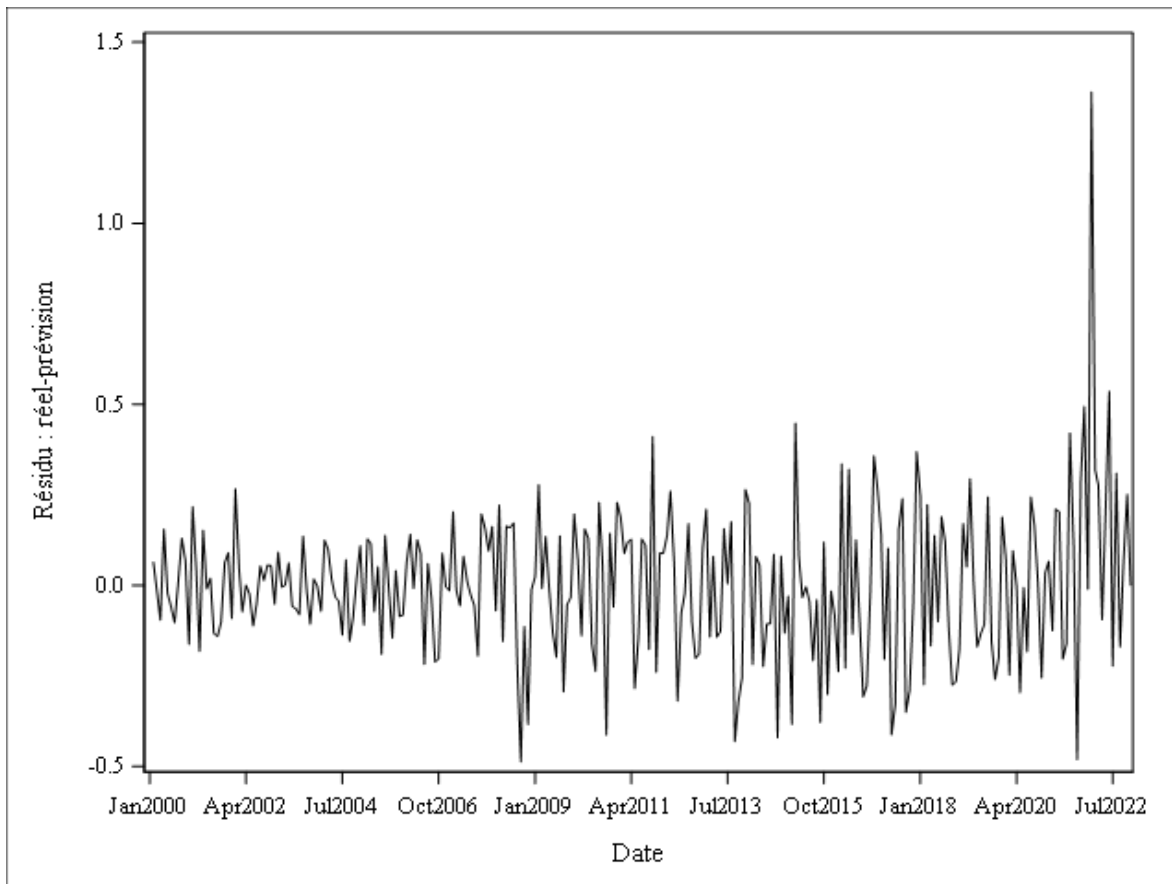


Figure 6: Erreur de prévision ARIMA-X.

5 SSA sous R

5.1 Composantes de notre modèle SSA

La Représentation de la décomposition de la série du CPI est la suivante :

$$Y_t = \tau_t + c_t + \varepsilon_t,$$

1. Tendence:

- **Application au CPI** : La tendance dans le CPI reflète l'évolution générale des niveaux de prix sur une longue période. Par exemple, une tendance à la hausse pourrait indiquer une inflation persistante. En regardant le graphique de la série du CPI de l'Italie, il est clair qu'il existe une tendance à la hausse dans la série.

2. Cycle:

- **Application au CPI** : Le CPI peut avoir des cycles ou des motifs saisonniers. Par exemple, les prix peuvent augmenter pendant la période des fêtes et diminuer ensuite, ou fluctuer en fonction des saisons agricoles ou touristiques.

