

# Dossier écrit d'économétrie avancée

Analyse en série temporelle de l'indice des prix à la consommation français

Francis OLOBA LOBELA & George MOUZONG

2024-09-19

---

## Abstract:

An important aspect of macroeconomic analysis involves using past and present information to predict the future. Thus, the economist can observe the past evolution of GDP, inflation rates, etc., and use this information to predict their future developments. To achieve this, specific procedures unique to time series econometrics are necessary.

This study presents a comprehensive analysis of the Consumer Price Index (CPI) in France based on time series data. The analysis covers the evolution of the CPI over the past decades, highlighting key trends, seasonal patterns, and the impact of significant economic events. By leveraging various statistical and econometric concepts, including decomposition, autocorrelation, and stationarity, the study leads to the development of a predictive model that is both relatively simple and accurate.

---

## Résumé:

Un aspect important de l'analyse macroéconomique consiste à utiliser les informations passées et présentes pour prédire l'avenir. Ainsi, les économistes sont en mesure d'observer l'évolution passée de certaines grandeurs économiques telles que le PIB, le taux d'inflation, etc. et de s'en servir pour prédire leur évolution future. Pour ce faire des procédés spécifiques, propres à l'économétrie des séries temporelles, sont nécessaires.

Cette étude présente une analyse complète des indices des prix à la consommation (IPC) en France à partir de données de séries temporelles. L'analyse couvre l'évolution de l'IPC au cours des dernières décennies, en mettant en lumière les principales tendances, les schémas saisonniers et l'impact des événements économiques significatifs. En mobilisant divers concepts statistiques et économétriques, notamment la décomposition, l'autocorrélation et la stationnarité, l'étude conduit à l'élaboration d'un modèle prédictif à la fois relativement simple et précis.

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Données et sources</b>	<b>3</b>
2.1	Définition . . . . .	3
2.2	Source des données . . . . .	3
<b>3</b>	<b>Analyse des données</b>	<b>4</b>
3.1	Évolution de l'IPC . . . . .	4
3.2	Analyse de l'autocorrelation . . . . .	4
3.2.1	Analyse du correlogramme complet . . . . .	4
3.2.2	Analyse du correlogramme partiel . . . . .	5
3.3	Analyse de la stationnarité . . . . .	6
3.3.1	Test de Dickey-Fuller (DF) . . . . .	6
3.3.2	Test de Dickey-Fuller augmenté (ADF) . . . . .	7
3.3.3	Test de Phillips-Perron . . . . .	8
3.3.4	Test de Kwiatkowski-Phillips-Schmidt-Shin . . . . .	9
<b>4</b>	<b>Stationnarisation</b>	<b>9</b>
<b>5</b>	<b>Modélisation</b>	<b>10</b>
<b>6</b>	<b>Prévision</b>	<b>13</b>
<b>7</b>	<b>Prolongements</b>	<b>14</b>
7.1	L'inflation . . . . .	14
7.2	La saisonnalité . . . . .	16

# 1 Introduction

L'analyse des séries temporelles occupe une place centrale en macroéconomie, permettant aux économistes de décrypter les dynamiques passées et actuelles de variables économiques cruciales telles que le produit intérieur brut (PIB), le taux d'inflation, et les indices des prix à la consommation (IPC). La particularité de ces analyses est qu'elles mobilisent les valeurs passées d'un indicateur pour expliquer et prévoir ces variations futures. Ces analyses ne sont pas seulement descriptives ; elles sont essentielles pour la prévision et la prise de décision économique à différents niveaux (politique monétaire, planification budgétaire, stratégies d'investissement, etc.).

En France, l'IPC constitue l'un des principaux indicateurs utilisés pour évaluer l'inflation, c'est-à-dire la variation des prix des biens et services consommés par les ménages. Suivre l'évolution de l'IPC sur de longues périodes permet non seulement de comprendre les tendances inflationnistes mais aussi d'identifier les chocs économiques et les cycles conjoncturels qui ont marqué l'économie française.

Cependant, pour que ces observations soient véritablement utiles, il est crucial de recourir à des outils et techniques appropriés pour analyser les données de séries temporelles. L'économétrie des séries temporelles fournit ces outils, permettant non seulement de modéliser les données historiques mais aussi de développer des modèles prédictifs robustes pour anticiper les fluctuations futures des prix à la consommation.

Cette étude se propose donc d'explorer l'évolution des indices des prix à la consommation en France au cours des dernières décennies. En mobilisant des méthodes économétriques telles que la décomposition, l'analyse de l'autocorrélation et de la stationnarité, nous chercherons à dégager les principaux déterminants des variations de l'IPC et à élaborer un modèle de prévision capable de fournir des estimations fiables pour les mois à venir. Cette analyse fournit un cas d'école, détaillant des étapes indispensables à toute analyse en série temporelle suivant principalement la méthode de Box et Jenkins.

Nous débuterons par une analyse approfondie des données, puis nous effectuerons une évaluation graphique de la stationnarité afin d'assurer la fiabilité des prévisions. Les tests de stationnarité seront utilisés pour évaluer la stabilité de la série temporelle ou pour déterminer si elle doit être stationnarisée, ce que nous ferons si nécessaire. Finalement, ces études permettront de créer un modèle prédictif, offrant une meilleure compréhension des tendances macroéconomiques et des prévisions plus précises.

## 2 Données et sources

Dans cette section, nous analyserons notre variable d'intérêt : l'indice des prix à la consommation (hors tabac).

### 2.1 Définition

L'IPC est l'outil principal de mesure de l'inflation. Il permet d'estimer la variation moyenne des prix des biens et services consommés par les ménages entre deux périodes données. L'IPC est calculé à partir d'un panier fixe de biens et services, qui est mis à jour chaque année pour refléter les habitudes de consommation des ménages. Chaque produit inclus dans ce panier est pondéré dans l'indice global en fonction de sa part dans les dépenses de consommation des ménages (Institut national de la statistique et des études économiques, INSEE)

### 2.2 Source des données

Les valeurs de l'IPC ont été obtenues sur le site de l'INSEE, l'organisme public français chargé de la production, de l'analyse et de la diffusion des statistiques officielles en France. Elles regroupent 416 observations allant de janvier 1990 à août 2024. De fait, la fiabilité des données semble garantie et la méthode de calcul des indices est la même sur toute la période. Afin de s'assurer de la fiabilité de la manœuvre, intentionnellement, les données de l'année 2024 ont été écartées. Ceci nous permettra de comparer nos prévisions pour l'année 2024 à nos observations et apprécier leur précision. Enfin, l'indice est calculé en base 100 en 2015, ce qui veut dire que, dans notre échelle de valeur, l'IPC moyen de l'année 2015 est de 100.

## 3 Analyse des données

Dans cette section, nous approfondirons notre connaissance des données. Nous allons ainsi explorer, décrire, analyser et représenter les données sur lesquelles nous travaillons. Plus précisément, nous traiterons de l'évolution, de l'autocorrélation et de l'éventuelle stationnarité de la série.

### 3.1 Évolution de l'IPC

Nous nous intéressons à l'évolution de l'IPC au fil du temps. D'un point de vue économique, l'augmentation de l'IPC d'une période  $t$  à une période  $t + 1$  traduit une inflation positive, tandis que sa diminution indique un taux d'inflation négatif, soit une déflation sur la période.

