


<u>Université</u> : UADB		<u>Etablissement</u> : UFR/SATIC
<u>Domaine</u> : SCIENCES ET TECHNOLOGIES		Niveau : <i>Master I</i>
<u>Mention</u> : MATHEMATIQUES APPLIQUEES		<u>Spécialité</u> : Statistique et Informatique Décisionnelle (SID)

Projet de séries temporelles : Modélisation de l'Indice de Production Industrielle du Sénégal



Présenté par :

Dalyo Sid Ousmane OURBA

Enseignant:

Dr. Souleymane FOFANA

Table des matières

Sigles et Abréviations.....	III
Tables des illustrations	IV
Liste des tableaux	IV
Liste des figures	IV
Introduction générale.....	1
Chapitre 1 : Cadre théorique et méthodologique	2
1.1 Revue de littérature.....	2
1.2 Présentation des données	2
1.2.1 Source de données	2
1.2.2 Présentation des variables	2
1.3 Méthodologie d'étude.....	2
1.3.1 Modélisation univariée	3
1.3.2 Modélisation multivariée.....	3
Chapitre 2 : Modélisation univariée de l'Indice de Production Industrielle du Sénégal.....	4
2.1 Analyse descriptive.....	4
2.2 Tests de stationnarité, Saisonnalité et corrections	5
2.2.1 Tests de stationnarité.....	5
2.2.2 Tests de Saisonnalité.....	8
2.3 Spécification du modèle	10
2.3.1 Détermination du paramètre q	10
2.3.2 Détermination du paramètre p	10
2.3.3 Estimation et Validation du modèle.....	11
2.3.2 Analyse des résidus.....	12
2.4 Prévisions et interprétations des résultats	15
Chapitre 3 : Modélisation multivariée de l'indice de Production Industrielle du Sénégal.....	17
3.1 Analyse descriptive de l'IPC	17

3.2 Tests de stationnarité de variable IPC et ordres d'intégration.....	18
3.3 Etude multivariée.....	21
3.3.1 Identification du nombre de retards du modèle	22
3.3.2 Estimations des paramètres et validation du modèle multivarié.....	22
a) Significativité des paramètres.....	22
b) Stabilité.....	25
c) Analyse des résidus	26
d) Causalité entre variables	27
e) Les fonctions de demande impulsionnelles	28
f) Décomposition de la variance.....	28
e) Prévisions.....	29
Conclusion.....	31
Bibliographie	A
Codes R utilisés	B

Sigles et Abréviations

ACF	Autocorrelation Function
ADF	Augmented Dickey–Fuller/
AIC	Akaike Information Criterion
ANSD	Agence Nationale de la Statistique et de la Démographie
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
FPE	Final Prediction error
HQ	Hann Quinn Information Criterion
IPC	Indice de la Production du Consommateur
IPI	Indice de la Production Industrielle
KPSS	Kwiatkowski-Phillips-Schmidt-Shin
PACF	Partial Autocorrelation Function
SARIMA	Seasonal Autoregressive Integrated Moving Average
SC	Schwarz Information Criterion
SSN	Système Statistique National
VAR	Vector autoregression/Vecteur Autoregressif

Tables des illustrations

Liste des tableaux

Tableau 1- listes des combinaisons SARIMA.....	11
Tableau 2- liste des modèles SARIMA	12
Tableau 3- des prévisions de l'IPI de 2017.....	15
Tableau 4 - Retards optimaux selon chaque critère d'information	22
Tableau 5 - liste des différentes valeurs de chaque critère en fonction des retards.....	22

Liste des figures

Figure 1 – Visualisation de la série IPI	4
Figure 2 - Visualisation de la série log(IPI) centrée.....	5
Figure 3 - Visualisation de la tendance du log (IPI) centrée	6
Figure 4 - fonction d'autocorrélation de la série log(IPI) centrée	6
Figure 5 - Visualisation de la tendance de la série log (IPI) centrée et différenciée.....	7
Figure 6 - fonction d'autocorrélation de la série log(IPI) centrée et différenciée	8
Figure 7 - fonction d'autocorrélation de la série log(IPI) centrée et différenciée et désaisonnalisée.....	9
Figure 8 - Visualisation de la tendance du log (IPI) centrée, différenciée et de desaisonnalisée	9
Figure 9 - fonction d'autocorrélation partielle de la série log(IPI) centrée, différenciée et désaisonnalisée.....	11
Figure 10 - Visualisation des résidus	13
Figure 11 - Fonction d'autocorrélation et Statistique de Ljung-Box des résidus	14
Figure 12- Prévisions de l'année 2017.....	16
Figure 13 - Prévisions de l'année 2017 et 2018.....	16
Figure 14 - Visualisation de la série IPC.....	17
Figure 15 - Visualisation de la série log (IPC) centrée	18
Figure 16 - Visualisation de la tendance du log (IPC) centrée.....	19
Figure 17- fonction d'autocorrélation de la série log(IPC) centrée	19
Figure 18 - Visualisation de la tendance de la série log (IPC) centrée et différenciée	20
Figure 19 - Observation simultanée de IPI et IPC	21

Figure 20 - Figure 19 - Observation simultanée de $\log(\text{IPI})$ centrée différenciée et $\log(\text{IPC})$ centrée différenciée	21
Figure 21 - Représentation de la série $\log(\text{IPI})$ centrée différenciée avec ses résidus	24
Figure 22 - Représentation de la série $\log(\text{IPC})$ centrée différenciée avec ses résidus.....	25
Figure 23 - Stabilité du modèle	26
Figure 24 - Visualisation des résidus	27
Figure 25 - Fonction d'impulsion	28
Figure 26 - Décomposition de la variance des erreurs des prévisions	29
Figure 27 - Prévision pour 12 mois	29
Figure 28 - Prévision pour 24 mois	30

