

Master 1 Econométrie et Statistique, parcours Econométrie  
Appliquée

---

L'indice des prix à la consommation de janvier 1990 à  
décembre 2022

---

Technique de prévision et conjoncture. O-DARNÉ

DAHMANI Amel

JAMIN Mathilde

ONNO Lilou

30 avril 2024

## Résumé

Notre étude s'intéresse à l'indice des prix à la consommation entre janvier 1990 et décembre 2022. Notre série mensuelle a été dans un premier temps corrigée des points atypiques puis dans un second temps désaisonnalisée et décomposée à l'aide de la méthode X13-ARIMA-SEATS dans l'objectif de réaliser des prévisions mensuelles sur la période de janvier 2023 à décembre 2023. Pour ce faire nous avons testé plusieurs modèles puis nous les avons comparés à l'aide d'indicateurs de précision tels que le MSE, CSSED et R2OOS. Nous les avons également confrontés grâce au test de précision de Diebold-Mariano. Le modèle ETS a été retenu comme étant le meilleur pour réaliser la prévision avec un pas de un mois.

## Abstract

Our study focuses on the consumer price index between January 1990 and December 2022. Our monthly series was first corrected for atypical points and then deseasonalized and decomposed using the X13-ARIMA-SEATS method in order to produce monthly forecasts for the period February 2023 to January 2024. To do this, we tested several models and compared them using precision indicators such as the MSE, CSSED and R2OOS. We also compared them using the Diebold-Mariano accuracy test. The ETS model was selected as the best for forecasting with a one-month time step.

# Sommaire

Introduction.....	3
I- Analyse exploratoire.....	5
II- Désaisonnalisation et décomposition.....	15
III- Prévion de la série saisonnière corrigée.....	17
IV- Représentation graphique de l'évolution des prévisions.....	26
V- Calcul de la qualité de prévision.....	28
VI- Test de précision : Le test de Diebold-Mariano.....	31
VII- Prévion sur une année avec un pas de 1 mois.....	32
Conclusion.....	40
Discussion.....	42
Annexes.....	43
Bibliographie.....	47
Table des matières.....	49

# Introduction

L'indice des prix à la consommation est un indicateur essentiel de l'évolution du coût de la vie et de la santé économique d'un pays, largement surveillé à la fois par les décideurs politiques et les entreprises afin d'anticiper les tendances économiques et d'ajuster en conséquence les politiques budgétaires.

En analysant cette série temporelle sur la période de janvier 1990 à décembre 2022, notre objectif est de dévoiler les tendances passées et d'élaborer des prévisions pour l'avenir. Cette modélisation nous permettra, à partir des informations sur les fluctuations de l'indice des prix à la consommation en France au cours des décennies écoulées, de mettre en place une meilleure anticipation des évolutions économiques pour l'année 2023 à l'aide de diverses méthodes de prévision.

Pour ce faire, nous allons suivre différentes étapes.

Tout d'abord, nous entamerons une analyse exploratoire des données pour décrire en détail les caractéristiques de l'IPC, y compris sa tendance, ses pics importants, ses variations saisonnières et toute valeur atypique éventuelle. Cette étape est cruciale pour mieux comprendre la série étudiée et pour préparer les étapes suivantes de manière plus juste.

Ensuite, nous procéderons à la désaisonnalisation et à la décomposition de notre série afin de séparer les différentes composantes. Nous utiliserons des méthodes spécifiques pour corriger les variations saisonnières.

La section suivante portera sur la prévision de la série saisonnière corrigée. Nous mettrons en œuvre plusieurs modèles de prévision, comme les modèles ARIMA et SARIMA, pour anticiper l'évolution de l'indice des prix à la consommation.

Nous illustrerons ensuite graphiquement l'évolution des prévisions obtenues par nos modèles, en comparant les résultats aux données réelles sur différentes périodes.

Par la suite, nous évaluerons la qualité de nos prévisions en utilisant des mesures standards telles que l'erreur quadratique moyenne et l'erreur absolue moyenne, ce qui nous donnera une vision quantitative de la performance de nos modèles.

Enfin, nous procéderons à un test approfondi de précision à l'aide du test de Diebold-Mariano, comparant les performances de nos modèles de prévision.

Ce rapport se conclura par une projection sur une année à venir avec un pas de un mois, permettant ainsi de visualiser les tendances de l'indice des prix à la consommation.

# I- Analyse exploratoire

## A) Présentation et caractérisation de la série étudiée

Notre série d'étude, trouvée sur le site de l'INSEE,<sup>1</sup> représente l'indice des prix à la consommation pour l'ensemble des ménages en France de l'année 1990 à 2022. Cette série est mise à jour chaque mois et permet d'analyser la variation moyenne des prix des produits consommés par les ménages entre deux périodes données.<sup>2</sup>

Les récentes polémiques entourant l'augmentation de l'inflation en France soulignent l'importance cruciale de surveiller l'indice des prix à la consommation, nommé aussi IPC, qui est par ailleurs un indicateur clé de la santé économique du pays. L'IPC est un indicateur essentiel puisqu'il va nous permettre d'avoir une mesure de l'évolution du coût de vie des ménages dans un pays.

Suite à la crise du COVID en 2020, la France a fait face à une période d'inflation très forte. Les autorités publiques ont réagi face à cela en mettant en œuvre des politiques économiques et monétaires visant à contenir cette inflation tout en encourageant la consommation et les investissements dans le but de stimuler la croissance.<sup>3</sup> Tout cela a été nécessaire pour éviter que le pouvoir d'achat se détériore et pour maintenir une certaine stabilité économique.

Bien analyser la tendance de l'indice des prix à la consommation est essentiel autant pour les décideurs politiques que pour les entreprises, car cela va permettre d'anticiper les tendances économiques et d'adapter les politiques budgétaires et monétaire en conséquence. A titre d'exemple, une augmentation soutenue de l'IPC peut indiquer des pressions inflationnistes, ce qui peut engendrer l'obligation de prendre des mesures appropriées pour maintenir une certaine stabilité économique et en outre le bien-être des ménages.

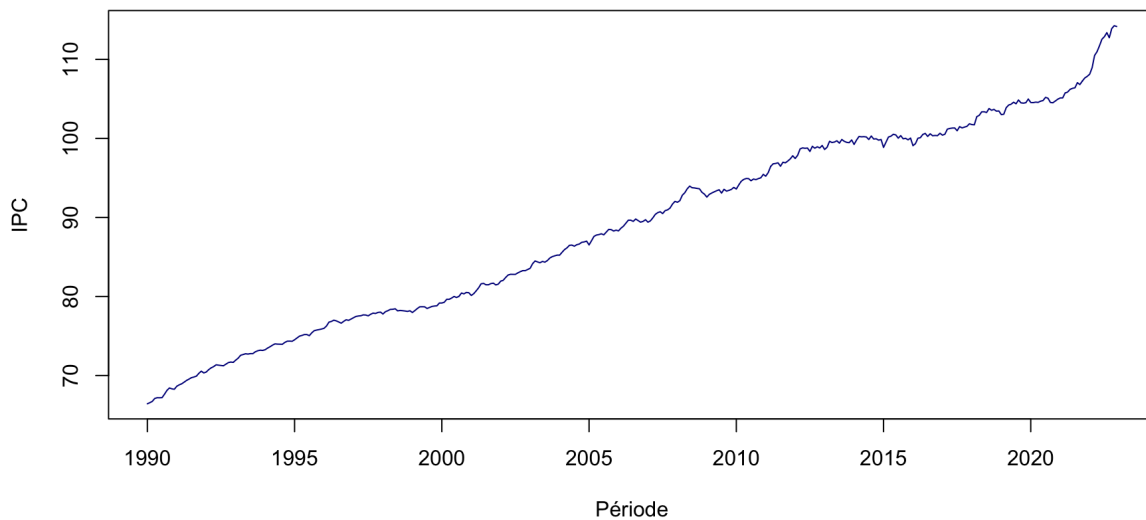
---

<sup>1</sup> INSEE (17-02-2024), Indice des prix à la consommation - Base 2015.

<sup>2</sup> INSEE (25-04-2024), Indice des prix à la consommation, IPC.

<sup>3</sup> Banque de France (10-01-2022), Covid-19 et économie, les clés pour comprendre.

**Graphique de la série brute Indice des prix à la consommation**



Graphique 1 : Série temporelle brut de l'indice des prix à la consommation.

L'indice des prix à la consommation a continuellement augmenté depuis 1990, en grande partie en raison de plusieurs facteurs. Tout d'abord, les prix en général ont augmenté au fil du temps. Par exemple, le salaire minimum horaire brut (SMIC) est passé de 6,67 euros à 11,65<sup>4</sup> euros en l'espace de 20 ans, mais le pouvoir d'achat n'a pas évolué de manière significative, voire a diminué.<sup>5</sup> Les politiques monétaires expansionnistes ont également contribué à cette tendance, tout comme l'augmentation constante de la demande, qui peut être liée au vieillissement de la population française, entraînant des besoins accrus en soins de santé et autres services.<sup>6</sup>

Il est important de noter quelques changements significatifs dans cette évolution. Par exemple, en 1999, nous observons un changement de pente plus abrupt, probablement dû à l'introduction de l'euro en France, ce qui a entraîné des ajustements de prix, notamment pour des produits de consommation courante comme le pain.<sup>7</sup>

---

<sup>4</sup> INSEE (15-02-2024), Salaire minimum interprofessionnel de croissance (Smic).

<sup>5</sup> INSEE (2024), Tableau de bord de l'économie française en 2021.

<sup>6</sup> Vie publique (03-06-2021), Vieillesse de la population : une adaptation nécessaire.

<sup>7</sup> La finance pour tous (07-07-2023), Inflation en France avant et après l'euro.

De même, un pic notable autour de 2008 peut être attribué à la crise des subprimes,<sup>8</sup> une crise financière mondiale qui a entraîné une instabilité économique et a impacté les prix à la consommation en France.

Certains pics négatifs vers 2014-2015 peuvent s'expliquer par des révisions à la hausse des salaires<sup>9</sup> qui n'ont pas été accompagnées d'une augmentation proportionnelle des prix, entraînant une amélioration du pouvoir d'achat pour les ménages.

Enfin, une croissance positive remarquable à partir de 2020 est due à la crise de la COVID-19 et aux tensions géopolitiques telles que le conflit entre l'Ukraine et la Russie. Ces événements ont créé des pressions inflationnistes en raison d'une forte demande et de problèmes d'approvisionnement dans divers secteurs. En 2021 et 2022, nous observons une progression continue, alimentée par une demande soutenue et des défis persistants en matière d'approvisionnement, contribuant ainsi à des hausses significatives des prix à la consommation.<sup>10</sup>

---

<sup>8</sup> Ministère de l'économie des finances et de la souveraineté industrielle et numérique (2013), Comment la crise a-t-elle commencé ? *La finance pour tous*.

<sup>9</sup> Le Point (16-06-2017), France : la croissance revue à la hausse en 2014 et à la baisse en 2015.

<sup>10</sup> Statista (15-04-2024), Indice des prix à la consommation et inflation en France - Faits et chiffres.