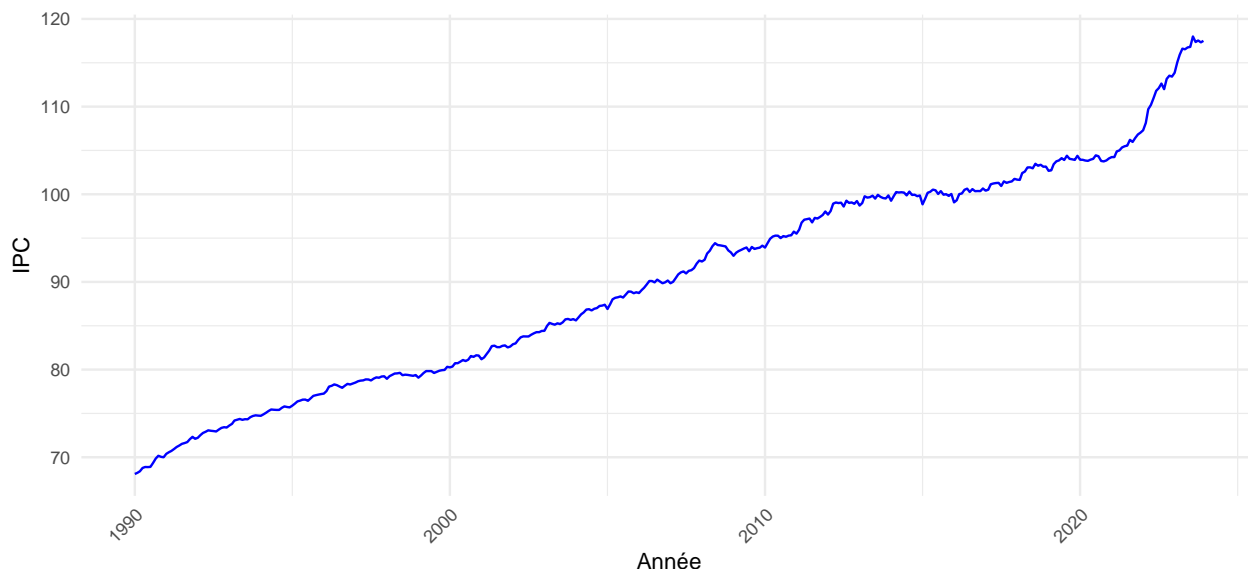


Figure 1: Représentation graphique de l'indice des prix à la consommation

Comme l'on pouvait s'y attendre, on constate que les valeurs de la série ont fortement augmenté entre 1990 et 2024, ce qui traduit une hausse du niveau général des prix cohérente avec ce que la théorie économique laisse présager. L'inflation est particulièrement forte à partir de 2020, où les effets combinés de la crise covid et surtout de la guerre en Ukraine nourrissent une forte haute hausse du niveau général des prix. La forte hausse des prix de l'énergie est vue par de nombreux économistes comme l'un des principaux éléments à l'origine de cette hausse. Enfin, la série apparaît, a priori, comme non stationnaire. L'analyse de l'autocorrélation (corrélogramme et corrélogramme partiel) et les tests de stationnarité nous permettront de savoir ce qu'il en est réellement.

### 3.2 Analyse de l'autocorrélation

L'autocorrélation est un concept statistique qui mesure la corrélation d'une série temporelle avec elle-même à différents décalages dans le temps. En d'autres termes, l'autocorrélation examine comment les valeurs d'une série temporelle sont liées à leurs propres valeurs passées.

Nous l'observerons au travers du corrélogramme et du corrélogramme partiel.

#### 3.2.1 Analyse du corrélogramme complet

Un corrélogramme complet est un graphique utilisé en statistique pour visualiser l'autocorrélation d'une série temporelle à différents décalages (ou lags).

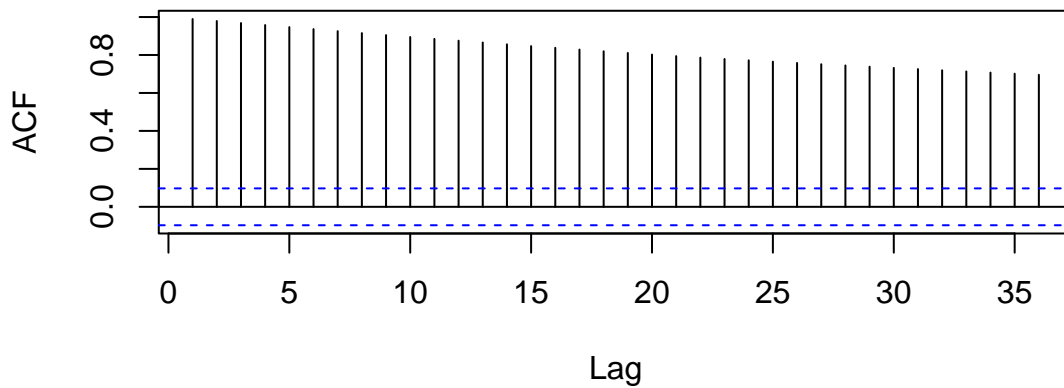


Figure 2 : Corrélogramme de la série temporelle

On constate que l'autocorrélation décroît lentement au fur et à mesure que le décalage augmente. Cela signifie que la série temporelle présente une dépendance significative à long terme. Plus précisément, la série est fortement corrélée avec ses valeurs passées, même sur plusieurs mois.

Le corrélogramme indique une forte autocorrélation persistante, avec une décroissance lente au fil des décalages. Cela suggère que la série temporelle est probablement non stationnaire et qu'elle a une dépendance à long terme.

### 3.2.2 Analyse du correlogramme partiel

Un corrélogramme partiel est un graphique utilisé pour visualiser les autocorrélations partielles d'une série temporelle à différents lags. Contrairement au corrélogramme complet, qui montre la corrélation entre une valeur et ses valeurs passées, le corrélogramme partiel montre la corrélation entre une valeur et ses valeurs passées en éliminant les effets des lags intermédiaires.

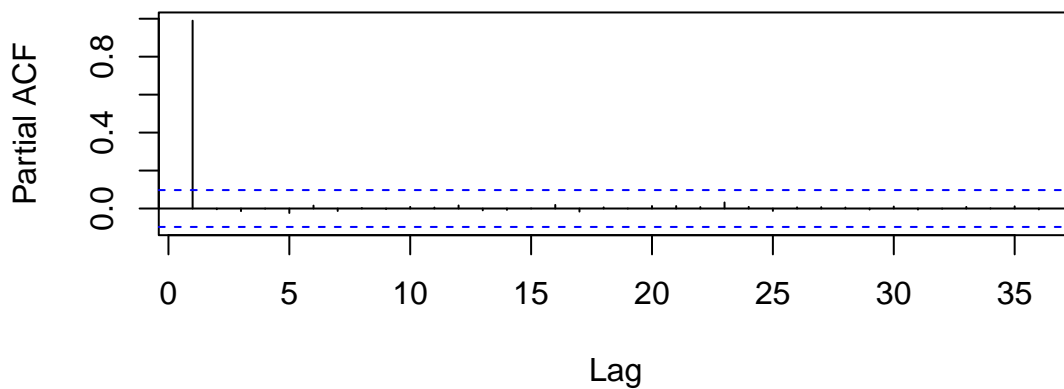


Figure 3 : Corrélogramme partiel de la série temporelle

La première barre (lag 1) a une autocorrélation partielle significative (barre qui dépasse nettement les lignes bleues). Cela suggère que la série a une dépendance à court terme, c'est-à-dire que la valeur de l'IPC en  $t$  est fortement influencée par la valeur de la période précédente  $t - 1$ . Toutefois, nous n'observons aucune autre barre significative après le lag 1. Cela signifie que les décalages suivants (2, 3, etc.) n'ont pas de corrélation partielle significative avec la série, une fois que l'effet du premier décalage (lag 1) est considéré. Nous pouvons à présent nous intéresser à la stabilité de la série.