Introduction générale

L'indice de la production industrielle se définit comme un indicateur qui permet de mesurer l'évolution de la production des unités industrielles exerçant sur le territoire national à une période bien définie. Il s'intéresse à l'activité de fabrication ou de transformation des unités industrielles et permet de mesurer les quantités physiques produites par ces unités au cours d'une période donnée. Il est aussi censé donner par conséquent l'évolution en volume de la production industrielle et présenter de façon assez représentative les mouvements observés au sein du tissu industriel. De plus, il s'intéresse au volume de la production des unités résidentes qui mènent une activité industrielle à titre principal ou secondaire. Il s'agit des entreprises et établissements industriels (de fabrication ou de transformation) qui résident et produisent dans le pays. En d'autres termes, L'indice de la production industrielle se rapporte aux unités industrielles exerçant leurs activités dans les usines, des chantiers, des carrières, des mines, etc. Cet indice, en effet, est l'un des principaux instruments statistiques qu'utilisent les pouvoirs publics et les responsables de l'économie pour dresser chaque mois le diagnostic de la situation économique.

Chapitre 1 : Cadre théorique et méthodologique

1.1 Revue de littérature

L'Indice de Production Industrielle (IPI), indicateur économique à court terme des statistiques officielles, est important à la fois en soi et parce qu'il sert à établir des comparaisons avec d'autres indicateurs à court terme, ou, en même temps que ces indicateurs, pour mesurer la performance d'une économie. Ainsi, plusieurs variables peuvent impacter sur son évolution. C'est le cas de l'Indice de Production du Consommateur (IPC). En effet, L'IPC représente la mesure officielle du taux d'inflation. L'augmentation du niveau général des prix peut être perçue par les industriels comme un facteur qui détériore le pouvoir d'achat des agents économiques, notamment les consommateurs qui sont sensés payer les biens produits. Ces tensions inflationnistes pourraient impacter négativement la production industrielle. De ce fait, il est attendu un coefficient négatif associé à l'inflation.

1.2 Présentation des données

1.2.1 Source de données

Le Sénégal, pays de l'Afrique de l'Ouest, est doté d'un Système Statistique National (SSN) dont le principal coordonnateur est l'Agence Nationale de la Statistique et de la Démographie (ANSD) créée en 2005. Toutes les statistiques publiques sont sous son contrôle.

1.2.2 Présentation des variables

Les données d'étude comportent deux variables parmi lesquelles l'indice de production industrielle (IPI) et l'indice de production du consommateur (IPC).

1.3 Méthodologie d'étude

Ce travail a pour objectif la modélisation de l'IPI du Sénégal allant du 1 janvier 2008 au 1 décembre 2017. Pour ce faire, une modélisation univariée démarrera l'étude. Ainsi, la série sur l'IPI servira de données d'étude. Ensuite, une modélisation multivariée sera effectuée pour expliquer les variations de l'IPI à l'aide de l'IPC.

1.3.1 Modélisation univariée

Nous adopterons la modélisation Box-Jenkins dans la suite de notre étude. Elle comprend 5 étapes : la familiarisation avec les données, l'identification du modèle, la validation du modèle, l'estimation des paramètres et les prévisions. Plus explicitement, elle se fera comme suit :

- Analyse descriptive de la variable
- Tests de stationnarité ainsi que son éventuelle correction
- Enumération de modèles possibles grâce aux graphiques (fonction d'auto corrélation, fonction d'autocorrélation partielle)
- Identification du modèle
- Validation du modèle grâce aux tests de significativité des coefficients et l'analyse des résidus
- Estimations des coefficients du modèle choisi pour ajuster la série
- Prévisions de l'IPI.

1.3.2 Modélisation multivariée

La modélisation multivariée consiste à trouver une forme fonctionnelle entre les variables IPI et IPC. Pour ce faire, une modélisation vectorielle de type VAR est réalisé.

Chapitre 2 : Modélisation univariée de l'Indice de Production Industrielle du Sénégal

Dans cette partie, il est question d'une modélisation temporelle univariée de la variable IPI. Ensuite, des tests stationnarité seront effectués dans ce processus de modélisation. Enfin, on procèdera à l'estimation des paramètres et l'analyse des résultats. Dans toute cette section on ne considèrera pour la modélisation que les valeurs de la série allant de janvier 2008 à décembre 2016.

2.1 Analyse descriptive

Il ressort principalement qu'en moyenne le Sénégal à un IPI de 111,72 entre janvier 2008 et décembre 2016. Ainsi, on remarque que notre série a une évolution comprise entre 82,69 et 139,72.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
82.69	102.66	112.36	111.72	120.37	139.72