## B) Détection des points atypiques.

Outliers	Type	Ind	Time	Coefhat	tstat
1	AO	237	2020:07	0.5048	4.403
2	LS	387	2022:03	0.7994	4.132

Tableau 1 : Détection des outliers pour la série temporelle brut par tsoutliers.

La fonction *tso* du package *tsoutliers* nous a permis de détecter deux points atypiques dans notre série de données. Nous distinguons un premier point atypique AO (Additive Outlier) qui impacte une seule observation à un moment précis, *t*, puis un second point atypique noté LS (Level Shift), qui lui, signale un changement structurel dans la série (cf annexe 1).

D'abord, l'année 2020 est significative en raison des conséquences liées à la pandémie de COVID-19. En effet, selon l'INSEE, sur un an, l'indice des prix à la consommation augmente de 0,8% en juillet 2020 contre +0,2% le mois précédent ce qui est relativement important au vu des tendances précédentes.<sup>11</sup> Cette hausse intervient en période de déconfinement total mais aussi au début des vacances estivales, ces deux facteurs sont les raisons de l'augmentation de la consommation ainsi de la hausse de l'IPC chez les français.

Nous remarquons un second point atypique, en mars 2022. D'après l'INSEE, sur cette période, l'IPC a augmenté de 4,5% sur un an. Cette hausse résulterait principalement de l'invasion de l'Ukraine par la Russie, ce qui a déclenché une guerre fin février 2022, ce qui a entraîné en France, une augmentation des prix de l'énergie, de l'alimentation et dans une moindre ampleur des services.<sup>12</sup>

---

<sup>11</sup> INSEE (31-07-2020), Indice des prix à la consommation - résultats provisoires (IPC) - juillet 2020.

<sup>12</sup> INSEE (31-03-2022), Indice des prix à la consommation - résultats provisoires (IPC) - mars 2022.

En ce qui concerne le t-statistique de nos deux valeurs atypiques, nous remarquons qu'elles sont élevées avec respectivement 4.403 et 4.132, le seuil d'importance étant fixé à 3, cela indique que nos outliers ont un impact important sur notre série. Par conséquent, il faut ajuster la série.

### C) Statistiques descriptives sur la série corrigée

<b>Minimum</b>	66.42
<b>1er quartile</b>	78.10
<b>Médiane</b>	89.45
<b>Moyenne</b>	88.82
<b>3eme quartile</b>	99.86
<b>Maximum</b>	113.31
<b>Variance</b>	148.29
<b>Ecart-type</b>	12.17759
<b>Skewness</b>	-0.04254325
<b>Kurtosis</b>	1.781468

Tableau 2 : Statistique descriptives de notre série corrigée.

La moyenne de notre série étant de 88.82 se rapprochant de la médiane suggérant que notre série de données est conforme et relativement bien équilibrée. De plus, avec un minimum de 66.42 et un maximum de 113.31, l'étendue de la série est de 46.89, ce qui peut montrer la tendance croissante de notre série temporelle.

La skewness qui est un indicateur de la symétrie de la distribution est négative, ce qui indique que la majorité des observations se trouvent à gauche de la distribution, ce qui est confirmé par l'histogramme (cf annexe 3).

La kurtosis qui mesure l'aplatissement de la distribution est quant à elle, positive, cela indique que la distribution est par conséquent leptokurtique, ce qui signifie que cette dernière est moins aplatie qu'une distribution normale (cf annexe 3).

Tests	p-value
Shapiro-Will	6.627e-10
Kolmogorov-Smirnov	5.929e-05

Tableau 3: Tests de normalité sur la série corrigée.

Nous pouvons conclure sur la base du tableau 3, que les p-values pour les deux tests sont inférieures à 5%, ce qui démontre le rejet de l'hypothèse nulle et ainsi la non normalité de notre distribution. Cette non normalité peut-être observée sur l'histogramme de notre série (cf annexe 3).

## D) Détection de la saisonnalité et schéma de décomposition (additif ou multiplicatif)

Nous pouvons à travers le graphique (cf. annexe 4) de notre série mensuelle brute de l'Indice des prix à la consommation, suspecter des tendances saisonnières traduites par certains pics qui se répercutent de façon régulière. Pour s'en assurer, nous allons effectuer le test Seasonal Dummies qui correspond au F test sur les variables dichotomiques saisonnières, ainsi que le test de Webel-Ollech, WO.

Le premier consiste en l'introduction de variables indicatrices saisonnières dans le modèle, permettant de capturer les effets saisonniers.<sup>13</sup> Le second combine les résultats des tests QS et KW, effectués sur les résidus d'un modèle ARIMA non saisonnier. Si la valeur p du test QS est inférieure à 0,01 ou si celle du test KW est inférieure à 0,002, le test WO considérera la série temporelle comme saisonnière.<sup>14</sup>

Dans les deux cas, l'hypothèse nulle est l'absence de saisonnalité dans la série chronologique.

Tests	Test statistic	p-value
Seasonal Dummies	20.7	0
Webel-Ollech	1	0 0 0

Tableau 4 : Résultats des tests de détection de saisonnalité.

Les résultats des deux tests effectués pour lesquels la p-value est égale à 0, nous indiquent que notre série présente sans aucun doute des tendances saisonnières. Il est donc nécessaire de procéder à la désaisonnalisation de la série. Nous désaisonnaliserons cette dernière à l'aide de la méthode X13-ARIMA-SEATS.

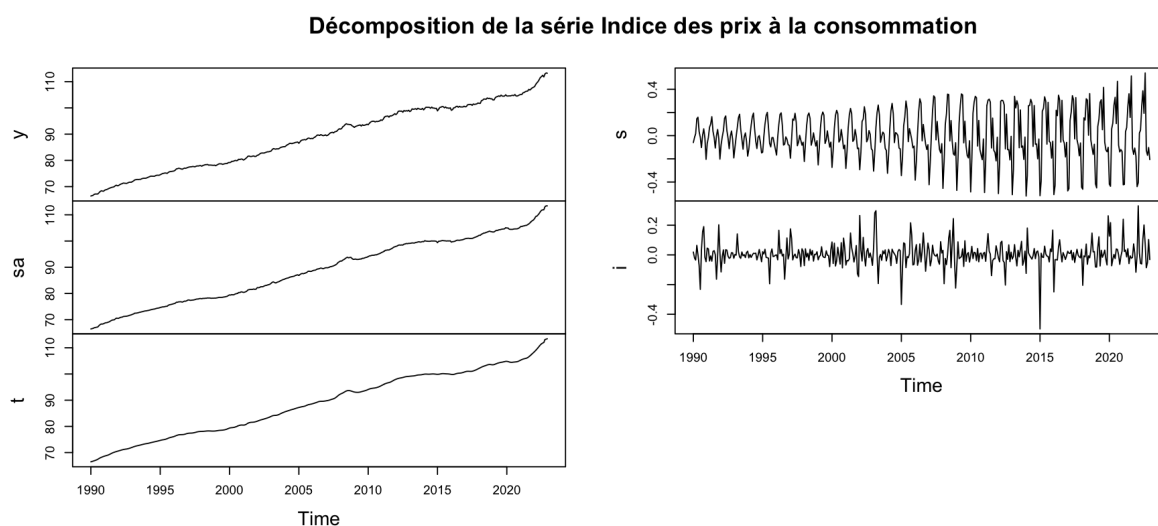
---

<sup>13</sup> SAS (11-08-2020), Mannequins saisonniers, *Utilisation de variables de prédiction*.

<sup>14</sup> Ollech.D, Test de saisonnalité combiné d'Ollech et Webel, *R documentation*.

Puisque la série présente des tendances saisonnières, il est pertinent de la décomposer afin de mieux comprendre ses composantes. La décomposition nous permettra de déterminer si le schéma de la série est additif ou multiplicatif et donc de modéliser et étudier la série de manière adéquate.

Pour déterminer le type de schéma, nous visualiserons la partie saisonnière d'un point de vue graphique. Dans un second temps, nous utiliserons la fonction *regarima\_x13* du package *RJDemetra* qui affiche le contenu de la sortie principale de la méthode X13-SEATS-ARIMA et la fonction *s\_transform()* qui affiche le résultat du test de log-level.



Graphique 2 : Décomposition de la série indice des prix à la consommation.

La décomposition de la série mensuelle de la figure 1 nous confirme la présence de saisonnalité au sein de la série. De plus, lorsque nous nous intéressons à la partie saisonnière, nous constatons que la série semble additive puisque nous n'observons pas d'amplification des variations de la courbe.

Toutefois, il est pertinent de vérifier cela à l'aide de la fonction *s\_transform()* qui affiche le résultat du test de log-level.

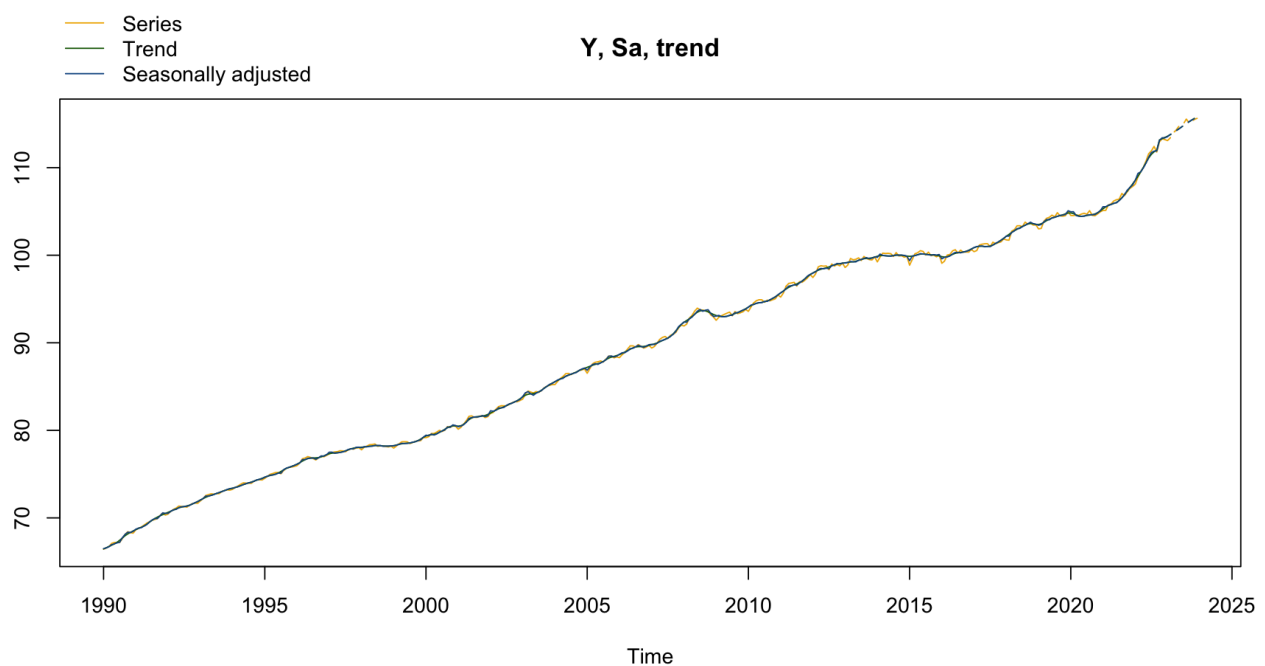
<b>tfunction</b>	<b>adjust</b>	<b>aicdiff</b>
Auto	None	-2

Tableau 5 : Test de log-level.