### 3.3 Analyse de la stationnarité

Une série temporelle est dite stationnaire si ses moments d'ordre 1 et 2, la moyenne et l'autocovariance, restent constants au fil du temps. L'analyse de l'évolution de l'IPC et celle de l'autocorrélation semblent indiquer que nous sommes en présence d'une série non stationnaire. Toutefois, il est impératif de recourir à de tests statistiques afin de confirmer cette intuition. À cet effet, nous utiliserons successivement les tests de Dickey-Fuller (DF), de Dickey-Fuller augmenté (ADF), de Phillips-Perron (PP) et le test de Kwiatkowski-Phillips-Schmidt-Shin (KPSS). L'utilisation de cette variété de tests n'est pas une remise en question de leur précision individuelle. Elle vise simplement à offrir au lecteur une vue globale des différentes manières de vérifier la stationnarité.

#### 3.3.1 Test de Dickey-Fuller (DF)

Le test de Dickey-Fuller est, avec le test de Dickey-Fuller augmenté et le test de Phillips-Perron, un test de racine unitaire. En ce sens, l'hypothèse nulle stipule que la série est générée par un processus comportant une racine unitaire, ce qui signifie qu'elle est non stationnaire. Il se décline en test de Dickey-Fuller simple (1), test de Dickey-Fuller avec dérive (2) et test de Dickey-Fuller avec tendance (3). Ils estiment respectivement les modèles suivants (avec  $\rho = 1 - \gamma$ ) :

$$\Delta Y_t = \gamma Y_{t-1} + \epsilon_t \quad \text{ou} \quad Y_t = \rho Y_{t-1} + \epsilon_t \quad (1)$$

$$\Delta Y_t = \alpha + \gamma Y_{t-1} + \epsilon_t \quad \text{ou} \quad Y_t = \alpha + \rho Y_{t-1} + \epsilon_t \quad (2)$$

$$\Delta Y_t = \alpha + \gamma Y_{t-1} + \beta t + \epsilon_t \quad \text{ou} \quad Y_t = \rho Y_{t-1} + \beta t + \epsilon_t \quad (3)$$

Dans les trois cas cependant, le jeu d'hypothèses reste le même

$$H_0 : \gamma = 0 \quad \text{ou} \quad \rho = 1$$

$$H_1 : |\gamma| < 0 \quad \text{ou} \quad |\rho| < 1$$

Le test de Dickey-Fuller vérifie l'existence d'une racine unitaire. L'hypothèse fondamentale est donc qu'il existe une racine unitaire. Celle-ci étant incompatible avec une série stationnaire, cela impliquerait la non-stationnarité de la série. En effet, lorsqu'une série temporelle possède une racine unitaire, cela signifie que ses valeurs sont hautement persistantes, ce qui la rend susceptible de présenter une tendance au fil du temps, rendant de facto la stationnarité impossible. En théorie, nous devrions dans un premier temps tester le modèle avec tendance (3). En fonction des résultats, nous vérifierions ensuite la significativité des termes du modèle pour vérifier que celui-ci est bien adapté à notre série. Si tel est le cas, nous en gardons le résultat ; sinon, nous procédons à l'identique au test du modèle avec dérive (2) puis à celui du modèle simple (1) le cas échéant<sup>1</sup>.

Dans notre cas, les résultats ne laissent planer aucun doute sur la nature de la série : elle n'est pas stationnaire. Le modèle (3) est rejeté pour car mal spécifié et c'est à partir du modèle (2) que nous arrivons à conclure.

Toutefois, le test de Dickey-Fuller est limité en ce sens qu'il ne tient compte que d'un seul lag. Si cette limite peut être sans conséquence pour des processus très simples, sans autocorrélation des résidus, elle devient très contraignante pour les processus plus complexes. Dès lors, dans l'hypothèse où la série temporelle serait générée par un processus plus complexe ou avec de l'autocorrélation dans les résidus, le test DF pourrait donner des résultats biaisés et peu fiables. Le test DF est donc limité dans sa capacité à modéliser la série de manière adéquate, ce qui explique qu'il n'est pas utilisé tel quel dans la pratique. On lui préférera dans ce cas une version plus performante, le test de Dickey-Fuller augmenté, dit ADF. Celui-ci permet de prendre en compte des processus autorégressifs d'ordres plus importants tout en tenant compte de la possibilité d'autocorrélation des résidus. Dans notre cas, nous constatons la présence d'autocorrélation dans les résidus ce qui nous contraint à abandonner nos conclusions pour recourir au test ADF.

<sup>1</sup>Ce procédé est également valable pour le test ADF

```

Test de Dickey-Fuller simple :
Statistique de test: 8.30857630021412
P-valeur: 1.0
Valeurs critiques : {'1%': -2.5712552617884805, '5%': -1.941679801779741, '10%': -1.6161834279747518}

Test de Dickey-Fuller (tendance) :
Statistique de test: 0.16257512075546626
P-valeur: 0.9955522322484323
Valeurs critiques : {'1%': -3.9811870946951964, '5%': -3.4213324456774834, '10%': -3.1334268521857362}

Test de Dickey-Fuller (dérive) :
Statistique de test: 1.489334608521645
P-valeur: 0.9974844806230153
Valeurs critiques : {'1%': -3.4465195891135845, '5%': -2.8686676281678634, '10%': -2.5705668101226085}

```

Tableau 1 : Résultats du test de Dickey-Fuller

### 3.3.2 Test de Dickey-Fuller augmenté (ADF)

Ce test, comme le précédent, vérifie l'existence d'une racine unitaire. L'hypothèse fondamentale reste donc le fait que la série possède une racine unitaire, c'est-à-dire qu'elle n'est pas stationnaire. Toutefois, à la différence du test DF, le test ADF tient compte d'un possible caractère autorégressif au-delà de l'ordre 1 et de l'éventuelle autocorrélation des résidus. Davidson [2004] montrent que le test ADF est celui qui fournit les meilleurs résultats dans le cas d'échantillons finis. Le modèle de base du test ADF (4) peut également être modifié pour inclure une dérive (5) ou une tendance (6). Les modèles estimés sont respectivement les suivants (avec  $\rho = 1 - \gamma$ ) :

$$\Delta Y_t = \gamma Y_{t-1} + \sum_{i=1}^p \beta_i \Delta Y_{t-i} + \epsilon_t \quad \text{ou} \quad Y_t = \rho Y_{t-1} + \sum_{i=1}^p \beta_i \Delta Y_{t-i} + \epsilon_t \quad (4)$$

$$\Delta Y_t = \alpha + \gamma Y_{t-1} + \sum_{i=1}^p \beta_i \Delta Y_{t-i} + \epsilon_t \quad \text{ou} \quad Y_t = \rho Y_{t-1} + \alpha + \sum_{i=1}^p \beta_i \Delta Y_{t-i} + \epsilon_t \quad (5)$$

$$\Delta Y_t = \alpha + \gamma Y_{t-1} + \phi t + \sum_{i=1}^p \beta_i \Delta Y_{t-i} + \epsilon_t \quad \text{ou} \quad Y_t = \rho Y_{t-1} + \alpha + \phi t + \sum_{i=1}^p \beta_i \Delta Y_{t-i} + \epsilon_t \quad (6)$$

Pour le reste, le test reste très proche du test DF, et leurs jeux d'hypothèses sont les mêmes, c'est-à-dire :

$$H_0 : \gamma = 0 \quad \text{ou} \quad \rho = 1$$

$$H_1 : |\gamma| < 0 \quad \text{ou} \quad |\rho| < 1$$

Le nombre de lags à utiliser constitue toutefois un dilemme. En effet, il est nécessaire de trouver un équilibre entre complexité du modèle et qualité de l'ajustement, de manière à éviter toute autocorrélation dans les résidus. Dans notre travail, nous choisissons le nombre de lags optimal en fonction du critère d'information bayésien (BIC). Cette méthode permet d'optimiser le nombre de lags tout en pénalisant les modèles avec trop de paramètres pour éviter le surajustement. Nous trouvons ainsi un compromis entre qualité d'ajustement et complexité. Nous devons toutefois nous assurer que les résidus du modèle ne sont pas autocorrélés.