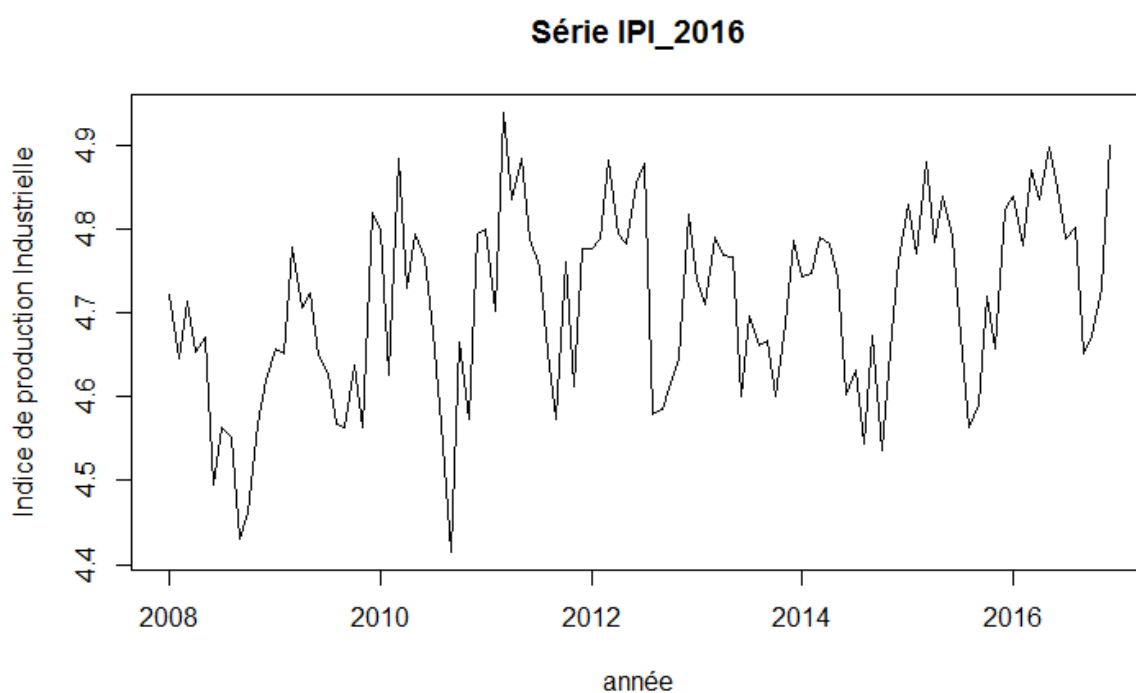


Figure 1 – Visualisation de la série IPI

La figure 1 nous montre une évolution en dent de scie tout au long de la trajectoire. En plus, on remarque que notre série présente une forte dispersion avec un écart type de 12,26. Dans la suite de notre travail, nous avons alors centrée notre série et appliqué une transformation logarithmique pour stabiliser la variance.

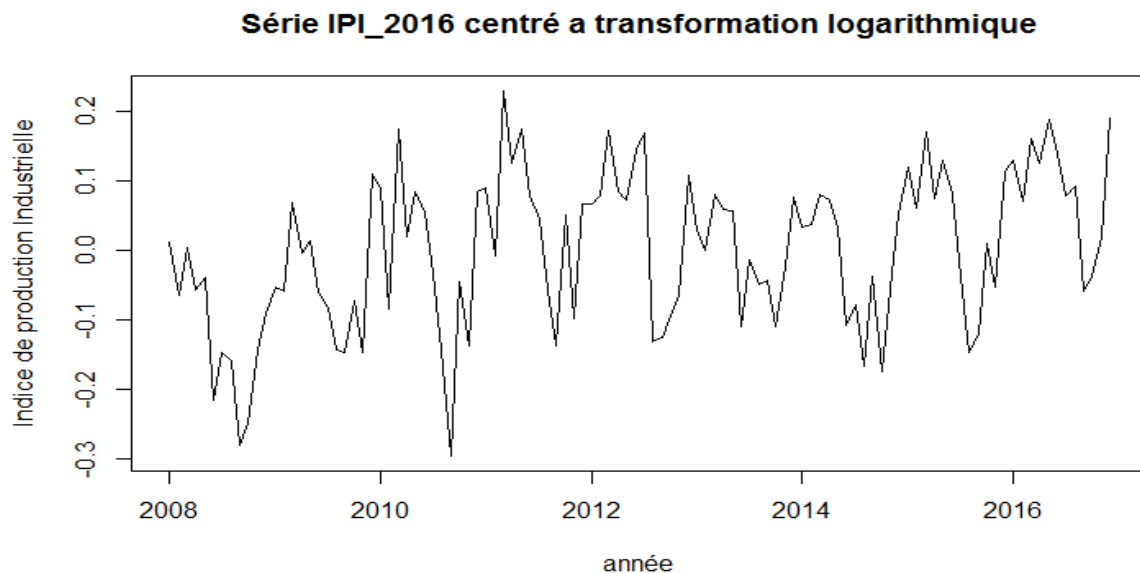


Figure 2 - Visualisation de la série $\log(IPI)$ centrée

2.2 Tests de stationnarité, Saisonnalité et corrections

2.2.1 Tests de stationnarité

La stationnarité ou encore l'homogénéité est un concept clé pour la validité externe d'une régression sur séries chronologiques. En effet, elle suppose que la distribution d'une série donnée ne change pas quels que soient le nombre de dates et les dates choisis, quand on décale ces dates d'une même quantité h donnée. De manière générale, un processus sera dit stationnaire au second ordre si sa moyenne, sa variance et sa covariance sont indépendantes du temps. Ceci permettra d'étudier la série en dehors de sa période en effectuant la prévision.