Les résultats présents au sein du tableau 5 nous indiquent à travers “None” qu’il n’y a pas de test de l’effet des années bissextiles et par la valeur “-2” qu’il y a une différence d’AICc entre les 2 spécifications AICc(level)-AICc(log). L’aicdiff étant inférieur à 0, la transformation en log n’est pas nécessaire. Le schéma de décomposition est donc additif.

## II- Désaisonnalisation et décomposition

Au sein de cette section, nous avons désaisonnalisé notre série de données en utilisant la méthode X13-ARIMA-SEATS. Cette approche nous permettra de corriger les variations saisonnières pour mieux mettre en évidence la tendance sous-jacente et les cycles inhérents à la série.



**Graphique 3** : Visualisation de la série ajustée par rapport à la saisonnalité.

Selon la méthode X13-ARIMA-SEATS, le meilleur modèle ajusté pour notre série est le modèle  $ARIMA(1,1,0)(0,1,1)_{12}$ . Ceci indique que la série nécessite une différenciation première ainsi qu'une différenciation saisonnière. De plus, le modèle intègre un terme autorégressif d'ordre 1 pour la composante non saisonnière AR, et un terme de moyenne mobile d'ordre 1 pour la composante saisonnière SMA.



La partie Auto Regressive correspondant à AR garantit que nous sommes capables de prédire les valeurs futures avec uniquement les valeurs précédentes. La prédiction à la date  $t+1$  va dépendre des valeurs à  $t$ ,  $t-1$  jusqu'à  $t-p$ . La Moving Average, quant à elle, permet d'intégrer les erreurs de prédiction de notre modèle pour effectuer des prédictions à la date  $t$ .<sup>15</sup>

Le modèle ARIMA est conçu pour capter à la fois les dynamiques internes de la série et les effets déterministes externes. Nous avons ici la présence de Easter effect, un effet déterministe externe qui prend compte des variations de la série temporelle qui pourraient être attribuées à la période de Pâques. Dans le cas de l'indice des prix à la consommation, cet effet déterministe est important puisque la série est influencée par le comportement des consommateurs et certains cycles commerciaux et événements spécifiques.

---

<sup>15</sup> Girard.D, Séries temporelles univariées, Semestre 2.

### III- Pr vision de la s rie saisonni re corrig e

#### A) Les m thodes param triques, semi param triques et non param triques.

Les m thodes param triques, semi param triques ainsi que non param triques comportent diff rents mod les avec par exemple la m thode na ve, X13-ARIMA-SEATS, STL ou encore STS. Ce sont ces m thodes qui vont  tre pr sent es et estim es par la suite.

La pr vision na ve est la m thode la plus simple des pr visions temporelles puisqu'elle consiste   extrapoler la derni re valeur observ e de la s rie chronologique dans le futur.

La m thode de pr vision X13-ARIMA-SEATS mod lise la s rie initiale  $X_t$  par un processus ARIMA pour ajuster les tendances et les saisons des donn es temporelles, permettant ainsi des pr visions pr cises en tenant compte des fluctuations saisonni res et des tendances   long terme des s ries chronologiques. Cette s rie est  galement prolong e en d but et fin de s rie, limitant ainsi les r visions des estimations, notamment lorsque nous disposons d'une observation suppl mentaire.

STL ou r gression polynomiale locale est une m thode non param trique de d composition des s ries chronologiques avec pond ration locale, bas e sur la m thode des  $k$  plus proches voisins. Elle s pare une s rie temporelle en trois composantes principales : la tendance, la saisonnalit  et les r sidus. Cette approche est particuli rement efficace pour d tecter les tendances et fluctuations saisonni res, facilitant l'analyse et la pr vision des s ries chronologiques.

L'approche Structural Time Series (STS) de mod lisation des s ries temporelles consiste   sp cifier directement des mod les ARIMA pour chaque variable inobservable. Elle repr sente alors les donn es comme une combinaison de composantes non

observées, telles que la tendance, la saisonnalité, les effets de calendrier ou encore les erreurs aléatoires. Elle s'inscrit dans la continuité des méthodes de régression outre le fait qu'au lieu d'utiliser des modèles déterministes pour estimer chacune des composantes, elle utilise des modèles stochastiques simples appartenant à la classe ARIMA.<sup>16</sup>

Modèle	Naïve	X13-ARIMA-S EATS (1, 1, 1, 0, 1, 1)	STL	STS
alpha			0.9999	
beta			0.1266	
sigma			0.0018	
gamma				
Phi		-0.88422		
Theta		-0.72111		
BTheta		-0.67530		
lambda				
residual sd	0.3078			
level				0.0229501
slope				0.0006372
seas				0.0012993
epsilon				0.0014660
distribution	Normale	Normale	Normale	Normale
AIC		-202.2	899.2908	
AICc		-201.8	899.4442	

**Tableau 6** : Estimation des modèles paramétriques, semi paramétriques et non paramétriques.

<sup>16</sup> Darné.O, Technique de prévision et conjoncture, Semestre 2.

Les modèles estimés présents au sein du tableau 6 ont tous quatre une distribution normale. La méthode X13-ARIMA-SEATS a estimé trois coefficients que sont Phi, Theta et BTheta correspondant respectivement aux paramètres AR(1), MA(1) et SMA(1). Ce dernier aussi présente une différence première ainsi qu'une différence saisonnière.

Le modèle STL présente quant à lui deux paramètres que sont alpha et bêta respectivement de 0.9999 et 0.1266, suggérant que la composante de tendance est très lissée, avec peu de réactivité aux variations à court terme, tandis que la saisonnalité est légèrement lissée, ce qui la rend relativement stable mais avec une certaine réactivité aux fluctuations saisonnières. Le paramètre sigma qui est faible représente l'écart-type estimé du terme d'erreur.

Le modèle STS comprend quatre paramètres. Tout d'abord, le level correspond au niveau qui représente la composante de la série temporelle qui capture le niveau moyen des données. Ensuite, la slope correspondant à la pente qui indique si les données ont une tendance croissante ou décroissante. Par la suite, la seas faisant référence à la saisonnalité, cette dernière capture les variations saisonnières, la valeur indique l'amplitude de ces variations. Enfin, epsilon correspond aux résidus après avoir retiré les composantes de niveau, de tendance et saisonnières. Il représente l'erreur de prédiction c'est-à-dire les fluctuations qui ne peuvent pas être expliquées par les composantes précédentes. Les valeurs des paramètres sont relativement faibles notamment epsilon ce qui suggère qu'il n'y a pas d'erreur des résidus significative.<sup>17</sup>

S'il convenait de choisir un meilleur modèle au sein de ce tableau, ce dernier serait X13-ARIMA-SEAT sur la base du critère AIC et AICc puisque ce dernier est meilleur.

---

<sup>17</sup> Darné.O, Technique de prévision et conjoncture, Semestre 2.

## B) Les méthodes de lissage exponentiel.

La méthode ETS propose des versions sous forme d'espace d'état de la méthode saisonnière de Holt-Winters (méthode légèrement différente du lissage exponentiel double adaptée aux séries sans saisonnalité pouvant être ajustées par une tendance) en introduisant un terme d'erreurs  $\epsilon$ . Ces modèles ETS offrent une approche flexible et robuste pour modéliser les tendances, les saisons et les erreurs dans les séries temporelles, que ce soit de manière additive ou multiplicative.

La méthode TBATS étend les modèles ETS en introduisant des ajustements pour traiter les comportements saisonniers non linéaires, l'autocorrélation des résidus et une plus grande variété de structures saisonnières. En incorporant des composantes trigonométriques pour modéliser les saisons ainsi qu'une transformation de Box-Cox pour stabiliser la variance, des erreurs ARMA pour capturer les corrélations résiduelles et une plus grande flexibilité dans la modélisation des tendances et des saisons, la méthode offre une approche puissante pour la prévision précise des séries temporelles.

L'extension des modèles ETS avec les modèles ADAM offre une flexibilité accrue en permettant l'utilisation de distributions alternatives pour le terme d'erreur, telles que la distribution Laplace, Log-Normale, Gamma, etc, au lieu de se limiter à une distribution normale comme dans les modèles ETS conventionnels. De plus, dans le cadre du modèle ADAM, il est possible de combiner les modèles ARIMA et ETS, ce qui permet une modélisation plus robuste et adaptative des séries temporelles, en prenant en compte à la fois les tendances, les saisons et les corrélations résiduelles complexes.

Le modèle SSARIMA offre une approche performante pour la modélisation des séries temporelles, notamment pour les petits échantillons. Il se distingue par sa flexibilité accrue, permettant de capturer plusieurs saisonnalités et d'ajuster les modèles ARIMA dans un espace d'état, ce qui offre une meilleure adaptabilité aux structures complexes des données temporelles, améliorant ainsi la précision des prévisions.

Les modèles de lissage exponentiel complexe (CES) sont fondés sur la théorie des variables complexes. Ils offrent une approche novatrice pour la modélisation des séries temporelles, permettant de traiter à la fois les séries stationnaires et non stationnaires, contrairement aux modèles ETS qui se limitent généralement aux séries non stationnaires. Cette flexibilité accrue permet une meilleure adaptation aux différentes structures des données chronologiques, améliorant ainsi la capacité de prédiction des modèles CES.<sup>18</sup>

Modèle	Holt-Winters	ETS(M,A,A)	TBATS	ADAM ETS(M,Ad,M)	ADAM ETS (A,M,N) + SARIMA (1,0,0)_12	SSARIMA(2,1,0)(3,0,3)_12	CES
alpha	0.7313084	0.8429	0.217077	0.8064	1		
beta	0.07109729	0.0993	0.008545095	0.0876	0.0709		
gamma	1	0.1571		0.1892			
gamma 1			5.053085e-05				
gamma 2			-0.000114258				
sigma			0.002310649				
phi				1	0.1667		
lambda			0.001654				
ar1			-0.106464			0.770	
ar2			0.874886			0.0982	
ar3						0.1267	
ma1			0.921381			-0.4506	
ma2			-0.1659			-0.0889	

<sup>18</sup> Darné.O, Technique de prévision et conjoncture, Semestre 2.

			56				
ma3			-0.0177 9			-0.0564	
ma4			0.11576				
distribu tion	Normal e	Normal e	Normal e	Normale Généralisé	Inverse Gaussian	Normale	full
AIC		1104.88 4	1159.17 4	-169.9916	-154.5853	-102.1848	-111.2088
AICc		1106.50 3		-167.9703	-152.7710	-101.7185	-111.0550

Tableau 7 : Estimation des méthodes de lissage exponentiel.

Le tableau 7 nous permet de comparer les différentes méthodes de lissage exponentiel. Les modèles Holt-Winters, ETS, TBATS, ADAM ETS et SSARIMA présentent tous cinq une distribution normale tandis que le modèle ARIMA ETS SARIMA présente une distribution inverse gaussienne de probabilité continue.

Pour l'approche CES, nous avons pour type de saisonnalité *full*, ce qui correspond à une saisonnalité complète avec des composantes saisonnières complexes. Cette dernière offre une flexibilité maximale en permettant d'incorporer à la fois des effets saisonniers multiplicatifs et additifs dans le modèle, selon les caractéristiques des données. Cette approche avancée garantit une meilleure capture des variations saisonnières complexes, ce qui peut conduire à des prévisions plus précises et plus adaptées aux particularités de la série temporelle étudiée.