3. Bruit:

- **Raison** : En plus des tendances et des cycles, il y a toujours des variations aléatoires ou inattendues dues à des événements imprévisibles ou à des anomalies dans les données.

5.2 L'approche utilisée pour la taille de la fenêtre et pour l'eigentriple grouping

La première étape (Embedding) consiste à transformer la série temporelle originale en une matrice appelée "matrice de trajectoire". On découpe la série temporelle en fenêtres superposées de taille fixe, ce qui permet de capturer la structure interne des données. Afin de déterminer la taille de cette fenêtre, on a suivi la recommandation d'Elsner et Tsonis (1996) qui suggère de fixer L tel que $L = T/4$.

Pour la partie du eigentriple grouping, on a choisi de suivre la méthode des w-correlations parceque l'utilisation des fréquences nécessite un ancrage académique et des connaissances sur les propriétés de la variable étudiées (Le CPI dans notre cas).

5.3 Les différentes composantes identifiées par SSA.

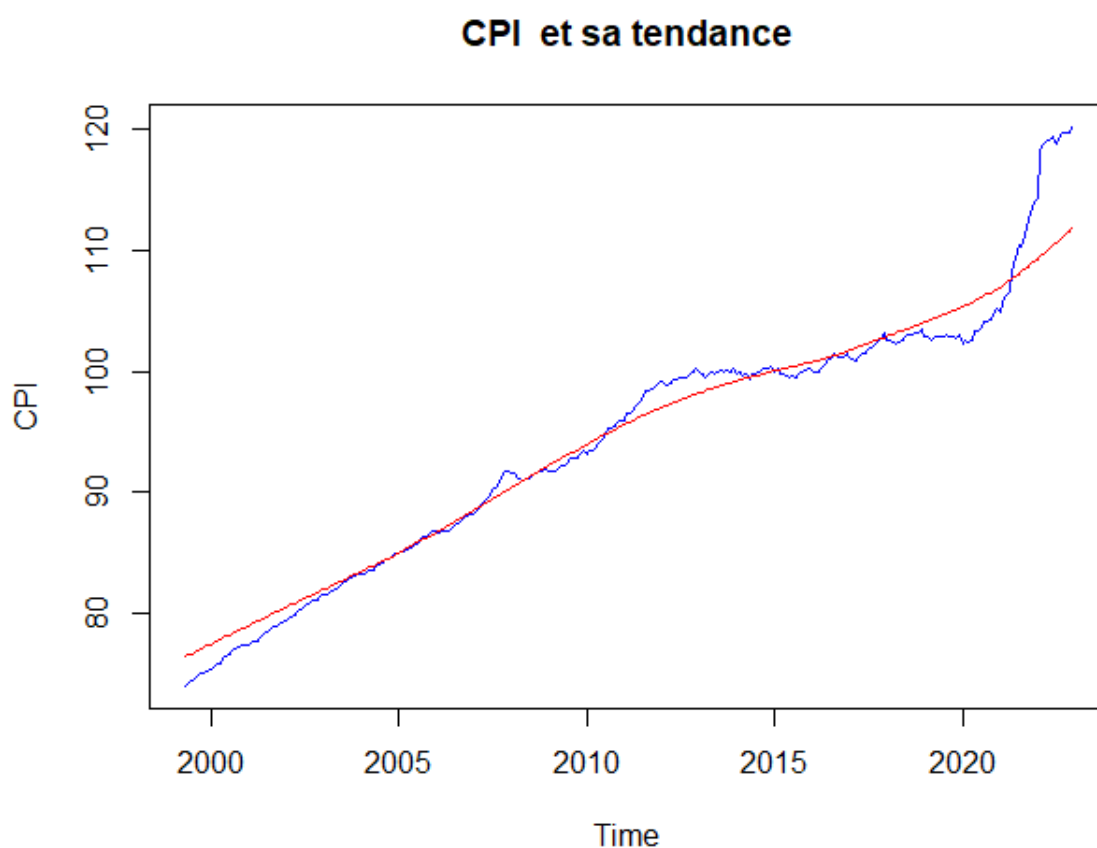


Figure 7: CPI et sa composante tendance.