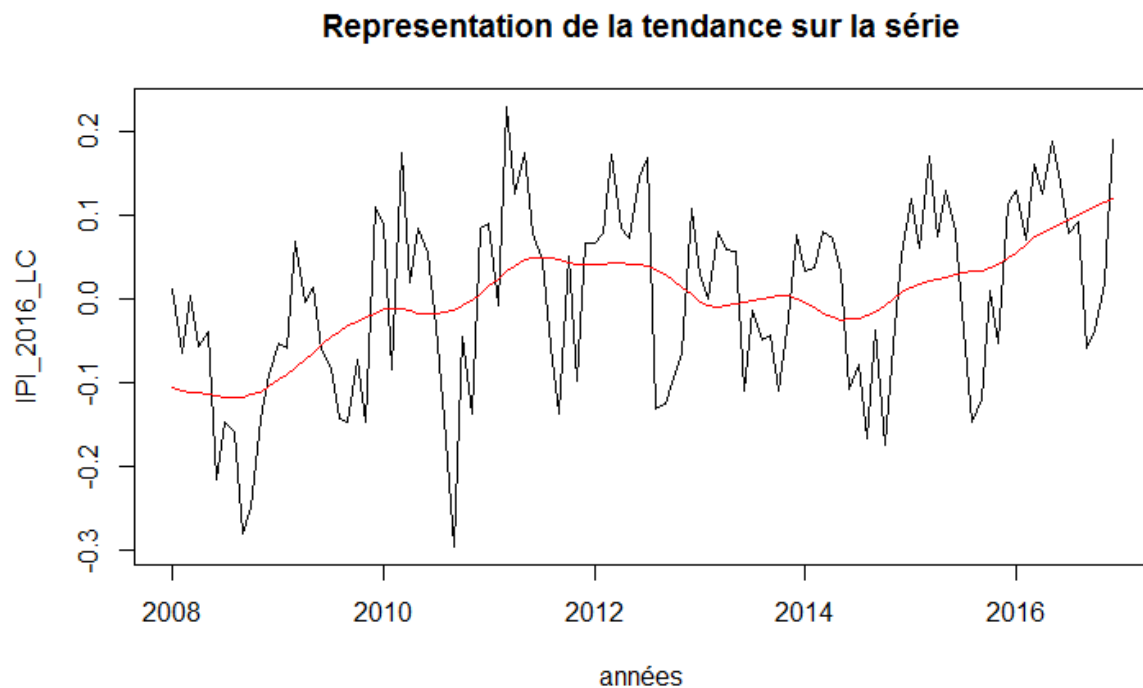


Figure 3 - Visualisation de la tendance du log (IPI) centrée

Une analyse minutieuse du graphique de la série permet à priori de postuler sur la présence d'une tendance. On remarque en effet une certaine tendance croissante.

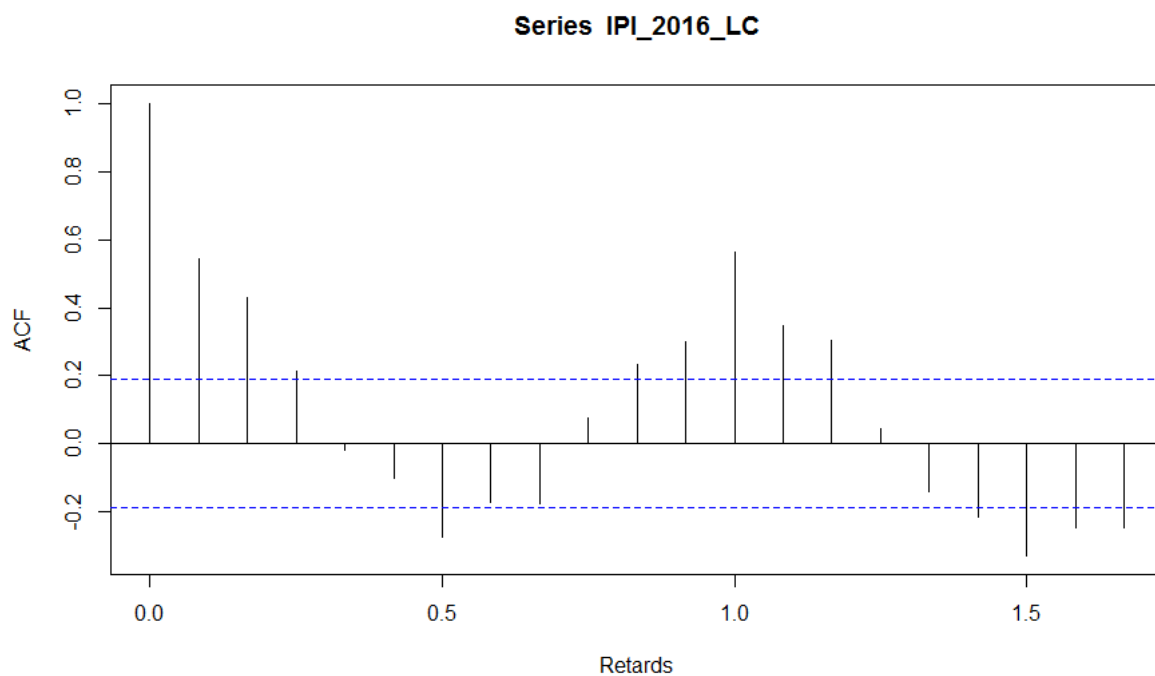


Figure 4 - fonction d'autocorrélation de la série log(IPI) centrée

L'ACF présente une décroissance lente vers 0, ce qui traduit un problème de non-stationnarité. Pour avoir une idée plus précise de la présence ou de l'absence de stationnarité, les tests de Dickey-Fuller augmenté, Phillips Perron et KPSS seront effectués sur la série.

- **Test augmenté de Dickey Fuller (ADF)** : Le test de Dickey-Fuller augmenté suppose au départ que la série soit non stationnaire. Les résultats de ce test nous donnent une p-value égale à 0,01. On rejette alors l'hypothèse H_0 . On peut donc affirmer avec 5% de risque de se tromper que la série est stationnaire.
- **Test de Phillips Perron** : Tout comme le test de Dickey-Fuller augmenté, le test de Phillips Perron admet pour hypothèse nulle la non stationnarité de la série. Les résultats nous donne une p-value de $0,01 < 0,05$. On peut donner affirmer avec un risque de 5% de se tromper que la série est stationnaire.
- **Test KPSS** : Contrairement aux deux tests précédents, le test KPSS admet comme hypothèse nulle que la série soit stationnaire. Ici la p-value de 0.03 traduit donc la non stationnarité de la série.