Concernant le modèle Holt-Winters, nous avons trois paramètres : alpha, bêta et gamma. Le paramètre alpha est proche de 1, ce qui indique que les valeurs les plus récentes ont un poids plus important que les autres. Pour bêta, sa proximité avec 0 suggère que la méthode de lissage n'accorde pas une grande importance à la tendance de la série. Enfin, gamma est égal à 1, ce qui signifie que notre modèle donne plus de

poids aux variations saisonnières récentes, rendant la composante de saisonnalité plus réactive aux changements saisonniers à court terme.

Notre modèle ETS présente des paramètres  $\alpha$ ,  $\beta$  et  $\gamma$  de ce modèle. Le paramètre  $\alpha$  qui est assez proche de 1 indique que le modèle réagit particulièrement aux nouvelles observations pour la composante de niveau. Le paramètre  $\beta$  qui est le paramètre de lissage pour la composante de tendance est proche de 0, il donne donc moins de poids aux observations les plus récentes. Le paramètre  $\gamma$  qui correspond au paramètre de lissage pour la composante saisonnière est également assez proche de 0, indiquant un ajustement plus lent aux changements saisonniers.

Dans le modèle TBATS, nous avons tout d'abord six paramètres :  $\alpha$ ,  $\beta$ ,  $\gamma_1$ ,  $\gamma_2$ ,  $\lambda$  et  $\sigma$ . Étant donné que  $\lambda$  n'est pas égal à 0, cela indique que notre modèle est non linéaire. Le paramètre  $\alpha$  est proche de 0, suggérant que les anciennes observations ont plus de poids dans le modèle. De même, le paramètre  $\beta$  est très proche de 0, ce qui signifie qu'il accorde moins d'importance aux observations les plus récentes. Les paramètres  $\gamma_1$  et  $\gamma_2$  contrôlent les effets de la saisonnalité multiplicative et additive. Une valeur positive de  $\gamma_1$  indique une saisonnalité multiplicative, où l'amplitude de la saisonnalité augmente proportionnellement avec le niveau moyen de la série temporelle. En revanche,  $\gamma_2$  ayant une valeur négative indique une saisonnalité dont l'amplitude diminue avec le niveau moyen de la série. Le paramètre  $\sigma$ , quant à lui, représente l'écart type des résidus du modèle, cette valeur est très faible dans notre cas.

Ce modèle comprend également six autres paramètres :  $AR(1)$ ,  $AR(2)$ ,  $MA(1)$ ,  $MA(2)$ ,  $MA(3)$  ainsi que  $MA(4)$ , ainsi il y a deux paramètres  $p$  et quatre paramètres  $q$ .



Le modèle ADAM ETS fournit quatre paramètres que sont : alpha, beta, gamma et phi. Avec une valeur de 0.8064, le paramètre alpha signifie que le modèle donne beaucoup de poids aux observations les plus récentes lors de la mise à jour de la composante de niveau. Le paramètre beta suggère, avec une valeur très proche de 0 que le modèle est assez prudent dans l'ajustement de la tendance. Il accorde moins de poids aux variations observées dans la tendance récente, ce qui peut indiquer une relative stabilité ou une préférence pour les tendances plus lentes à évoluer. Par ailleurs, avec une valeur de Gamma qui se rapproche de 0, le modèle réagit de manière assez rapide aux variations saisonnières. Enfin, Phi correspond au sein de ce modèle à un paramètre de régularisation introduit pour aider à stabiliser l'estimation de la saisonnalité. Une valeur de 1 pour ce dernier indique une régularisation assez forte de la composante saisonnière, ce qui peut être approprié dans certaines situations pour atténuer les fluctuations saisonnières brusques ou inhabituelles.

Le modèle combiné ADAM ETS met complètement à jour le niveau ( $\alpha=1$ ) avec une certaine prudence dans l'ajustement de la tendance puisque le paramètre  $\beta$  est assez proche de 0. Toutefois, SARIMA régularise la saisonnalité avec un paramètre  $\phi$  AR(1) de 0.1667.

Le modèle SSARIMA comprend six paramètres : AR(1), AR(2), AR(3), MA(1), MA(2), MA(3) au sein de la partie classique respectant les conditions de stationnarité pour AR et d'inversibilité pour MA puisque la somme des coefficients des paramètres est inférieure à 1.

Pour déterminer la meilleure méthode de lissage exponentiel, il convient de comparer l'AIC des différents modèles. Au vu de ce critère, le meilleur modèle serait SSARIMA.

### C) Le modèle SARIMA

Ce modèle a été estimé avec une transformation logarithmique.

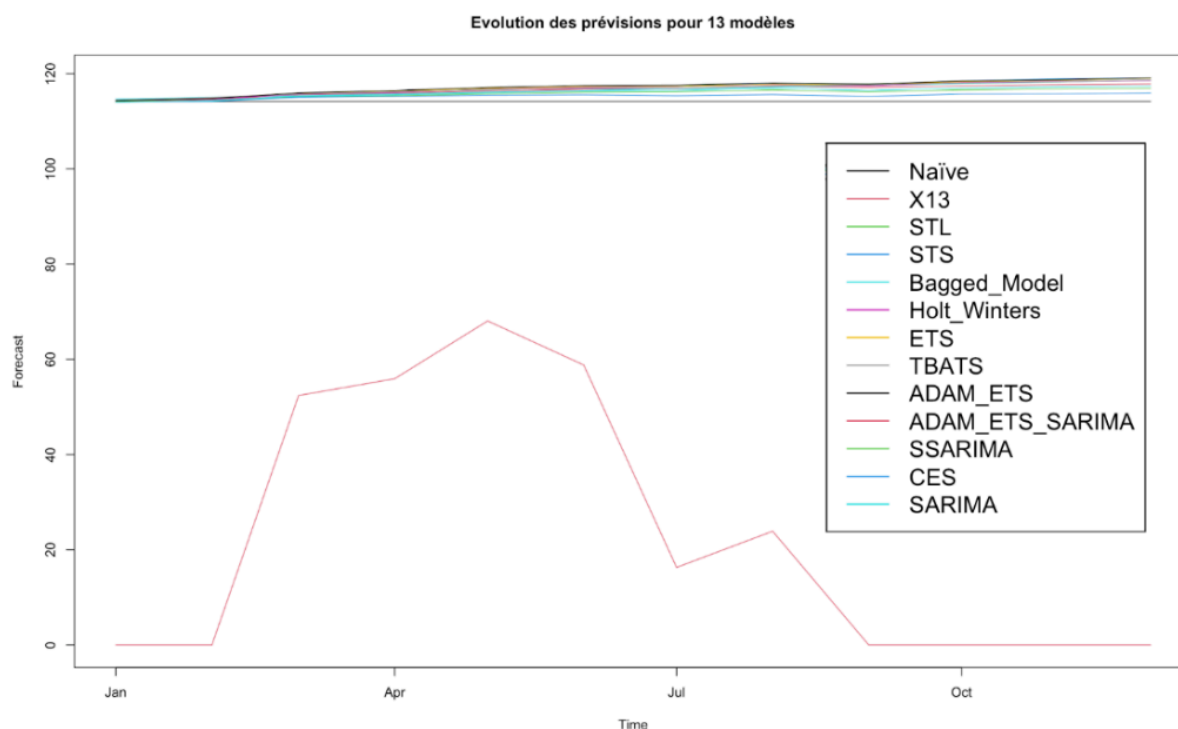
Modèle	SARIMA (0,1,1)(0,1,1)
Theta	0.15839
BTheta	-0.66850
AIC	-155.6
AICc	-155.5

Tableau 8 : Estimation du modèle SARIMA(p,d,q)(P,D,Q)<sub>12</sub>.

Le modèle SARIMA présente une différenciation première classique ainsi qu'une différenciation saisonnière. Deux paramètres sont estimés, MA(1) ainsi que SMA(1). Le critère d'AIC est plutôt faible, ce qui en fait un modèle satisfaisant.

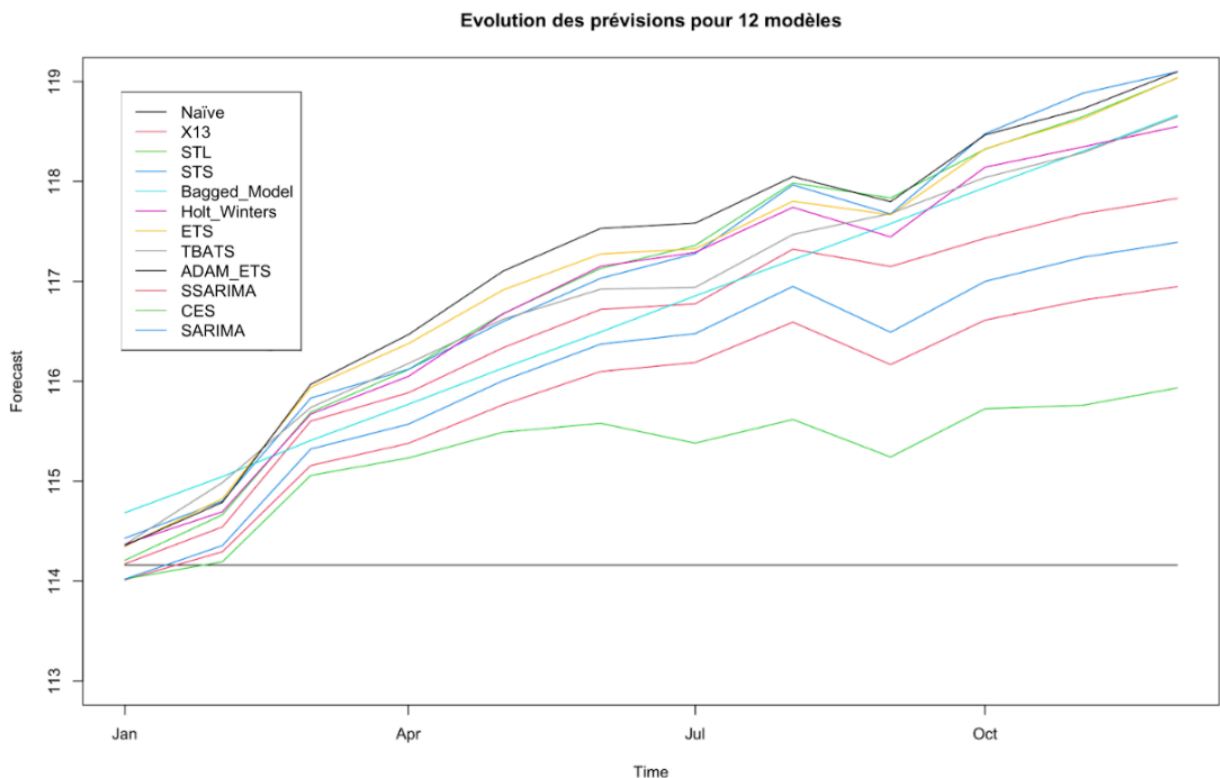
Sur l'ensemble des modèles estimés et sur la base du critère AIC et AICc, nous pouvons conclure que le meilleur modèle est SSARIMA.