La tendance suit les valeurs observés mais se détache de ces-derniers à partir de 2020. Cela suggère une période d'inflation plus élevée notamment dû la crise Covid-19.

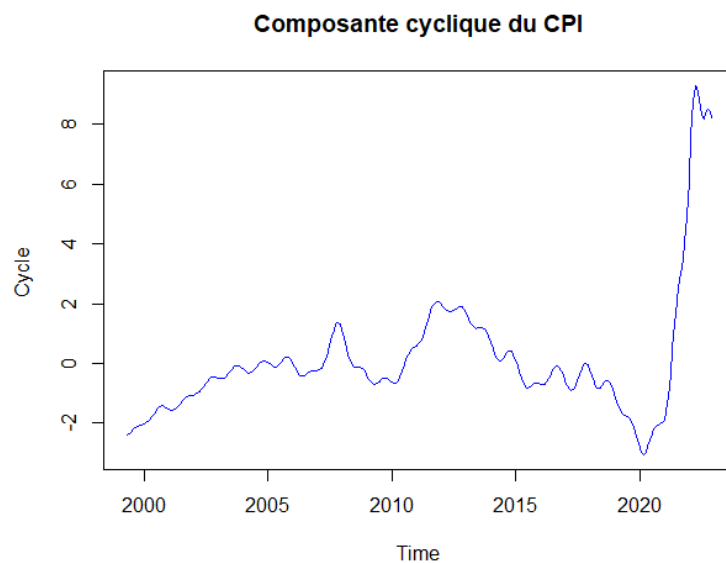


Figure 8: Composante cyclique du CPI.

Le graphique suggère que la composante cyclique du CPI ne montre pas les caractéristiques régulières attendues d'un cycle économique typique en raison d'une volatilité marquée. Le choc significatif survenu après 2020 semble contribuer fortement à cette irrégularité.

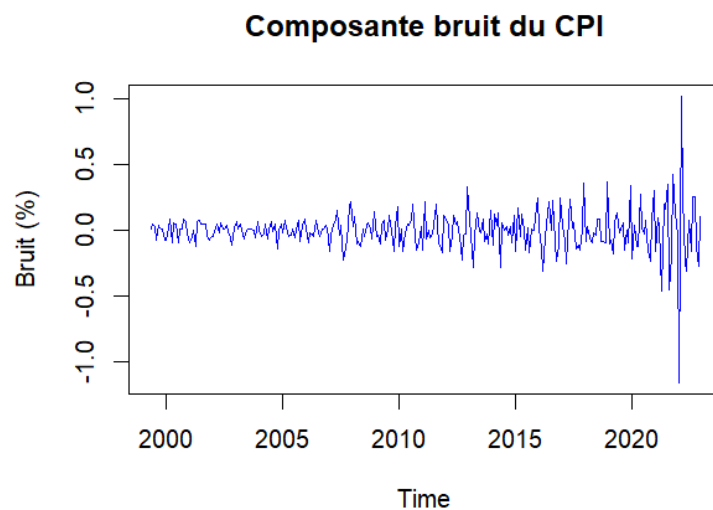


Figure 9: Composante bruit du CPI.

Le graphique montre la composante de bruit de l'Indice des Prix à la Consommation (CPI), qui représente les variations aléatoires et irrégulières de l'indice qui ne sont pas expliquées par la tendance ou les composantes cycliques. Jusqu'à environ 2020, le bruit fluctue autour de zéro. Cependant, il y a une forte augmentation de la volatilité après 2020, ce qui peut refléter la turbulence des marchés ou des chocs économiques imprévus liés à des événements comme la pandémie COVID-19.

5.4 Qualité du fit du modèle.

Sur le graphique, la ligne bleue représente la série réelle de l'Indice des Prix à la Consommation (CPI) et la ligne rouge indique la série de CPI estimée ou prédite. Les deux séries semblent très proches l'une de l'autre, ce qui suggère que la prédiction du modèle est assez précise et qu'elle suit bien les tendances réelles du CPI.