Les deux premiers tests ci-dessus nous montrent que la série est stationnaire. Cependant, le test de KPSS postule en faveur de la présence d'une tendance sur notre série. Pour uniformiser tous les tests, on effectue donc une différenciation de type $(I-B)$. La reprise de ses trois tests nous donne respectivement des p-values de 0,01, 0,01 et 0,1. Nous pouvons donc conclure que la série est stationnaire. Ainsi, la moyenne et la variance de la série ne dépendent pas du temps. Le graphique nous permet de confirmer visuellement cette stationnarité.

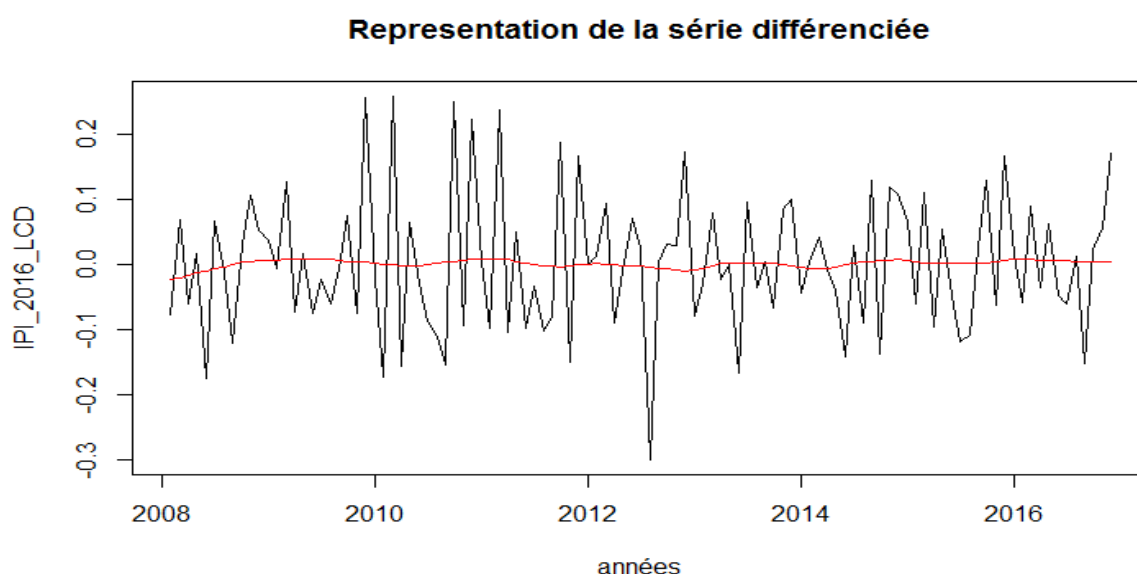


Figure 5 - Visualisation de la tendance de la série log (IPI) centrée et différenciée

2.2.2 Tests de Saisonnalité

La saisonnalité est une variation cyclique prévisible dépendant de la période de l'année. La nature de nos données étant mensuelle, ce qui nous permet de soupçonner de la présence d'une saisonnalité. Pour juger de la présence ou l'absence de la saisonnalité, nous allons observer l'ACF de la nôtre série stationnaire.

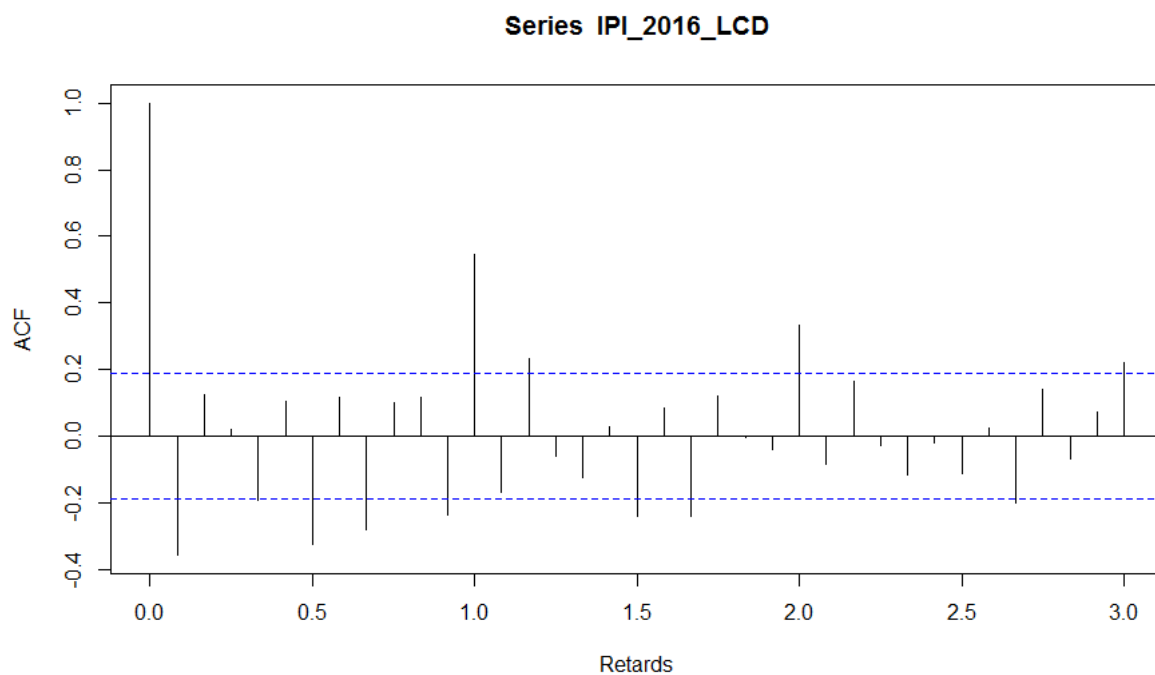


Figure 6 - fonction d'autocorrélation de la série $\log(IPI)$ centrée et différenciée

La série ainsi différenciée présente encore une décroissance lente vers 0 pour les multiples de 12. Cela traduit donc la présence d'une saisonnalité. On effectue cette fois une différenciation de type $(I - B^{12})$.