Nous procédons tel qu'expliqué précédemment : dans un premier temps tester le modèle avec tendance (6). En fonction des résultats, vérifier la significativité des termes du modèle pour vérifier que celui-ci est adapté à notre série. Si tel est le cas, il faut garder le résultat ; sinon, procéder à l'identique au test du modèle avec dérive (5) puis à celui du modèle simple (4), le cas échéant. De plus, nous devons tester l'autocorrélation des résidus, car des résidus autocorrélés remettraient en cause toutes nos conclusions. En présence d'autocorrélation des résidus, deux possibilités s'offrent à nous : (a) considérer que cela provient d'un problème de spécification du modèle et passer à un modèle plus simple ou (b) juger que cela provient d'une faible prise en compte des valeurs précédentes et rajouter des lags. Dans notre cas, l'utilisation de 12 lags nous permet d'obtenir des résidus non autocorrélés. Dans ce travail, tous les tests ADF sujets à interprétation ont été testés robustes à l'autocorrélation.

L'utilisation du BIC pour sélectionner le nombre optimal de lags est automatisée sur les logiciels Python et RStudio. Sur Python, le procédé conduit à retenir 12 lags contre un seul lag sur RStudio. Nous rencontrons alors une difficulté : le choix du nombre de lags.

```
ADF avec tendance :
Statistique ADF : -4.23464251615976
Valeur critique : -3.4216635867385743
Nombre de lags utilisés : 12
P-value : 0.003966489712371071
```

```
ADF sans tendance ni dérive :
Statistique ADF : 1.8440256623416462
Valeur critique : -1.9417010604353815
Nombre de lags utilisés : 12
P-value : 0.9851571433480307
```

Tableau 3 : Résultats du test de Phillips-Perron

En ne retenant qu'un lag, nous observons bientôt que les résidus restent fortement autocorrélés. En l'occurrence, le test de Ljung-Box nous donne une p-value de l'ordre de  $10^{-16}$ . Les résultats restent les mêmes quels que soient les modèles sous-jacents retenus (tendance, dérive ou simple). En procédant par la méthode (b), nous réalisons le test ADF, avec tendance et dérive, sur 12 lags. Ce test conclut à la stationnarité de la série et le coefficient de la tendance,  $\phi$ , sont significatifs. Le modèle apparaît donc comme correctement spécifié. Le test de Ljung-Box conclut à l'absence d'autocorrélation dans les résidus (p-value de 0,2226). Selon le test ADF, la série est stationnaire en tendance. En d'autres termes, il est possible d'exprimer le processus générateur de la série sous la forme :  $Y_t = f(t) + z_t$ , où  $z_t$  est une série stationnaire. Dès lors, il nous faudra choisir une méthode adaptée pour stationnariser la série.

Dans la suite du travail, nous procéderons en considérant que la série est stationnaire en tendance. Nous estimerons différents modèles et comparerons leurs pouvoirs prédictifs afin de décider lequel des modèles fournit les meilleurs résultats. Avant cela, poursuivons notre analyse de la stationnarité avec le test de Phillips-Perron.

### 3.3.3 Test de Phillips-Perron

À l'instar du test ADF, le test de Phillips-Perron est un test de racine unitaire utilisé pour vérifier la stationnarité d'une série temporelle. Cependant, il diffère du test ADF dans la façon dont il traite l'autocorrélation et l'hétéroscédasticité dans les erreurs. En effet, le test PP utilise une approche non paramétrique pour corriger les erreurs pour l'autocorrélation et l'hétéroscédasticité, contrairement au test ADF qui utilise un modèle autorégressif avec des lags.

Le test de Phillips-Perron se déroule en quatre étapes principales. Dans la première étape, les trois modèles de base du test DF sont estimés à l'aide des Moindres Carrés Ordinaires (MCO). Cela permet de tester la stationnarité de la série temporelle dans différents contextes. Ensuite, la deuxième étape consiste à estimer la variance à court terme, notée  $\sigma^2$ , à partir des résidus des modèles précédemment estimés. Cette variance capture la dispersion immédiate des résidus autour de leur moyenne. Dans la troisième étape, une variance à long terme,  $s_t^2$ , est calculée. Cette estimation tient compte de l'autocorrélation et de l'hétéroscédasticité dans les résidus, corrigeant ainsi les biais potentiels liés à ces phénomènes. Enfin, lors de la quatrième étape, la statistique de Phillips-Perron est calculée en ajustant la statistique du test de Dickey-Fuller, grâce aux estimations des variances à court et long terme. Cette statistique est ensuite comparée aux valeurs critiques issues de la table de Mackinnon. Pour ce test, il n'est pas nécessaire de s'assurer de la sphéricité des résidus car les possibles effets d'une non-sphéricité sur les résultats sont neutralisés par la procédure de test. Ainsi, le test de Phillips-Perron offre une alternative robuste au test de Dickey-Fuller pour tester la stationnarité en présence d'autocorrélation et d'hétéroscédasticité dans les résidus. Pour ce test, nous avons retenu 17 lags par application du paramètre de troncature de Schwert [Schwert, 2002]. Celui-ci recommande de sélectionner le nombre de lags par :

$$l_{12} = 12\left(\frac{N}{100}\right)^{1/4}$$

où N est la taille de l'échantillon.



```

Statistique Phillips-Perron : 0.9426500844135858
Valeur critique (5%) : -2.8686676281678634
P-value : 0.9936224265845313

```

Tableau 4 : Résultats du test de Phillips-Perron

Le test de Phillips-Perron révèle la non-stationnarité de la série. La p-value de 0,99 confirme ce résultat. En particulier, les résultats suggèrent que la série est intégrée d'ordre 1. Toutefois, nous retiendrons principalement le résultat du test ADF, plus précis en échantillon fini [Davidson, 2004].

Dans la partie suivante, nous mobilisons le test de Kwiatkowski-Phillips-Schmidt-Shin (KPSS) qui a la particularité d'être un test de stationnarité, à la différence des tests de racine unitaire mobilisés jusqu'alors.

### 3.3.4 Test de Kwiatkowski-Phillips-Schmidt-Shin

Le test KPSS (Kwiatkowski-Phillips-Schmidt-Shin) est utilisé pour vérifier la stationnarité d'une série temporelle. Contrairement aux tests de racine unitaire où l'hypothèse nulle est que la série est non stationnaire, l'hypothèse nulle dans le test KPSS est que la série est stationnaire. Le test peut être ajusté selon que l'on souhaite faire l'hypothèse nulle d'une stationnarité en tendance ou en niveau :

$H_0$  : La série est stationnaire (en niveau ou en tendance)

$H_1$  : La série est non-stationnaire (processus DS)

```

Statistique KPSS en tendance : 0.112053718089509
Valeur critique (10%, 5%, 2.5%, 1%) : {'10%': 0.119, '5%': 0.146, '2.5%': 0.176, '1%': 0.216}
P-value : 0.1

```

```

Statistique KPSS en niveau : 3.211664672110235
Valeur critique (10%, 5%, 2.5%, 1%) : {'10%': 0.347, '5%': 0.463, '2.5%': 0.574, '1%': 0.739}
P-value : 0.01

```

Tableau 5 : Résultats du test KPSS

Le test KPSS conclut que la série est stationnaire en tendance, l'hypothèse de la stationnarité en tendance n'étant pas rejetée au seuil de risque 5%. Pour ce test, nous appliquons le paramètre de troncature de Schwert dans le choix du nombre de lags.

Dans la suite de notre travail, nous procédons à la stationnairisation de la série en extrayant la tendance détectée dans les tests.

## 4 Stationnarisation

Dans cette section, nous effectuons les changements nécessaires afin de rendre notre série stationnaire en niveau. Pour ce faire, nous retenons que notre série est stationnaire en tendance (TS).