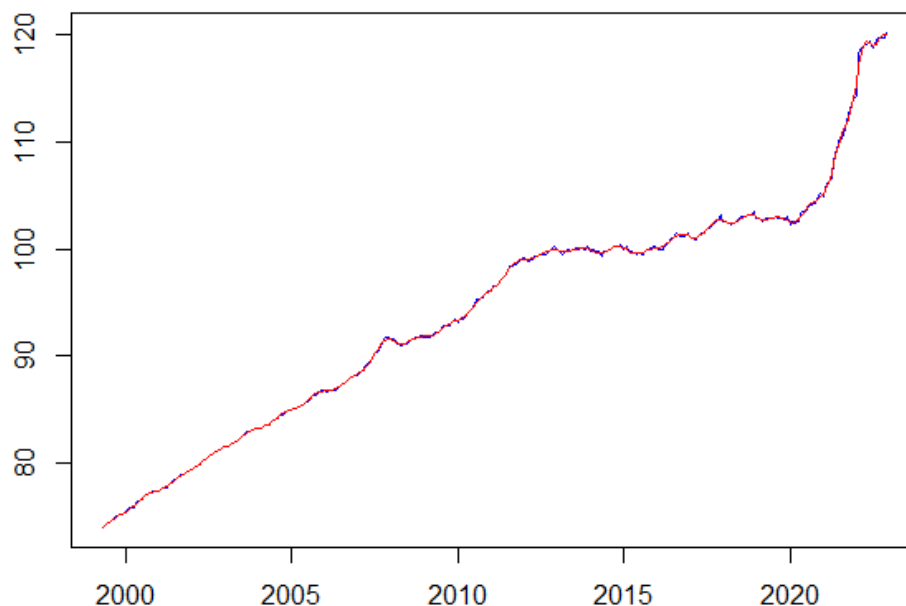


Figure 10: Qualité du fit du modèle.

5.5 L'erreur de prévision à un pas in sample.

La plupart des erreurs fluctuent autour de zéro, ce qui indique que la prédiction est, en général, assez précise. Cependant, il y a une augmentation notable de l'erreur de prévision vers la fin de la série, ce qui peut signifier que le modèle de prévision n'a pas bien capturé un changement récent dans les données ou qu'un événement imprévu a causé une variation significative entre les valeurs réelles et les valeurs prédites (Dans notre cas la crise Covid19).

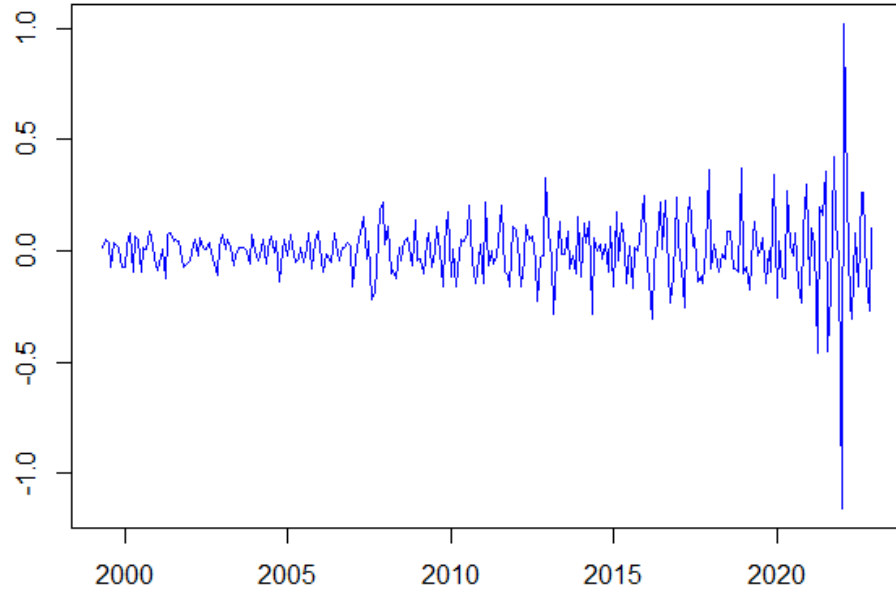


Figure 11: Erreur de prévision.