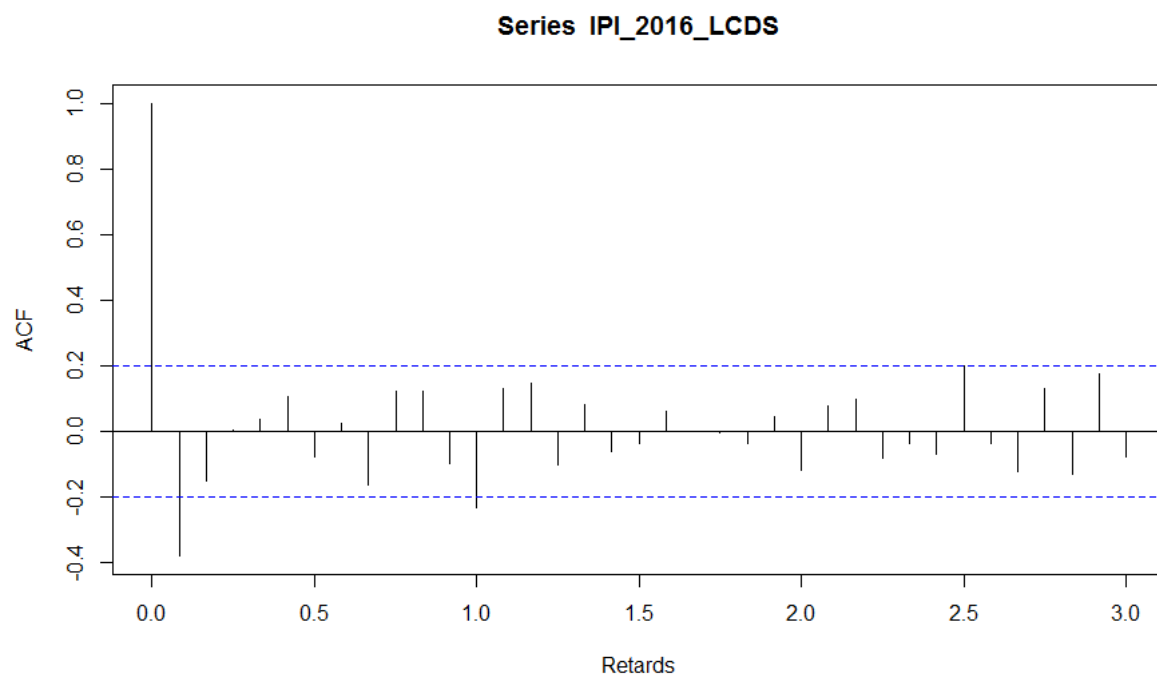


Figure 7 - fonction d'autocorrélation de la série $\log(IPI)$ centrée et différenciée et désaisonnalisée

L'ACF de la figure 7 semble pouvoir être interprétée comme un autocorrélogramme simple empirique.

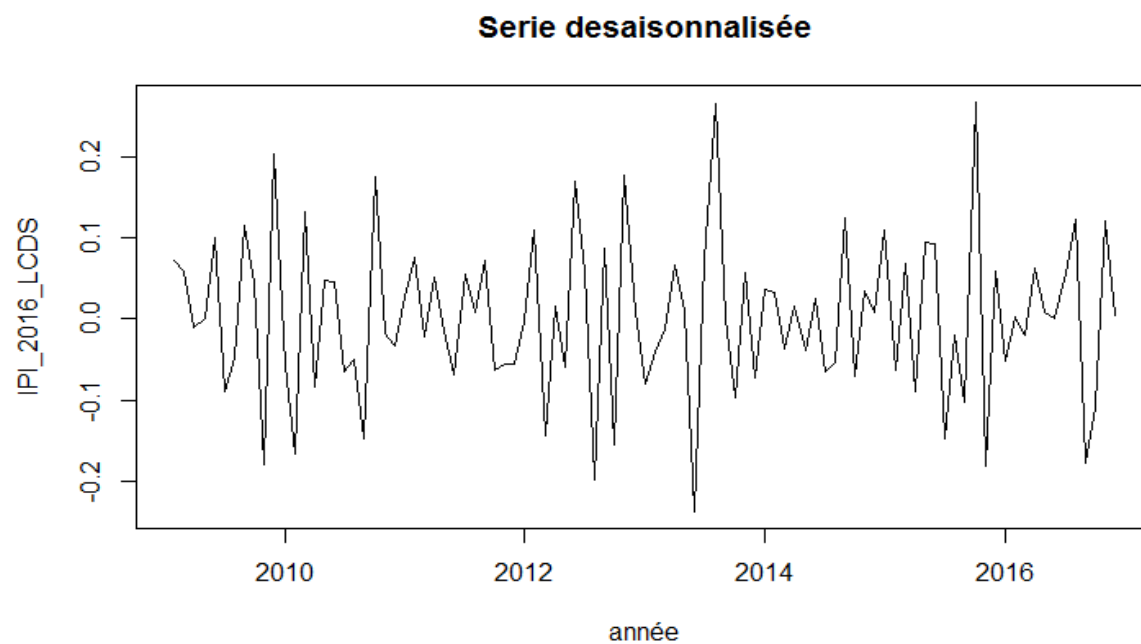


Figure 8 - Visualisation de la tendance du $\log(IPI)$ centrée, différenciée et de désaisonnalisée

2.3 Spécification du modèle

L'identification du modèle ARMA se fait à travers l'observation de l'ACF et de la PACF de la série stationnaire désaisonnalisée. Les ordres de ces fonctions vont permettre de postuler sur les paramètres p et q du processus ARMA; p , d et q s'il s'agit d'un processus ARIMA et p , P , d , D , q et Q pour un processus SARIMA. Pour le présent modèle, la présence de saisonnalité et la différenciation effectuée plus haut nous permet déjà d'opter pour un SARIMA de paramètres $d = 1$ et $D = 1$.