Un processus TS est un processus qui est stationnaire autour d'une tendance. Dans notre cas, nous retenons l'hypothèse que cette tendance est linéaire. Une telle tendance peut rendre une série non-stationnaire, car la moyenne évolue au fil du temps. Il est donc nécessaire de la supprimer pour stationnariser la série d'entamer la modélisation.

Dans notre cas, la nature TS du processus générateur de la série implique que celle-ci peut être représentée comme :

$$IPC_t = \alpha + \beta t + \epsilon_t$$

où  $\epsilon_t$  est un bruit blanc ou résidu stationnaire. Un modèle peut alors être estimé par la méthode des moindres carrés ordinaires pour estimer cette relation :

$$\hat{IPC}_t = \hat{\alpha} + \hat{\beta}t \quad (7)$$

Pour un tel processus, l'estimateur de  $\epsilon_t$  donné par :

$$\hat{\epsilon}_t = IPC_t - \hat{IPC}_t$$

est stationnaire, avec  $\hat{IPC}_t$  la valeur prédite de l'IPC en t par le modèle MCO.

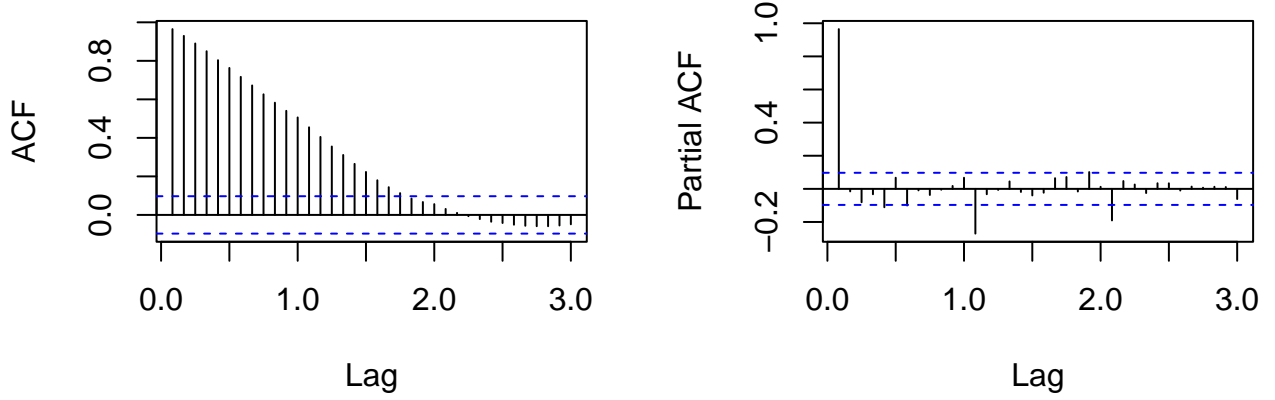


Figure 4 : ACF et PACF de la série temporelle corrigée de la tendance

Nous constatons que le corrélogramme et le corrélogramme partiel sont compatibles avec une série stationnaire, et semblent suggérer un processus AR. De plus, les tests ADF et KPSS menés sur cette nouvelle série concluent que la série est stationnaire en niveau, confirmant ainsi la réussite du procédé. La figure 6 permet de constater la nouvelle évolution de notre variable.

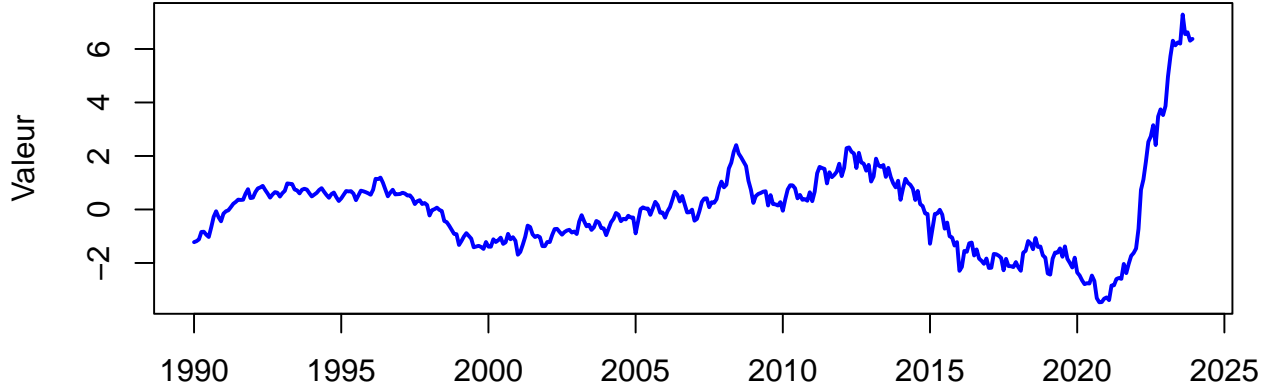


Figure 5 : Évolution de la série temporelle sans tendance

## 5 Modélisation

Dans cette section nous procédons, pour notre série corrigée de la tendance, à l'estimation des modèles suivants de modèles ARMA, dont l'expression générale est la suivante pour un ARMA (p, q) :

$$IPC_t = \phi_1 IPC_{t-1} + \phi_2 IPC_{t-2} + \dots + \phi_p IPC_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q} \quad (8)$$

ou

$$IPC_t = \sum_{i=1}^p \phi_i IPC_{t-i} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j} \quad (9)$$

où  $\varepsilon_t$  est un bruit blanc.

Les modèles estimés sont :

- ARMA (12,3) : effets conjoncturels horizon annuel et faible persistance des chocs. Du fait de l'hétéroscédasticité détectée dans ce modèle, nous rajoutons un lag ce qui nous conduit à estimer un ARMA (13,3).
- AR (25) ou ARMA (25,0) : sur base de l'observation de l'ACF et le PACF de la série
- ARMA (24,12) : effets conjoncturels sur deux ans et persistance importante des chocs

Les estimations de ces modèles sont présentées dans la table 1.

Table 1: Résultats des Modèles ARMA

	Modèle 1 (ARMA 13, 3)	Modèle 2 (ARMA 25, 0)	Modèle 3 (ARMA 24, 12)
Intercept	0.0351(0.3105)	0.0807(0.3475)	0.0036(0.3599)
AR1	0.8953*** (0.0675)	1.1036*** (0.0475)	0.5946** (0.2357)
AR2	0.1244(0.0789)	-0.1139(0.0708)	-0.0132(0.0872)
AR3	0.0457(0.0748)	-0.0031(0.0710)	0.0467(0.0757)
AR4	-0.0521(0.0709)	0.0977(0.0710)	0.0833(0.0881)
AR5	-0.0243(0.0498)	-0.1011(0.0713)	0.0467(0.0951)
AR6	0.0658(0.0495)	0.1265(0.0718)	0.0111(0.0780)
AR7	-0.0619(0.0503)	-0.0890(0.0721)	-0.0311(0.0973)
AR8	-0.0096(0.0511)	0.0097(0.0727)	0.0708(0.0812)
AR9	-0.0166(0.0511)	-0.0884(0.0738)	-0.2381*** (0.0827)
AR10	0.0706(0.0519)	0.1502** (0.0743)	0.2705*** (0.1017)
AR11	-0.0858* (0.0504)	-0.1011(0.0750)	-0.0646(0.0987)
AR12	0.6526*** (0.0527)	0.4490*** (0.0747)	1.0129*** (0.0654)
AR13	-0.6442*** (0.0415)	-0.5728*** (0.0710)	-0.6104*** (0.2327)
AR14	---	0.1233(0.0735)	0.0156(0.0855)
AR15	---	0.0080(0.0741)	-0.0351(0.0777)
AR16	---	-0.1309(0.0755)	-0.1254(0.0851)
AR17	---	0.1175(0.0756)	-0.0336(0.0939)
AR18	---	-0.1160(0.0758)	-0.0356(0.0799)
AR19	---	0.0508(0.0762)	0.0139(0.0952)
AR20	---	0.0215(0.0752)	-0.0702(0.0774)
AR21	---	0.0720(0.0764)	0.1850** (0.0830)
AR22	---	-0.1238(0.0780)	-0.2518*** (0.0951)
AR23	---	0.0358(0.0793)	0.0589(0.0926)
AR24	---	0.3525*** (0.0789)	-0.0505(0.0660)
AR25	---	-0.3089*** (0.0537)	---
MA1	1.5888*** (0.0857)	---	0.4797** (0.2332)
MA2	1.3994*** (0.1227)	---	0.4423* (0.2292)
MA3	0.5491*** (0.0796)	---	0.4150* (0.2204)
MA4	---	---	0.4621** (0.2165)
MA5	---	---	0.3383(0.2123)
MA6	---	---	0.4541** (0.1813)
MA7	---	---	0.4564** (0.2004)
MA8	---	---	0.4261** (0.2121)
MA9	---	---	0.6154*** (0.2019)
MA10	---	---	0.4289* (0.2348)
MA11	---	---	0.3719(0.2162)
MA12	---	---	-0.3882* (0.2027)
AIC	-15.2	-39.53	-63.57
BIC	57	68.78	88.86
Ljung-Box p-value	0.4235	0.9945	1.0000
Jarque-Bera p-value	1.113e - 10	< 2.2e - 16	1.513e - 12
ARCH p-value	< 2.2e - 16	1.435e - 09	3.308e - 10
Breusch-Pagan p-value	0.07596	0.1822	0.2149