6 Comparaison des modèles

6.1 Mean absolute Error (MAE) et Root Mean Squared Error (RMSE)

Le MAE est défini comme suit :

$$\text{MAE} = \frac{1}{m} \sum_{i=T+1}^{T+m} |e_i|, \quad (1)$$

où les e_i sont les erreurs de prévision (à un pas) sur un échantillon de taille m .

Le RMSE est défini comme suit :

$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=T+1}^{T+m} e_i^2}. \quad (2)$$

On a calculé ces critères d'informations sur logiciel SAS et voici le résultat :

Table 12: Comparaison des MAE et RMSE pour différents modèles

Modèle	MAE	RMSE
ARIMA-X	0,151950078	0,201366498
SSA	0.086534468	0,150099871
Holt ESM	0.18625945	0,306298678

Le modèle qui minimise le MAE/RMSE est le SSA. Le modèle ESM maximise ces critères.

6.2 Test de Diebold-Mariano

Le test de Diebold-Mariano est utilisé pour comparer les performances prédictives de deux modèles de séries temporelles. Il évalue si la différence dans les erreurs de prévision entre les deux modèles est statistiquement significative.

- statistique du test :

$$DM = \frac{\bar{d}}{\hat{\sigma}_d} \sim N(0, 1), \quad (3)$$

avec $\bar{d} = \frac{1}{N} \sum_{t=1}^N d_t$ où N est le nombre d'observations composant d_t , et $\hat{\sigma}_d$ est une estimation (convergente) de l'écart-type de d .

Pour la mise en place du test, nous avons utilisé la fonction de perte suivante : $g(e_t) = e_t^2$ où e_t correspond à l'erreur de prévision.

Les résultats du test sur SAS sont les suivants :

Table 13: Le test de significativité de la cte pour les différents modèles

Modèles	P-Value du test de significativité de student
ESM/ARIMA	0,1595
ARIMA/SSA	0,1189
SSA/ESM	0,0195

On rejette l'hypothèse nulle pour ESM/ARIMA et ARIMA/SSA (sous H_0 , les modèles $M1$ et $M2$ conduisent à des prévisions de précision identique). Par contre pour SSA/ESM, l'hypothèse nulle n'est pas rejetée. C'est à dire que les modèles SSA et ESM conduisent à des prévisions de précision identique. Mais finalement, on choisit de retenir le modèle SSA pour faire nos prévisions puisque c'est celui qui minimise les critères d'informations MAE et RMSE.

7 Analyse prédictive

7.1 Prévisions jusqu'à décembre 2023 avec le modèle SSA

Durant cette partie, on effectue des prévisions jusqu'à décembre 2023 avec le modèle retenu SSA suite au critères d'informations MAE et RMSE et au test de Diebold-Mariano.

Toujours à l'aide de R, on fait nos prévisions accompagné d'un intervalle de confiance au seuil de 90% (Nos données ne sont pas normalement distribuées), et on obtient le résultat suivant :