## IV- Représentation graphique de l'évolution des prévisions



Graphique 4 : Evolution des prévisions pour tous les modèles.

Ce graphique met en évidence nos prévisions mensuelles pour l'année 2023, basées sur les différents modèles étudiés. Ces dernières suivent une tendance similaire, à l'exception du modèle ARIMA ETS SARIMA. Ce modèle prédit l'indice des prix à la consommation à 0 pour les mois de janvier et février, ainsi que pour la période de septembre à décembre. Cette observation soulève des questions quant à l'efficacité de ce modèle pour notre série. Une évaluation plus approfondie de la qualité des prévisions sera effectuée ultérieurement, en utilisant notamment le MSE et le R2OOS. Néanmoins, à la lumière de ce graphique, nous pouvons déjà exclure ce modèle. Par conséquent, nous allons générer un second graphique mettant en avant les prévisions des autres modèles, afin d'avoir une meilleure visualisation de leurs performances.



Graphique 5 : Evolution des prévisions à l'exception de ADAM ETS SARIMA.

Dans ce second graphique, nous constatons une cohérence générale parmi tous nos modèles, à l'exception de la méthode de prévision naïve, comme prévu. Nous pouvons identifier des regroupements au sein de nos prévisions. En effet, les prévisions générées par les modèles SARIMA, CES, SSARIMA, ADAM ETS présentent des similitudes. De même, une observation similaire peut être faite pour les modèles Holt-Winters, STS, et STL.

Nous constatons également que l'ensemble des modèles, à l'exception de TBATS, Bagged Model et Naïf ont un deux creux pour la prévision de juillet et septembre 2023 ainsi qu'un pic pour le mois d'août de cette même année.

Nous pouvons aussi remarquer que le Bagged Model croît de manière linéaire.

## V- Calcul de la qualité de prévision

	<b>MSE</b>	<b>CSSED</b>	<b>R200S</b>
Naïve	12.35876	0	0
X13-ARIMA-SEATS	1.167215	27.49464	0.9055556
STL	0.5916396	39.81257	0.9521279
STS	0.6018477	39.84123	0.9513019
Bagged model	1.102663	56.59015	0.9103713
Holt-Winters	0.6423258	36.64513	0.9480267
ETS	0.4710399	41.76266	0.9618861
TBATS	0.6597915	35.11355	0.9466134
ADAM ETS	0.3831946	46.23058	0.9689941
ADAM ETS + SARIMA	9641.822	52955.42	-779.1709
SSARIMA	2.859516	15.0728	0.7686244
CES	5.314774	7.231639	0.5699589
SARIMA	2.031381	19.99988	0.8356323

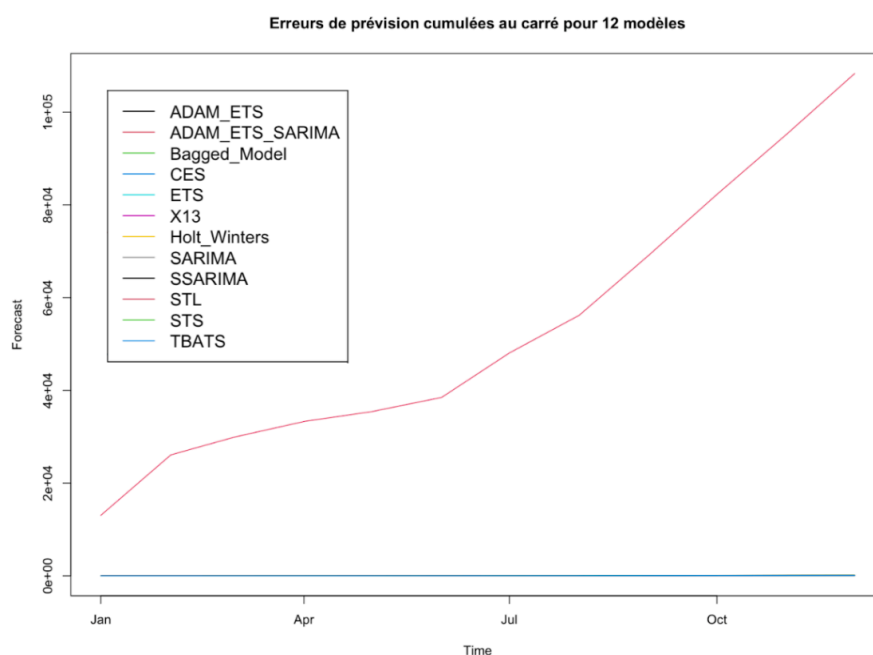
Tableau 9 : Qualité de prévision de nos modèles.

Dans cette présentation, une variété de méthodes a été utilisée pour évaluer la précision des prévisions de nos modèles.

Nous avons commencé par l'utilisation du MSE, qui mesure l'erreur quadratique moyenne. Un MSE élevé indique une performance moins satisfaisante du modèle. Il est à noter qu'un modèle se distingue : ARIMA ETS SARIMA, comme prévu à partir du graphique 4, avec un MSE extrêmement élevé. En revanche, plusieurs modèles présentent un MSE satisfaisant : ADAM ETS, ETS, STL, STS, Holt-Winters ainsi que TBATS.

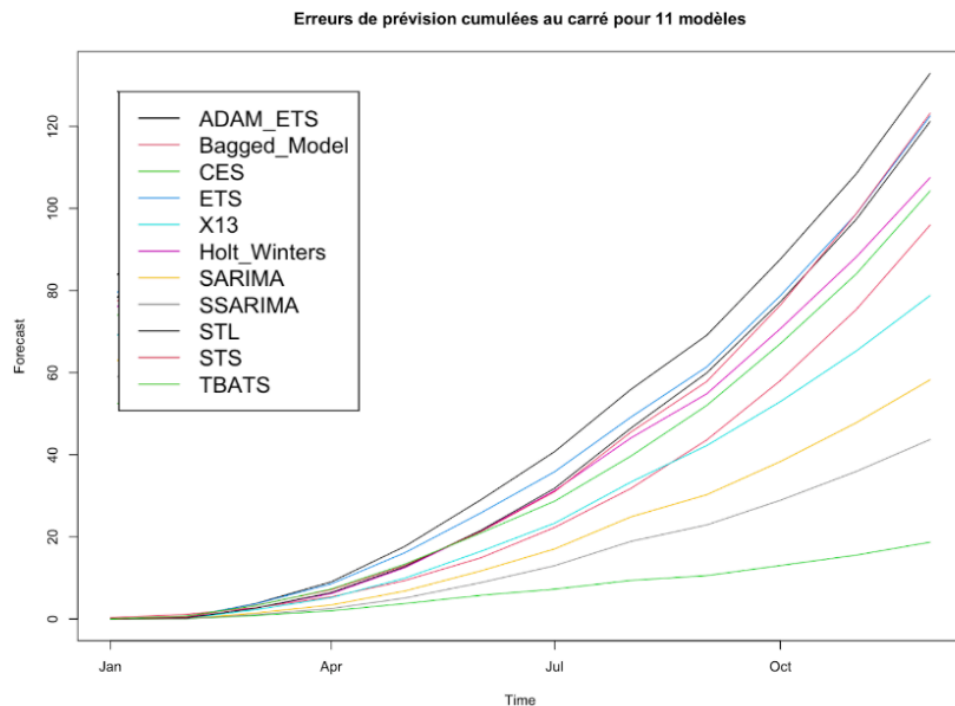
Nous avons ensuite continué notre analyse en utilisant le CSSED, qui mesure la somme cumulative des différences des erreurs au carré, permettant d'évaluer la stabilité de la précision des prévisions. Tout comme le MSE, un CSSED plus faible est préférable, et là encore, le modèle ADAM ETS SARIMA se démarque avec une forte valeur. En utilisant ce critère, plusieurs nouveaux modèles se distinguent par une faible valeur : SARIMA, Bagged Model, X13-ARIMA-SEATS et SSARIMA.

Enfin, pour conclure notre analyse, nous avons utilisé le R2OOS, qui compare les prévisions des différents modèles à un modèle de référence, naïf dans notre cas. Ainsi, par rapport à ce dernier, les deux meilleurs modèles seraient le ADAM ETS ainsi que ETS car ces derniers s'éloignent le plus de notre prévision naïve. Nous pouvons également souligner encore une fois que le modèle ADAM ETS SARIMA se distingue avec une forte valeur négative, cela signifie que ce dernier fait moins bien que la prévision naïve.



**Graphique 6 :** Erreurs de prévision cumulées au carré (CSPE) pour 12 modèles.

Sur ce graphique 6, nous pouvons observer comme attendu, que l'erreur de prévision cumulée au carré du modèle ADAM ETS SARIMA est très élevée ce qui en fait un modèle de mauvaise qualité qui se distingue des 11 autres modèles. Au vu de cette illustration, nous ne pouvons pas comparer les modèles entre eux ainsi, nous allons refaire ce graphique en retirant le modèle ADAM ETS SARIMA.



Graphique 7 : Erreurs de prévision cumulées au carré (CSPE) pour 11 modèles.

Le graphique 7 nous permet alors de comparer les différentes prévisions. Sur la base de ce dernier, les erreurs de prévision cumulées qui croissent moins exponentiellement représentent les meilleurs modèles. C'est le cas des modèles TBATS, SARIMA et SSARIMA qui par ailleurs présentent un R2OOS assez élevé, cela reste cohérent.

## VI- Test de précision : Le test de Diebold-Mariano

	DM test	DM p-value (x e <sup>-05</sup> )
X13 ARIMA SEATS	-5.9453	4.829
STL	-6.0062	4.423
STS	-6.1062	3.834
Bagged Model	-6.4746	2.292
Holt Winters	-6.0472	4.171
ETS	-6.2361	3.192
TBATS	-6.1994	3.361
ADAM ETS	-6.2274	3.231
ADAM ETS SARIMA	6.8942	1*
SSARIMA	-5.6322	7.64
CES	-5.9822	4.579
SARIMA	-5.7097	6.812

\*unique valeur qui n'est pas multipliée par e<sup>-05</sup>.

Tableau 10 : Test de Diebold-Mariano.

Le test de Diebold-Mariano permet de comparer la précision prédictive de deux méthodes de prévision. Ce test compare les performances prédictives des modèles en examinant les différences entre les erreurs de prévision produites par chaque modèle sur une période donnée. Ainsi, il teste l'hypothèse nulle selon laquelle les modèles ont la même précision prédictive autrement dit, une prévision naïve.

Nous pouvons nous focaliser sur la comparaison de nos modèles prédictifs par rapport à la prévision naïve. Nous pouvons remarquer qu'il y a pour tous nos modèles à l'exception du modèle ADAM ETS SARIMA le rejet de l'hypothèse nulle, cela signifie que la prévision pour nos différents modèles est différente de la prévision naïve.



## VII- Pr vision sur une ann e avec un pas de 1 mois

Nous explorerons deux m thodes distinctes de pr vision   un pas pour l'ann e 2023, en utilisant des techniques de s rie temporelle pour pr dire les valeurs futures. Nous utiliserons et comparerons les approches Rolling et R cursive, qui sont deux strat gies pr visionnelles populaires dans l'analyse des donn es temporelles.