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

De manière à obtenir des modèles plus simples, avec moins de termes nous retirons les termes non significatifs de chaque modèle. Ceci est notamment possible car le test de Breush-Pagan et celui de Ljung-Box ont permis de conclure à une absence d'autocorrélation et d'hétéroscédasticité. Nos modèles sont donc estimés à nouveau, une première fois sans les coefficients non significatifs, ce qui nous donne les résultats en Table 2. Sur base de ces résultats, nous pouvons comparer les modèles afin de choisir notre modèle benchmark. Pour ce faire, nous commençons par différents tests de validation.

Premièrement, nous utilisons le test de Ljung-Box pour vérifier la présence d'autocorrélation. L'hypothèse nulle de ce test est l'absence d'autocorrélation. Nous constatons que le modèle 1 présente de l'autocorrélation contrairement aux modèles 2 et 3, où les résidus ne sont pas autocorrélés. En utilisant le test de Jarque-Bera<sup>2</sup>, nous vérifions la normalité des résidus. Il apparaît qu'aucun des modèles ne présente des résidus normaux, ce qui n'est pas problématique dans la mesure où l'échantillon est de grande taille. Le test d'effet ARCH nous permet de vérifier la présence d'hétéroscédasticité conditionnelle<sup>3</sup>. Selon les résultats du test, modèles présentent tous de l'hétéroscédasticité conditionnelle, ce qui indique que des modèles de type ARCH/GARCH seraient indiqués pour capturer cette volatilité. Ces modèles sont utilisés pour modéliser des séries temporelles présentant une volatilité instantanée dépendante du passé. Ils permettent une prévision dynamique de la variance du processus. Enfin, le test de Breush-Pagan nous permet de vérifier la présence d'hétéroscédasticité dans nos résidus<sup>4</sup>. Il conclut que le seul modèle 3 a des résidus d'hétéroscédastiques, tandis que ceux des modèles 1 et 2 sont homoscedastiques.

Table 2: Résultats des Modèles ARMA trouvés

	Modèle 1 trouvé	Modèle 2 trouvé	Modèle 3 trouvé
AR1	0.9953*** (0.0091)	1.0315*** (0.0126)	0.9769*** (0.0212)
AR9	---	---	0.5990*** (0.1744)
AR10	---	-0.0051 (0.0300)	-0.6077*** (0.1693)
AR12	0.6433*** (0.0410)	0.4106*** (0.0583)	0.6679*** (0.0465)
AR13	-0.6791*** (0.0387)	-0.5047*** (0.0502)	-0.6795*** (0.0446)
AR21	---	---	-0.3473** (0.1221)
AR22	---	---	0.3670** (0.1202)
AR24	---	0.3496*** (0.0554)	---
AR25	---	-0.3131*** (0.0534)	---
MA1	0.0933 (0.0526)	---	0.1218** (0.0399)
MA2	-0.0042 (0.0466)	---	---
MA3	-0.0069 (0.0480)	---	---
MA4	---	---	0.0786 (0.0459)
MA6	---	---	0.0945 (0.0508)
MA7	---	---	0.1167 (0.0570)
MA8	---	---	0.0707 (0.0458)
MA9	---	---	-0.6150*** (0.1561)
AIC	-26.86	-62.22	-28.45
BIC	1.22	-34.14	27.71
Ljung-Box p-value	0.02396	0.396	0.4324
Jarque-Bera p-value	$2.62e - 14$	$< 2.2e - 16$	$7.673e - 12$
ARCH p-value	$1.956e - 14$	$5.545e - 08$	$1.253e - 14$
Breusch-Pagan p-value	0.1822	0.08159	0.03369

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

L'utilisation des critères AIC et BIC nous conduit à retenir le modèle 2, c'est-à-dire celui issu de l'AR (25). En effet, ce modèle présente un AIC et un BIC nettement inférieur aux autres, ce qui indique qu'il propose le meilleur compromis entre précision et simplicité. En outre, ce modèle a des résidus sphériques (non autocorrélation et homoscedasticité), ce qui rend l'inférence fiable. Nous constatons que la composante AR10 n'est plus significative, et nous estimons à nouveau le modèle 2 sans elle. Nous adoptons le modèle présenté ci-dessous, que nous utiliserons dans la section suivante consacrée aux prévisions.

<sup>2</sup>L'hypothèse nulle est la normalité des résidus

<sup>3</sup>L'hypothèse nulle est l'absence d'hétéroscédasticité conditionnelle

<sup>4</sup>L'hypothèse nulle est l'absence d'hétéroscédasticité

Table 3: Résultats du Modèle AR(13)

Modèle 2 (AR 25)	
AR1	1.0307*** (0.0118)
AR12	0.4062*** (0.0523)
AR13	-0.5049*** (0.0502)
AR24	0.3502*** (0.0553)
AR25	-0.3134*** (0.0533)
$\sigma^2$	0.04741
Log-vraisemblance	38.09
AIC	-64.19
BIC	-40.12
Test de Ljung-Box (p-value)	0.3957
Test de Jarque-Bera (p-value)	$< 2.2e - 16$
Test ARCH (p-value)	$6.521e - 08$
Test de Breusch-Pagan (p-value)	0.0827
Note:	*p<0.1; **p<0.05; ***p<0.01

## 6 Prévision

Dans cette section, nous utilisons le modèle 2 ajusté afin de prédire les huit prochaines valeurs de l'IPC. Ces valeurs seront ensuite comparées aux valeurs de l'IPC de janvier à août 2024 afin d'apprécier le pouvoir prédictif du modèle.

Pour prédire les valeurs futures de l'IPC, nous procédons comme suit :

1. Au moyen de notre modèle *benchmark* nous prédisons les huit valeurs futures de notre série. Celles-ci sont les prévisions des chocs, des écarts futurs à la tendance.
2. En nous servant du modèle utilisé pour extraire la tendance de la série, nous prédisons les valeurs de l'IPC pour les mois de janvier à août 2024.
3. Enfin, en additionnant valeur prévue et choc prévu, nous obtenons une prédiction précise de l'indice des prix à la consommation pour chacun des huit premiers mois de l'année 2024.