2.3.1 Détermination du paramètre q

Le paramètre q est déterminé à l'aide de l'ACF. En effet, l'ACF d'un processus ARMA est nulle à partir de l'ordre $q+1$. Plus précisément, c'est le test de Bartlett qui a permis ici de déterminer q en calculant des intervalles des intervalles autour de 0 pour l'ACF à partir de l'ordre $q+1$.

En observant la figure 7, on remarque ainsi que les deux valeurs 0 et 1 sont les seuls retards significatifs qui sortent de l'intervalle de confiance. En plus, on remarque que le douzième retard est significatif ce qui permet de conclure que Q prend les valeurs 0 et 1.

2.3.2 Détermination du paramètre p

Le paramètre p est déterminé à l'aide de la PACF. En effet, la PACF d'un processus ARMA est nulle à partir de l'ordre $p+1$. Ici c'est la formule de Quenouille qui calcule des intervalles de confiance autour de 0 pour la PACF à partir de l'ordre $p+1$.

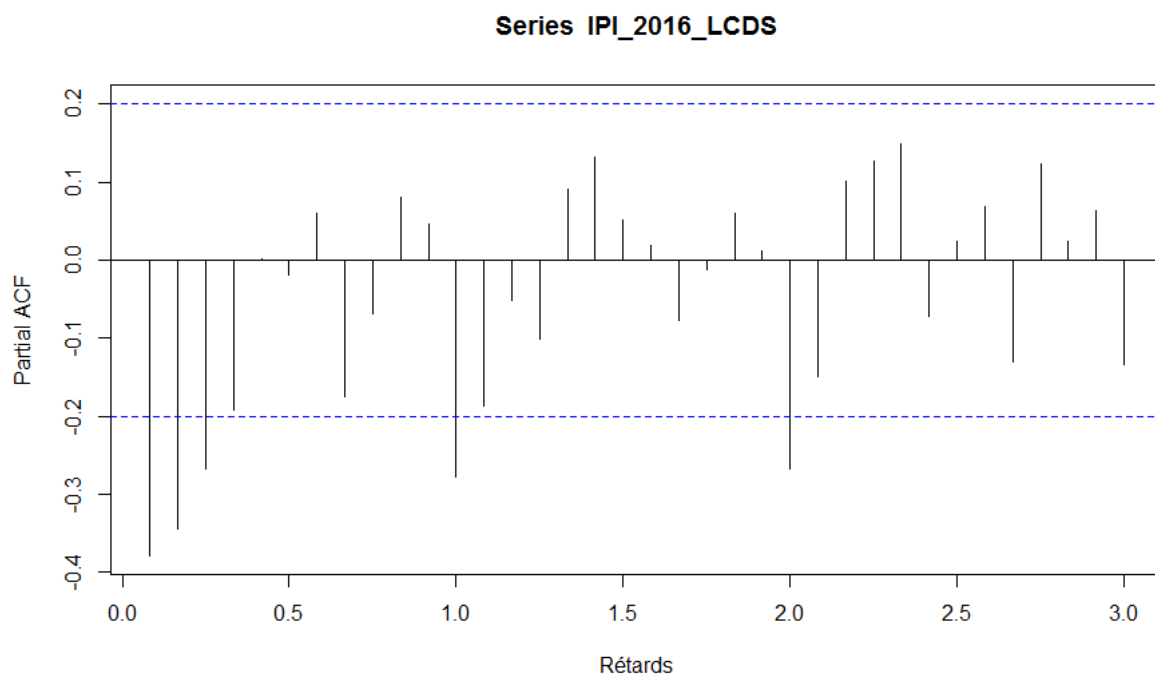


Figure 9 - fonction d'autocorrélation partielle de la série $\log(IPI)$ centrée, différenciée et désaisonnalisée

On observe ainsi trois valeurs qui sont les seuls retards qui sortent de l'intervalle de confiance. Notre p prend donc comme valeurs : 0, 1, et 2. En plus, on remarque que le douzième et le vingt-quatrième retard est significatif ce qui permet de conclure que P prend les valeurs 0, 1 et 2.