Au vu de la partie pr c dente, nous avons d cid  de s lectionner comme meilleur mod le, le mod le ETS car ce dernier pr sente une des p.value les plus faibles. Nous avons fait le choix de ne pas retenir les mod les ARIMA ETS SARIMA et Bagged Model qui ont une p.value encore plus faible que le mod le ETS car, selon les diff rents graphiques et des crit res correspondant   la qualit  des pr visions, ces mod les n' taient pas satisfaisants.

### A) M thode rolling

La m thode Rolling repose sur la mise   jour continue du mod le de pr vision en utilisant une fen tre glissante de donn es historiques, ce qui permet d'int grer les informations les plus r centes tout en conservant une taille de fen tre constante.

Mois	Pr�vision
Janvier	114.3474
F�vrier	114.8349
Mars	115.8725
Avril	116.2729
Mai	116.7804
Juin	117.0930
Juillet	117.1195
Ao�t	117.5166

Septembre	117.3808
Octobre	118.0197
Novembre	118.3131
Décembre	118.7328

**Tableau 11** : Prévisions sur une année avec un pas de un mois avec la méthode rolling.

Le tableau 11 met en évidence les résultats de la prévision qui semblent assez proches des valeurs réelles. A présent, nous allons mettre en évidence la différence ainsi que les erreurs de prévision entre la prévision à un pas de la méthode rolling et les vraies valeurs.

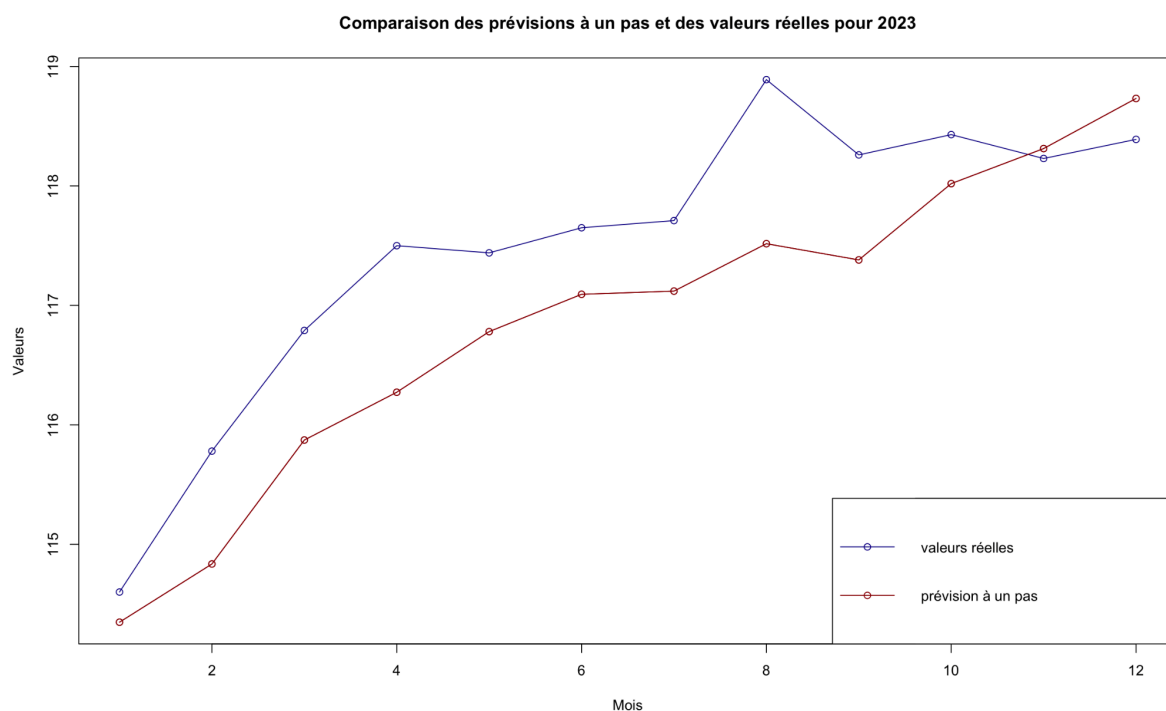
Mois	Différence	Erreur de prévision
Janvier	-0.25258169	-0.22%
Février	-0.94513117	-0.82%
Mars	-0.91748017	-0.79%
Avril	-1.22710756	-1.04%
Mai	-0.65955873	-0.56%
Juin	-0.55700911	-0.47%
Juillet	-0.59050049	-0.50%
Août	-1.37336414	-1.16%
Septembre	-0.87915881	-0.74%
Octobre	-0.41033612	-0.35%
Novembre	0.08309718	0.07%
Décembre	0.34279078	0.29%

**Tableau 12** : Comparaison des erreurs de prévisions avec la méthode rolling.

Le tableau 12 ci-dessus nous permet de rendre compte de la différence ainsi que de l'erreur de prévision entre la prévision à un pas effectuée à partir du meilleur modèle et les valeurs réelles de l'indice des prix à la consommation en 2023.

L'analyse des résultats révèle que, dans l'ensemble, la méthode du rolling fournit généralement des prévisions robustes malgré des variations dans la précision d'un mois à l'autre. La majorité des mois montrent des prévisions sous-estimées, avec des erreurs plus marquées en août et avril, tandis que novembre et décembre affichent des prévisions légèrement supérieures aux valeurs réelles.

Malgré ces fluctuations, les erreurs restent généralement faibles en pourcentage, indiquant une bonne adéquation du modèle aux données sous-jacentes. Cette méthode semble ainsi offrir une base solide pour la prévision continue des données temporelles.



**Graphique 8 :** Comparaison de la prévision avec la méthode rolling et les valeurs réelles.

Le graphique 8 compare les prévisions à un pas et les valeurs réelles sur l'année 2023. Les valeurs réelles, représentées en bleu, sont initialement plus élevées et continuent de surperformer les prévisions, qui sont en rouge, tout au long de l'année

sauf en novembre et décembre. Cela suggère que la méthode de prévision utilisée a tendance à sous-estimer les résultats réels mois après mois.

Quant à l'écart entre les deux courbes, au début de l'année, l'écart est relativement faible, mais il augmente progressivement jusqu'à atteindre un maximum autour des mois de juin et juillet. Après cela, bien que les deux courbes continuent de croître, l'écart commence à se réduire légèrement, suggérant une amélioration dans la précision de la prévision à un pas vers la fin de l'année.

## B) Méthode récursive

La méthode récursive construit le modèle à partir d'un ensemble initial de données et l'actualise progressivement en intégrant les prévisions antérieures comme si elles étaient de nouvelles observations, étendant ainsi la base de données à chaque pas de temps. Cette méthode adapte le modèle en continu pour refléter les changements dans les tendances sous-jacentes et améliorer la précision des prédictions futures.

Mois	Prévision
Janvier	114.3474
Février	114.8470
Mars	115.9310
Avril	116.3476
Mai	116.8640
Juin	117.2011
Juillet	117.2462
Août	117.7190
Septembre	117.5693
Octobre	118.2374
Novembre	118.5339
Décembre	118.9696

Tableau 13 : Prévisions sur une année avec un pas de un mois avec la méthode récursive.

Le tableau 13 représente les résultats de la prévision à un pas de la méthode récursive. Ces derniers semblent assez proches des valeurs réelles. Il convient maintenant d'examiner de près les écarts et les erreurs de prévision entre la méthode récursive à un pas et les valeurs réelles.

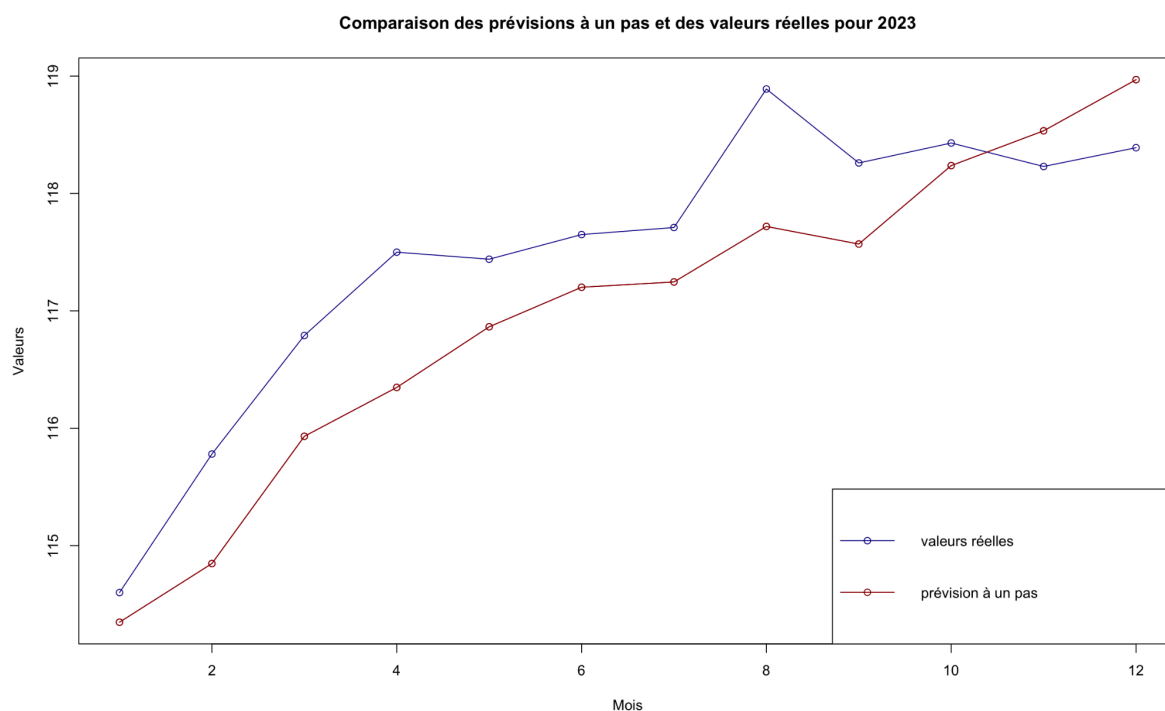
Mois	Différence	Erreur de prévision
------	------------	---------------------

Janvier	-0.2525817	-0.22%
Février	-0.9329709	-0.81%
Mars	-0.8589689	-0.74%
Avril	-1.1523963	-0.98%
Mai	-0.5760125	-0.49%
Juin	-0.4488861	-0.38%
Juillet	-0.4638259	-0.39%
Août	-1.1709565	-0.98%
Septembre	-0.6907039	-0.58%
Octobre	-0.1925609	-0.16%
Novembre	0.3038788	0.26%
Décembre	0.5796030	0.49%

Tableau 14 : Comparaison des erreurs de prévisions avec la méthode réursive.

L'analyse des différences et des erreurs de prévision entre la méthode réursive à un pas et les valeurs réelles présente au sein du tableau 14 révèle plusieurs tendances significatives.

Globalement, les différences entre les valeurs prévues et les valeurs réelles fluctuent d'un mois à l'autre, mais tendent à rester relativement proches de zéro, indiquant une certaine précision dans les prévisions. Cependant, certains mois présentent des écarts plus importants, notamment avril et en août, où les différences dépassent 1 unité. Les erreurs de prévision exprimées en pourcentage sont également relativement faibles, généralement inférieures à 1%, ce qui suggère une bonne précision globale du modèle.