Table 4: Prévisions de l'IPC et erreurs relatives

Date	IPC prédit	Écart prédit	Observation	Prédiction	Erreur	Erreur relative
2024-01-01	111.226	6.360	117.160	117.586	-0.426	-0.364%
2024-02-01	111.328	6.824	118.150	118.153	-0.003	-0.002%
2024-03-01	111.431	7.335	118.400	118.767	-0.367	-0.310%
2024-04-01	111.534	7.397	119.010	118.931	0.079	0.066%
2024-05-01	111.637	7.203	119.050	118.840	0.210	0.177%
2024-06-01	111.739	7.190	119.140	118.929	0.211	0.177%
2024-07-01	111.842	6.942	119.370	118.784	0.586	0.491%
2024-08-01	111.945	7.232	120.010	119.177	0.833	0.694%

Nous constatons que les erreurs commises sont toujours d'amplitude inférieure à l'unité. Avec ces valeurs, nous obtenons une erreur quadratique moyenne de 0.18 et à une erreur absolue moyenne de 0.39. En termes relatifs, les erreurs de prédiction sont toujours d'un ordre inférieur à 1%, ce que nous trouvons très satisfaisant. Cette section conclut la partie purement académique de ce rapport. Dans la section suivante, nous proposons des prolongements liés à la nature économique de la série en présence.

## 7 Prolongements

Dans cette section, nous abordons la question de l'IPC avec un regard plus économique. De fait, le niveau des prix ne représente pas une variable d'intérêts dans l'analyse macroéconomique. On lui préférera notamment l'inflation, représentée par la variation relative de l'indice de prix à la consommation. La modélisation de d'agrégats macroéconomiques tels que l'inflation, et l'utilisation de modèles prédictifs sont au centre des analyses qu'effectuent des organismes tels que l'OCDE ou l'OFCE. Une partie de cette section sera dédiée à la prédiction de l'inflation.

De plus, l'indice des prix à la consommation a pour caractéristique d'avoir une forte saisonnalité. En effet, le prix d'un panier de bien peut varier en fonction de la période de l'année, du fait, par exemple, de l'influence des conditions climatiques sur l'approvisionnement les conditions d'approvisionnement. Les prix de l'énergie (gaz, pétrole, etc.), sont souvent hauts pendant l'hiver, car la demande de chauffage est forte. Ceci influe à la hausse sur l'IPC, toutes choses égales par ailleurs. Dans une seconde partie, nous aborderons la question de la saisonnalité de l'IPC.

### 7.1 L'inflation

Afin de prédire les valeurs futures du taux d'inflation, il convient de commencer par expliquer le lien entre IPC et inflation. En effet, l'IPC étant une indication niveau général des prix, le taux de croissance de l'IPC correspond au taux d'inflation. Aussi, pouvons-nous écrire :

$$\text{inflation} = \Delta \log IPC \quad (\text{p1})$$

En appliquant cette transformation, nous obtenons la série des taux d'inflation dont nous observons les corrélogrammes simples et corrélogrammes partiels, et dont nous étudions la stationnarité au moyen du test ADF.

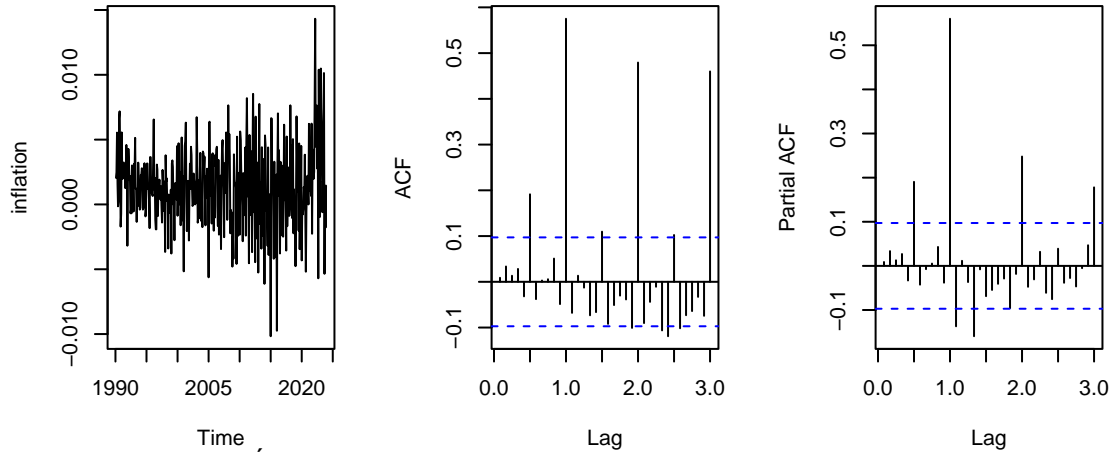


Figure 6: Évolution, ACF et PACF de la série temporelle des taux d'inflation

Les résultats du test ADF permettent de conclure que la série est stationnaire en niveau, mais avec une déviation (*drift*). Nous estimons ensuite un modèle AR (13), résultat d'une procédure de sélection similaire à celle déployée pour l'indice des prix. Le modèle estimé est donc :

$$\Delta \log IPC = \alpha + \sum_{i=1}^{13} \phi_i \Delta \log IPC_{t-i} + \varepsilon_t \quad (\text{p1.1})$$

où la composante  $\alpha$  permet de stationnariser la série.

Table 5: Résultats du Modèle AR(13)

	<b>AR 13</b>
AR1	0.1344**(0.0490)
AR2	-0.0112(0.0391)
AR3	0.0160(0.0390)
AR4	0.0064(0.0392)
AR5	-0.0116(0.0395)
AR6	0.0822**(0.0389)
AR7	-0.0077(0.0398)
AR8	-0.0035(0.0398)
AR9	-0.0203(0.0398)
AR10	0.0421(0.0402)
AR11	-0.0414(0.0404)
AR12	0.6118*** (0.0402)
AR13	-0.1494**(0.0510)
Intercept	0.0014*** (0.0003)
$\sigma^2$	$5.711e - 06$
Log-vraisemblance	1876.29
AIC	-3724.59
Test de Ljung-Box (p-value)	0.2377
Test de Jarque-Bera (p-value)	0.0004028
Test ARCH (p-value)	$9.367e - 12$
Test de Breusch-Pagan (p-value)	0.07161
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

Contrairement à ce que nous avons fait dans la section précédente, nous ne recourons au modèle trouvé car cela crée de l'hétéroscédasticité dans notre cas. Notre modèle a des résidus sphériques, malgré une hétéroscédasticité conditionnelle, et non-normaux.

Table 6: Erreurs de Prédiction de l'Inflation

Date	Observation	Prédiction	Erreur	Erreur relative
2024-01-01	-0.00841	0.00290	-0.011	134.446%
2024-02-01	-0.00211	0.00754	-0.010	456.608%
2024-03-01	-0.00514	0.00365	-0.009	170.982%
2024-04-01	-0.00034	0.00372	-0.004	1208.453%
2024-05-01	-0.00076	-0.00056	-0.000	25.843%
2024-06-01	-0.00193	0.00230	-0.004	219.141%
2024-07-01	-0.00535	0.00042	-0.006	107.945%

Le modèle ainsi obtenu est utilisé pour réaliser des prédictions (sur janvier à juillet 2024) dont les résultats, comparés aux observations faites en 2024, sont beaucoup trop imprécis. Le modèle se trompe en moyenne de 332%, ce qui traduit bien son pouvoir prédictif médiocre. Ceci semble indiquer que la dynamique inflationniste est le résultat d'interactions complexes entre des variables tant macroéconomiques que microéconomiques, ce qui la rend particulièrement difficile à modéliser sur base de ses seules valeurs passées.

Nous pouvons à présent nous intéresser au second aspect retenu dans cette section : la saisonnalité de la série.