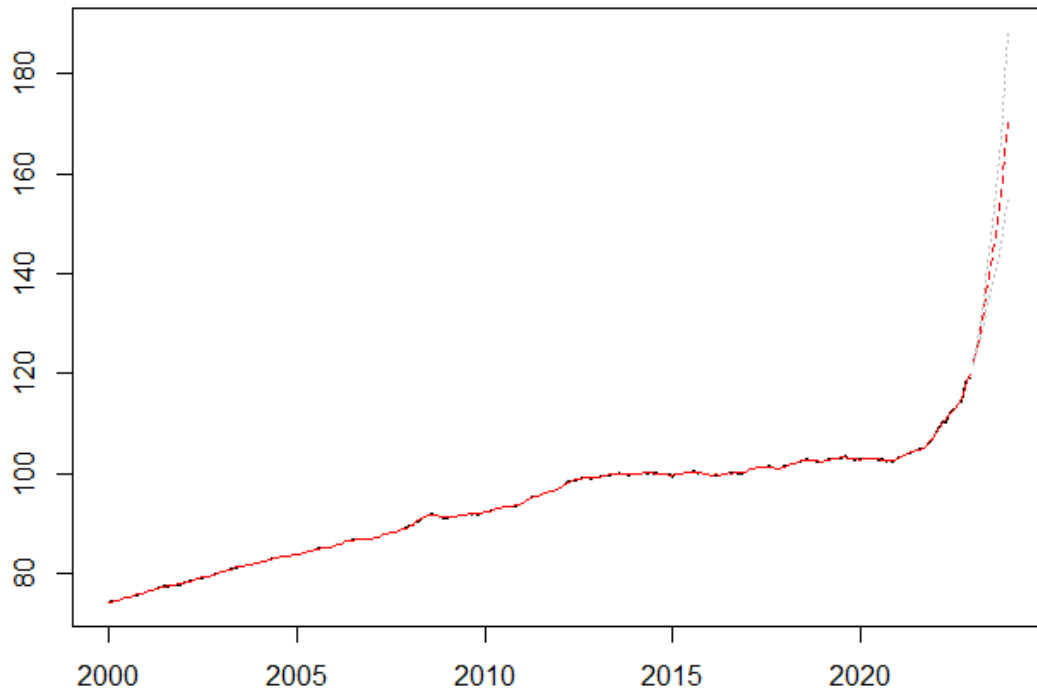


Figure 12: Pr vision avec le mod le SSA.

Le mod le SSA pr voit une hausse assez significative du CPI sur 2023. L'intervalle de confiance s' largit   mesure que l'on se projette dans l'avenir, indique que plus on s' loigne dans le temps, plus les pr visions deviennent incertaines.

7.2 Comparaison des pr vision SSA avec les pr visions du site TradingEconomics

Nos pr visions avec le mod le SSA ne semblent  tre de bonne qualit . En les comparant avec les vraies donn es observ es d'une part, et avec les pr visions du site "trading economics" d'autre part. On s'est rendu compte que nos pr visions s' loignent de la r alit . Voici nos pr visions avec SSA :

Table 14: CPI SSA mensuel

Mois	CPI SSA
Janvier	121,99073
F�vrier	124,09454
Mars	126,32729
Avril	128,50953
Mai	130,58843
Juin	132,56599
Juillet	134,60191
Ao�t	136,68354
Septembre	138,76967
Octobre	140,89409
Novembre	143,2484
D�cembre	145,74144

A partir de ces pr visions on peut calculer le taux d'inflation annuel comme suit :

$$\text{Taux d'inflation annuel (\%)} = \left(\frac{\text{CPI de la fin de décembre 2023} - \text{CPI de décembre 2022}}{\text{CPI de décembre 2022}} \right) \times 100$$

Cela nous fait une inflation annuel de 19%. Ce qui est totalement biaisé bien sûr. Le site de trading economics nous montre que sur le mois de novembre 2023 par exemple, on a eu un CPI réel de 119.6 contre un CPI de 143.24 estimé par notre modèle SSA avec une prévision de taux d'inflation de 0.8%.

Pour consulter l'indice des prix à la consommation (CPI) pour l'Italie, veuillez visiter le site [Trading Economics](#).

Pour consulter le taux d'inflation pour l'Italie, veuillez visiter le site [Trading Economics](#).

7.3 Comparaison des prévision avec les données réelles in-sample

Mois	CPI SSA	CPI Réel	Différence
Janvier	121,99073	119,1	-2,89073
Février	124,09454	119,3	-4,79454
Mars	126,32729	118,8	-7,52729
Avril	128,50953	119,3	-9,20953
Mai	130,58843	119,7	-10,88843
Juin	132,56599	119,7	-12,86599
Juillet	134,60191	119,7	-14,90191

Table 15: Comparaison entre le CPI SSA et le CPI Réel

On constate que le modèle SSA surévalue systématiquement le CPI. En janvier, la différence est de -2,89073, ce qui signifie que la prévision était supérieure à la valeur réelle par cette marge. Cette tendance s'accroît avec le temps, la différence devenant de plus en plus négative jusqu'à atteindre -14,90191 en juillet.

Ces écarts croissants entre les prévisions et les valeurs réelles peuvent être mis en relation avec l'intervalle de confiance qui s'élargit avec le temps, comme observé dans les précédentes analyses graphiques.