**Graphique 9** : Comparaison de la prévision avec la méthode récursive et les valeurs réelles.

Sur ce graphique 9, l'évolution de l'écart entre les deux courbes montre une dynamique légèrement différente par rapport au graphique précédent.

Initialement, les prévisions à un pas et les valeurs réelles débutent avec un écart assez léger. Cet écart augmente légèrement et diminue légèrement au mois de juin, après quoi il commence à augmenter de nouveau. Les deux courbes semblent se croiser aux alentours d'octobre et novembre, mais divergent à nouveau vers la fin, indiquant que l'écart entre les prévisions et les valeurs réelles varie au fil des mois sans maintenir un parallélisme constant. Cela suggère des périodes où la méthode de prévision est moins précise par rapport aux valeurs réelles.

	<b>MSE</b>	<b>R2OOS</b>
ETS	0.4710399	0.9618861
ETS avec un pas (méthode rolling)	0.6125364	0.9504371
ETS avec un pas (méthode récursive)	0.5051728	0.9591243

Tableau 15 : Comparaison de la qualité des modèles ETS.

Les différents modèles ETS ont été évalués en fonction de leur performance prédictive à l'aide de l'erreur quadratique moyenne (MSE) et du coefficient de détermination hors échantillon (R2OOS). Sur la base des résultats fournis par le tableau 15, le modèle ETS utilisant la méthode récursive pourrait être considéré comme étant légèrement meilleur, car il affiche une MSE plus faible et un R2OOS légèrement plus élevé par rapport au modèle ETS utilisant la méthode du pas roulant.



## Conclusion

L'objectif de cette étude était de modéliser et de prévoir notre série temporelle concernant l'indice des prix à la consommation. Pour ce faire, nous avons, dans un premier temps, désaisonnalisé la série à l'aide de la méthode X13-ARIMA-SEATS ainsi que corrigé cette dernière des points atypiques. Par la suite, nous avons entrepris différents modèles afin de prévoir notre série sur une période de un an, nous avons comparé ces modèles à l'aide de différents outils tel que le MCE, le CSSED ou encore le R2OOS, nous avons complété cette analyse concernant la qualité de prévision par un test de précision de Diebold-Mariano. Enfin, nous avons effectué une prévision sur une année avec un pas de un mois à l'aide d'une boucle.

L'analyse des différents modèles de prévision appliqués à notre série temporelle nous a permis de déterminer que le modèle SSARIMA est le plus approprié selon les critères AIC et AICc. De plus, l'examen des prévisions pour l'année 2023 met en lumière des performances cohérentes parmi plusieurs modèles, confirmant ainsi la fiabilité du modèle SSARIMA. En comparaison avec d'autres approches, le SSARIMA se distingue par sa capacité à modéliser de manière robuste et efficiente les comportements saisonniers et non saisonniers observés dans nos données.

En outre, les évaluations telles que le R2OOS ont renforcé la supériorité du modèle SSARIMA par rapport à d'autres alternatives, démontrant ainsi sa capacité à surpasser un modèle de référence naïf dans la prévision des tendances futures. Cette analyse renforce donc notre recommandation en faveur du modèle SSARIMA pour les prévisions et l'analyse de notre série temporelle.

Néanmoins, selon le test de Diebold-Mariano, le modèle ETS affiche une meilleure performance, ce qui nous a conduit à le sélectionner pour réaliser les prévisions pas à pas pour une année. Nous avons effectué pour cela deux approches pour cette analyse : la méthode Rolling qui assure une mise à jour continue du modèle de prévision, et la méthode récursive qui permet une actualisation progressive en se basant sur les prévisions précédentes. En conclusion, le modèle ETS utilisant la méthode

réursive semble légèrement plus performant, affichant une MSE plus faible et un R2OOS légèrement supérieur par rapport au modèle ETS utilisant la méthode Rolling.

## Discussion

Dans cette section nous aborderons les défis rencontrés au cours de notre étude. Tout d'abord, nous avons rencontré des difficultés liées à la prévision pas à pas, surtout sur le fait de ne pas avoir d'indication entre les bons résultats et ceux que nous avons trouvés. Nous nous sommes aussi demandé, si les résultats auraient été les mêmes si nous avions fait une prévision plus étendue, cela nous aurait donné une indication supplémentaire sur la robustesse de notre prévision. Une autre discussion porte sur la pertinence du modèle ETS par rapport à d'autres modèles de prévision, il faudrait prendre en compte différents critères et se baser sur cela. Enfin, il aurait été intéressant d'inclure des variables explicatives pour approfondir notre étude.

## Annexes

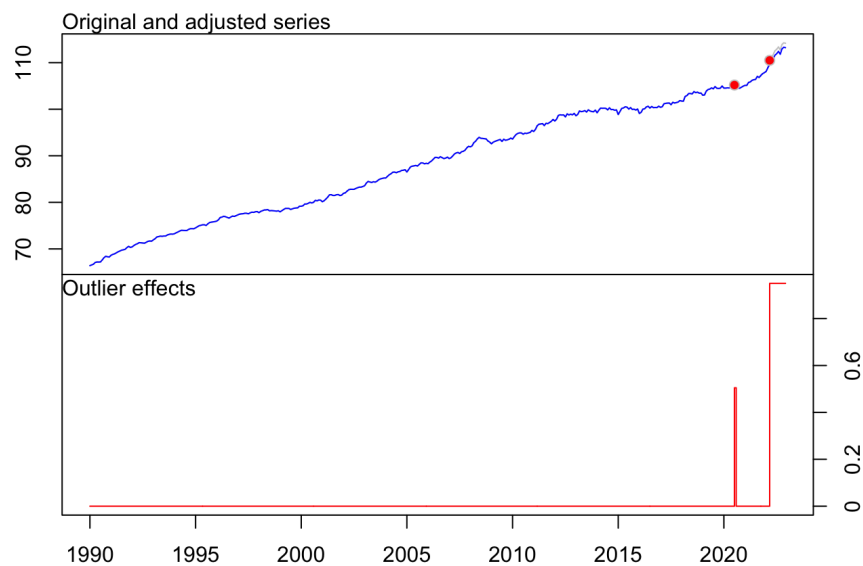
```
Regression with ARIMA(1,1,1)(2,0,0)[12] errors

Coefficients:
      ar1      ma1      sar1      sar2      A0367      LS387
      0.8922 -0.7712  0.4776  0.3198  0.5048  0.9500
s.e.  0.0594  0.0825  0.0519  0.0517  0.1147  0.201

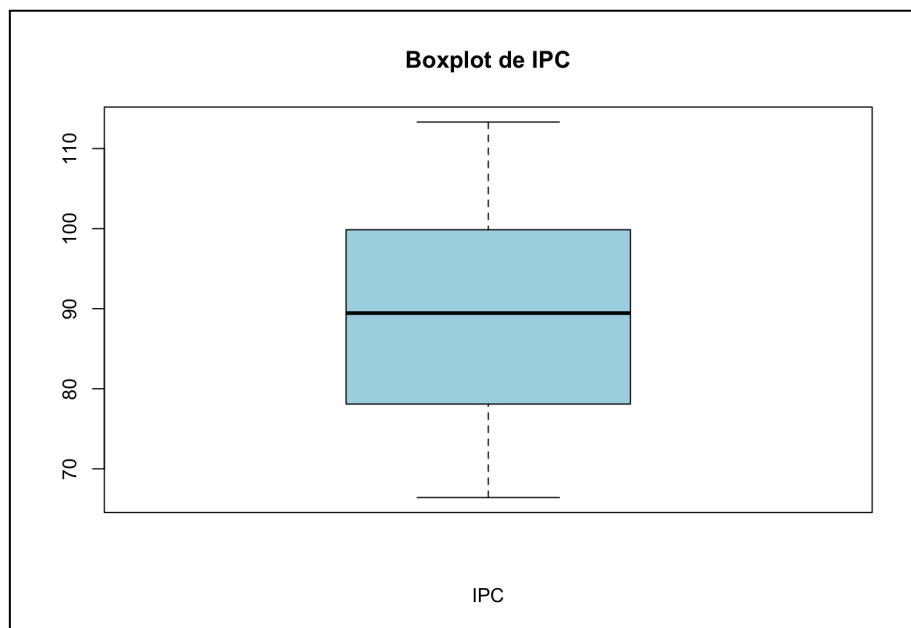
sigma^2 = 0.03984:  log likelihood = 73.52
AIC=-133.03  AICc=-132.75  BIC=-105.18

Outliers:
  type ind    time coefhat tstat
1  A0 367 2020:07  0.5048 4.403
2  LS 387 2022:03  0.9500 4.727
```

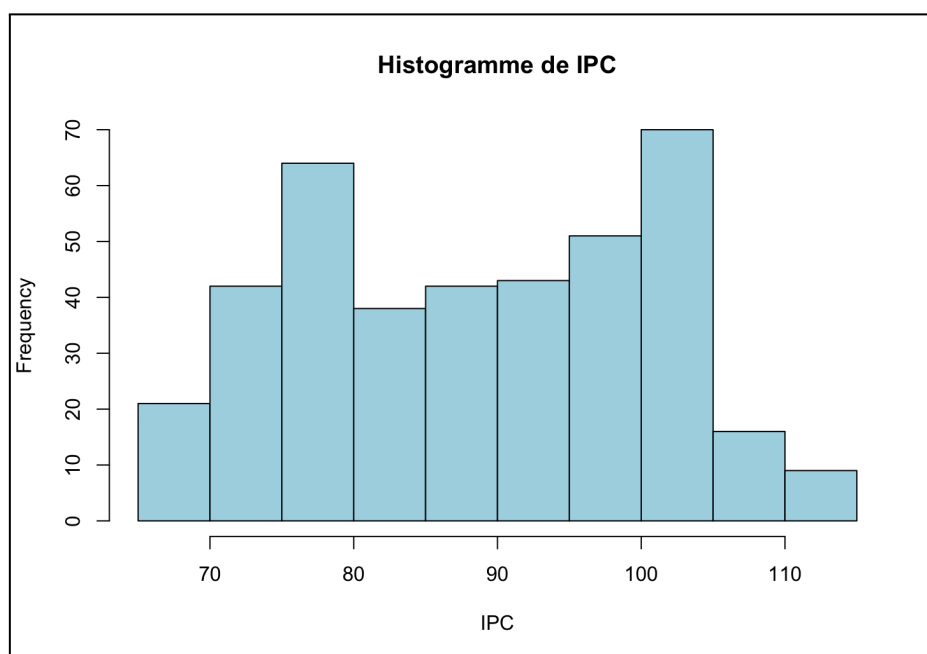
Annexe 1 : Régression ARIMA, détection outliers avec le package R tsoutliers.