## 7.2 La saisonnalité

L'indice des prix à la consommation (IPC) présente une saisonnalité structurelle, résultant de fluctuations récurrentes des prix liées aux variations de l'offre et de la demande tout au long de l'année. Cette dynamique est influencée par des facteurs climatiques, culturels et économiques, qui engendrent des cycles prévisibles dans l'évolution des prix de certains biens et services. Par exemple, les produits alimentaires frais, tels que les fruits et légumes, connaissent des variations saisonnières en raison des cycles de production agricole : en hiver, la rareté de certaines denrées locales entraîne une hausse des prix, tandis qu'en été, une production plus abondante tend à les réduire. De même, le secteur du tourisme et des transports est marqué par des hausses tarifaires significatives durant les périodes de vacances scolaires et les congés estivaux, lorsque la demande pour les déplacements et l'hébergement s'intensifie. Ces variations cycliques, observables chaque année, justifient l'ajustement de l'IPC pour la saisonnalité, permettant ainsi une meilleure analyse des tendances sous-jacentes de l'inflation. Dans cette section nous transformons notre série afin d'obtenir l'inflation en glissement annuel.

L'utilisation de l'inflation en glissement annuel, définie par la formule :

$$\text{Inflation} = \frac{\text{IPC}_t - \text{IPC}_{t-12}}{\text{IPC}_{t-12}} \quad (\text{p2})$$

permet de résoudre le problème de saisonnalité inhérent à l'indice des prix à la consommation (IPC). En comparant le niveau des prix d'un mois donné à celui du même mois de l'année précédente, cette approche neutralise les fluctuations saisonnières récurrentes qui pourraient fausser l'interprétation des variations de prix à court terme. Par exemple, si l'on compare les prix de décembre à ceux de novembre, une hausse peut simplement refléter l'augmentation saisonnière des prix liée aux fêtes de fin d'année et non une tendance inflationniste réelle. En revanche, en comparant décembre à décembre, on s'assure que l'effet saisonnier est présent dans les deux valeurs et que la variation mesurée est davantage attribuable aux dynamiques de l'inflation sous-jacente plutôt qu'à des variations cycliques. Cette méthode permet donc d'obtenir une vision plus fiable et plus pertinente de l'évolution des prix à long terme.

La figure 7 donne les ACF et PACF de la série ainsi obtenue. Au moyen d'un test ADF, nous étudions la stationnarité de la série et concluons qu'elle est stationnaire avec une déviation. Nous estimons un modèle ARMA (24,6)<sup>5</sup> dont nous retirons quelques coefficients non significatifs de manière à garder des résidus sphériques. Les résultats sont présentés dans la table ci-dessous.

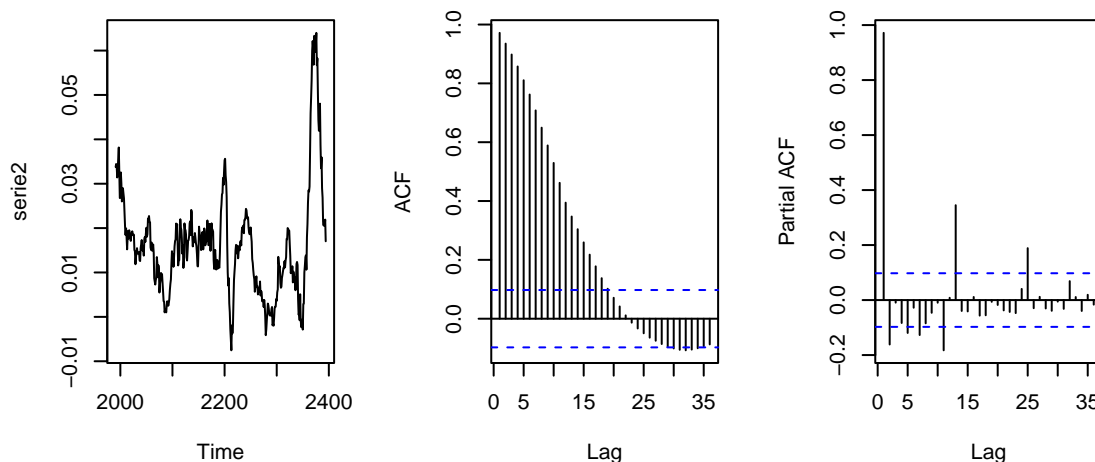


Figure 7 : Évolution, ACF et PACF de la série temporelle des taux d'inflation en glissement annuel

<sup>5</sup>Avec une déviation à l'instar de l'équation p1.1



Table 7: Résultats du Modèle AR(24,6)

	AR(24,6)
Intercept	0.0158*** (0.0020)
AR2	1.0894*** (0.0647)
AR4	-0.6634*** (0.1352)
AR5	0.0521 (0.0556)
AR6	0.6188*** (0.1124)
AR12	-0.5874*** (0.0550)
AR14	0.4862*** (0.0703)
AR15	-0.0456 (0.0401)
AR16	-0.3658*** (0.0822)
AR17	0.0610 (0.0464)
AR18	0.3246*** (0.0646)
AR24	-0.1632*** (0.0367)
MA1	1.0998*** (0.0506)
MA3	-0.1425 (0.0772)
MA4	0.6053*** (0.1157)
MA5	0.6874*** (0.1251)
MA6	0.0946 (0.0571)
$\sigma^2$	$5.409e - 06$
Log-vraisemblance	1872.51
AIC	-3711.01
Test de Ljung-Box (p-value)	0.9781
Test de Jarque-Bera (p-value)	$5.717e - 08$
Test ARCH (p-value)	0.005889
Test de Breusch-Pagan (p-value)	0.05553
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

En utilisant ce modèle, nous pouvons prédire les valeurs futures de l'inflation annuelle glissante. Ces prédictions, présentées sur la table 8, nous permettent d'apprécier la précision du modèle. Il apparaît que le modèle est plus précis que le modèle employé dans la section précédente. En moyenne, l'erreur relative est de l'ordre de 45.5%, un résultat nettement meilleur que les 332% du modèle précédent. Pour autant, les erreurs restent très importantes, traduisant l'influence importante d'autres variables économiques sur le processus de inflationniste. On peut toutefois s'interroger sur les raisons pour lesquelles nos prédictions des IPC sont si précises alors que celles de l'inflation ne le sont pas, ce qui soulève une question que nous laissons à l'appréciation des auteurs.

Table 8: Prédictions et erreurs du modèle

Date	Observation	Prediction	Erreur	Erreur relative
2024-01-01	0.02898	0.01416	0.0148	51.13%
2024-02-01	0.02686	0.01561	0.0112	41.87%
2024-03-01	0.02139	0.01544	0.0059	27.82%
2024-04-01	0.02058	0.01065	0.0099	48.27%
2024-05-01	0.02154	0.01188	0.0097	44.84%
2024-06-01	0.02047	0.00958	0.0109	53.22%
2024-07-01	0.02192	0.01166	0.0103	46.79%
2024-08-01	0.01703	0.00846	0.0086	50.32%

## Bibliographie

Artur Bolganbayev, Baltain Sabenova, Gulmira Mombekova, Gulnur Sultankhanova, and Tazhibayeva Raikhan Musamatovna. The effect of electricity generation, thermal energy production, fixed capital investment, and consumer price index on economic growth in kazakhstan. *International Journal of Energy Economics and Policy*, 12(6):67–72, 2022.

R Davidson. Econometric theory and methods, 2004.

Jama Mohamed. Time series modeling and forecasting of somaliland consumer price index: a comparison of arima and regression with arima errors. *American Journal of Theoretical and Applied Statistics*, 9(4): 143–153, 2020.

G William Schwert. Tests for unit roots: A monte carlo investigation. *Journal of business & economic statistics*, 20(1):5–17, 2002.