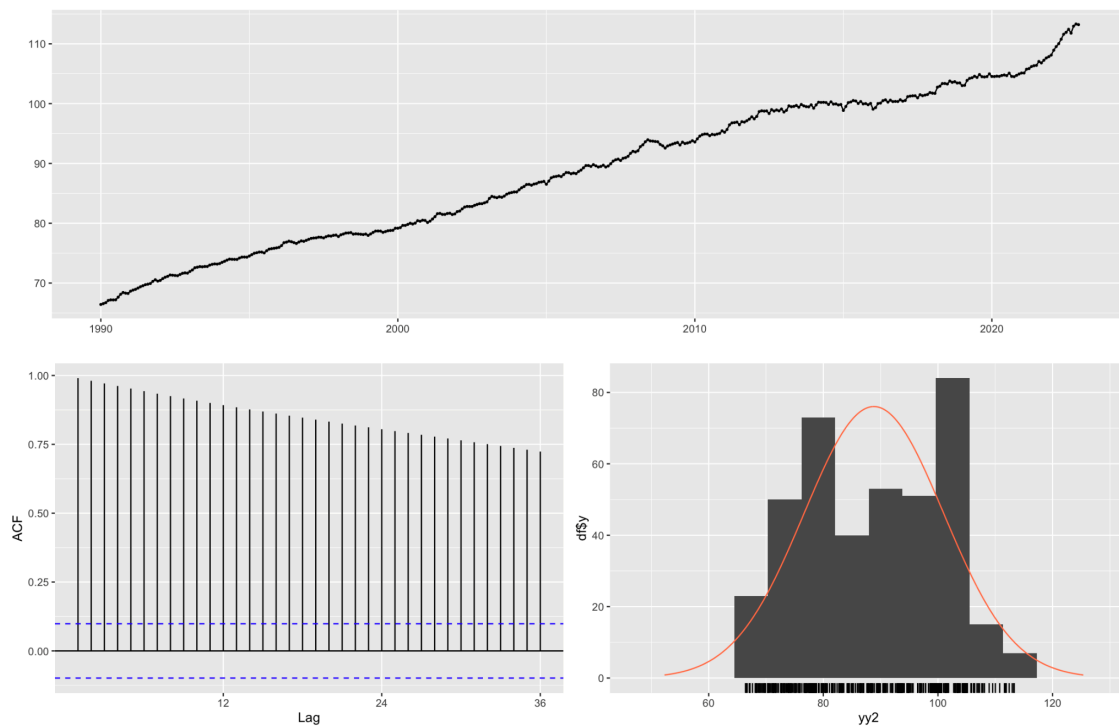
Annexe 2 : Détection des outliers pour la série temporelle brut par tsoutliers.



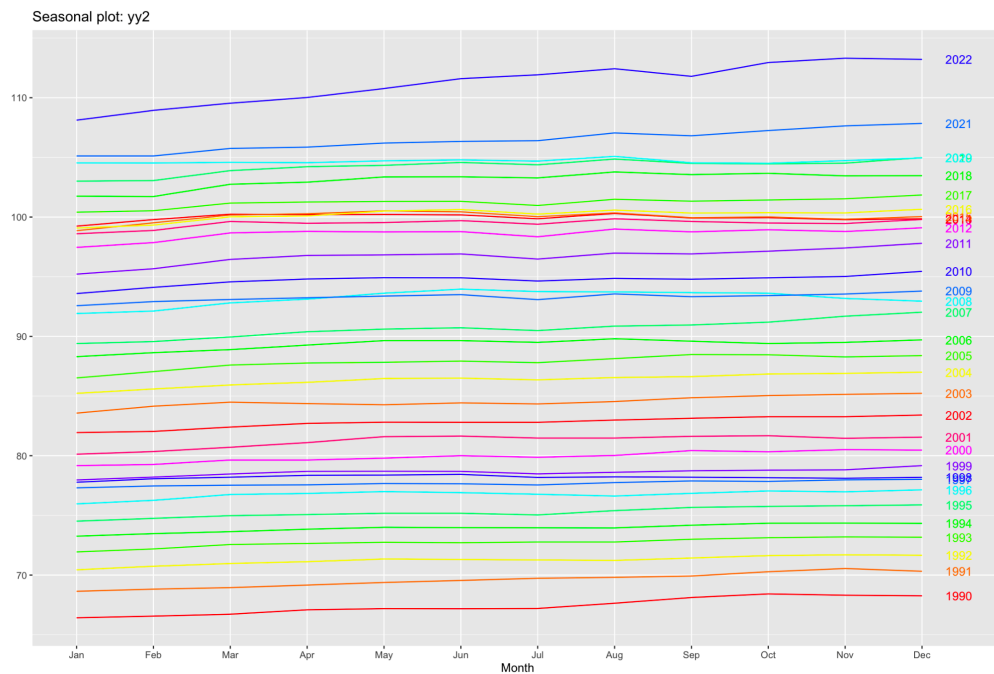
Annexe 3 : Boxplot de la série corrigée.



Annexe 4 : Histogramme de la série ajusté.



Annexe 5 : Graphiques d'ensemble de la série mensuelle.



Annexe 6 : Graphique de la comparaison de la saisonnalité par année.

Test used: SeasonalDummies

Test statistic: 20.7

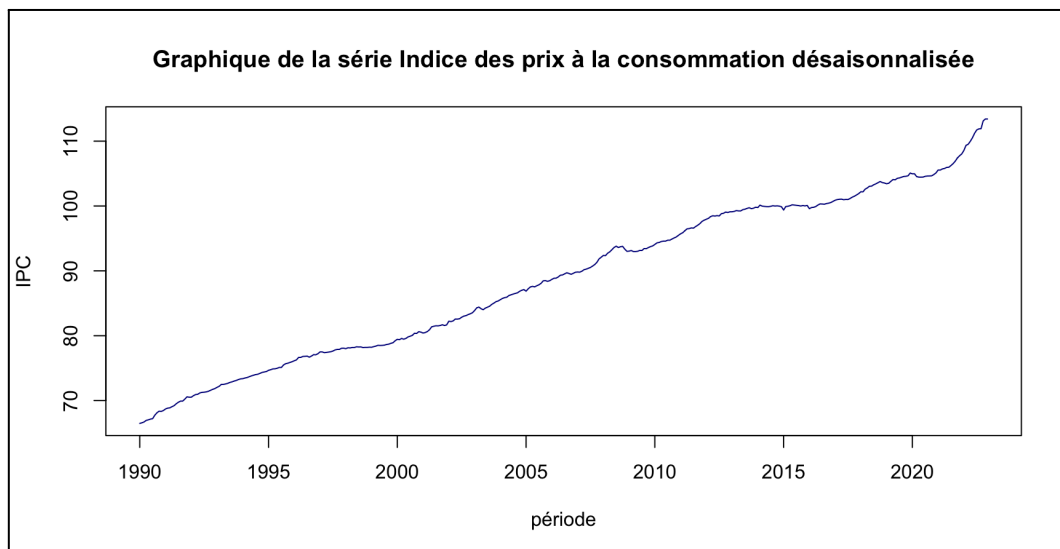
P-value: 0

Test used: W0

Test statistic: 1

P-value: 0 0 0

### Annexe 7 : Tests SeasonalDummies et W0 de détection de saisonnalité.



### Annexe 8 : Graphique de la série indice des prix à la consommation après désaisonnalisation.

```
y = regression model + arima (1, 1, 0, 0, 1, 1)

Model: RegARIMA - X13
Estimation span: from 1-1990 to 12-2022
Log-transformation: no
Regression model: no mean, no trading days effect, no leap year effect, Easter effect, outliers(2)

Coefficients:
ARIMA:
      Estimate Std. Error T-stat Pr(>|t|)
Phi(1)  -0.21771    0.05050  -4.311 2.07e-05 ***
BTheta(1) -0.64849    0.04432 -14.633 < 2e-16 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Regression model:
      Estimate Std. Error T-stat Pr(>|t|)
Easter [1]    0.04799    0.02261  2.122  0.0345 *
LS (10-2022)  1.02566    0.18082  5.672 2.78e-08 ***
A0 (1-2015)  -0.44368    0.10565 -4.200 3.33e-05 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1838 on 377 degrees of freedom
Log likelihood = 101.9, aic = -191.8, aicc = -191.6, bic(corrected for length) = -3.31
```

### Annexe 9 : Modèle RegARIMA - X13.

# Bibliographie

## Ressources internet

Banque de France (10-01-2022), Covid-19 et économie, les clés pour comprendre.

INSEE (17-02-2024), Indice des prix à la consommation - Base 2015.

INSEE (25-04-2024), Indice des prix à la consommation, IPC.

INSEE (31-03-2022), Indice des prix à la consommation - résultats provisoires (IPC) - mars 2022.

INSEE (31-07-2020), Indice des prix à la consommation - résultats provisoires (IPC) - juillet 2020.

INSEE (15-02-2024), Salaire minimum interprofessionnel de croissance (Smic).

INSEE (2024), Tableau de bord de l'économie française en 2021.

La finance pour tous (07-07-2023), Inflation en France avant et après l'euro.

Le Point (16-06-2017), France : la croissance revue à la hausse en 2014 et à la baisse en 2015.

Ministère de l'économie des finances et de la souveraineté industrielle et numérique (2013), Comment la crise a-t-elle commencé ? *La finance pour tous*.

Ollech.D, Test de saisonnalité combiné d'Ollech et Webel, *R documentation*.

SAS (11-08-2020), Mannequins saisonniers, *Utilisation de variables de prédiction*.

Statista (15-04-2024), Indice des prix à la consommation et inflation en France - Faits et chiffres.

Vie publique (03-06-2021), Vieillessement de la population : une adaptation nécessaire.



## Cours du Master ECAP

Darné.O, Technique de prévision et conjoncture, Semestre 2.

Girard.D, Séries temporelles univariées, Semestre 2.

## Packages

corrplot

dplyr

EnvStats

FactoMineR

factoextra

forecast

gvlma

ggplot2

gridExtra

lubridate

moments

nortest

openxlsx

outliers

RColorBrewer

readxl

RJDemetra

seastests

smooth

tidyr

tidyverse

tsoutliers

# Table des matières

<b>Sommaire.....</b>	<b>2</b>
<b>Introduction.....</b>	<b>3</b>
<b>I- Analyse exploratoire.....</b>	<b>5</b>
A) Présentation et caractérisation de la série étudiée.....	5
B) Détection des points atypiques.....	8
C) Statistiques descriptives sur la série corrigée.....	10
D) Détection de la saisonnalité et schéma de décomposition (additif ou multiplicatif).....	12
<b>II- Désaisonnalisation et décomposition.....</b>	<b>15</b>
<b>III- Prévion de la série saisonnière corrigée.....</b>	<b>17</b>
A) Les méthodes paramétriques, semi paramétriques et non paramétriques.....	17
B) Les méthodes de lissage exponentiel.....	20
C) Le modèle SARIMA.....	25
<b>IV- Représentation graphique de l'évolution des prévisions.....</b>	<b>26</b>
<b>V- Calcul de la qualité de prévision.....</b>	<b>28</b>
<b>VI- Test de précision : Le test de Diebold-Mariano.....</b>	<b>31</b>
<b>VII- Prévion sur une année avec un pas de 1 mois.....</b>	<b>32</b>
A) Méthode rolling.....	32
B) Méthode récursive.....	36
<b>Conclusion.....</b>	<b>40</b>
<b>Discussion.....</b>	<b>42</b>
<b>Annexes.....</b>	<b>43</b>
<b>Bibliographie.....</b>	<b>47</b>
Ressources internet.....	47
Cours du Master ECAP.....	48
Packages.....	48