2.3.3 Estimation et Validation du modèle

L'estimation des paramètres de notre processus SARIMA se fait par la fonction `arima()`. On est donc amené à énumérer les potentiels modèles suivants:

Tableau 1- listes des combinaisons SARIMA

Modèles
SARIMA(0,1,1)(0,1,1) ₁₂
SARIMA(1,1,1)(0,1,1) ₁₂
SARIMA(2,1,1)(0,1,1) ₁₂
SARIMA(1,1,0)(0,1,1) ₁₂
SARIMA(2,1,0)(0,1,1) ₁₂
SARIMA(0,1,1)(1,1,1) ₁₂
SARIMA(1,1,1)(1,1,1) ₁₂
SARIMA(2,1,1)(1,1,1) ₁₂
SARIMA(1,1,0)(1,1,1) ₁₂
SARIMA(2,1,0)(1,1,1) ₁₂

SARIMA(0,1,1)(2,1,1) ₁₂
SARIMA(1,1,1)(2,1,1) ₁₂
SARIMA(2,1,1)(2,1,1) ₁₂
SARIMA(1,1,0)(2,1,1) ₁₂
SARIMA(2,1,0)(2,1,1) ₁₂
SARIMA(0,1,1)(1,1,0) ₁₂
SARIMA(1,1,1)(1,1,0) ₁₂
SARIMA(2,1,1)(1,1,0) ₁₂
SARIMA(1,1,0)(1,1,0) ₁₂
SARIMA(2,1,0)(1,1,0) ₁₂
SARIMA(0,1,1)(2,1,0) ₁₂
SARIMA(1,1,1)(2,1,0) ₁₂
SARIMA(2,1,1)(2,1,0) ₁₂
SARIMA(1,1,0)(2,1,0) ₁₂
SARIMA(2,1,0)(2,1,0) ₁₂

Nous avons donc estimé les différents coefficients de ces modèles pour ne retenir que le modèle le plus pertinent. Des tests de significativité des paramètres ont ainsi été effectués pour ne retenir que les modèles ayant seulement tous leurs paramètres significatifs. Les AIC des modèles ainsi retenus sont consignées dans le tableau 2 suivant :

Tableau 2- liste des modèles SARIMA

Modèles	AIC
SARIMA(0,1,1)(0,1,1) ₁₂	662.6214
SARIMA(2,1,0)(0,1,1) ₁₂	672.5845
SARIMA(0,1,1)(1,1,0) ₁₂	682.0604
SARIMA(1,1,0)(1,1,0) ₁₂	701.454
SARIMA(2,1,0)(1,1,0) ₁₂	693.3654
SARIMA(0,1,1)(2,1,0) ₁₂	677.8713
SARIMA(1,1,0)(2,1,0) ₁₂	696.2427
SARIMA(2,1,0)(2,1,0) ₁₂	689.8677

Le modèle possédant une bonne qualité d'ajustement aux données fournit une variance résiduelle faible et par conséquent un critère d'information d'Akaike faible. Il ressort donc que le modèle SARIMA(0,1,1)(0,1,1)₁₂ est celui ayant le plus petit AIC.

2.3.2 Analyse des résidus

L'économétrie des séries temporelles repose sur une hypothèse fondamentale qui est celle de considérer les résidus comme des bruits blancs, c'est-à-dire qu'il n'y a pas de

dépendance entre les erreurs et leurs passés, et que les résidus sont en moyenne nuls. Pour pouvoir apporter plus de poids au modèle prédictif, il sera question dans cette partie de vérifier d'abord la normalité, ensuite l'homoscédasticité et l'autocorrélation des résidus issus de la modélisation.

Les résidus de notre modèle SARIMA(0,1,1)(0,1,1)₁₂ se présentent comme suit :

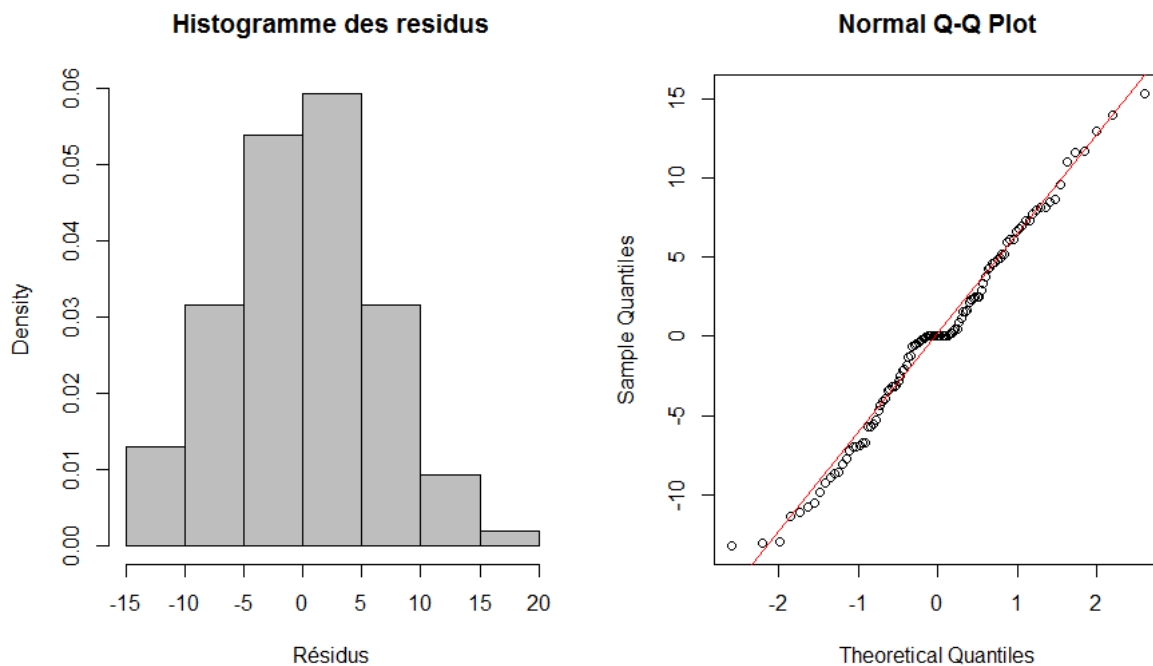


Figure 10 - Visualisation des résidus

L'histogramme des résidus et les quantiles théoriques des résidus de la figure nous poussent à penser que les résidus ici sont normaux. Aussi, Le test du t de Student dont l'hypothèse nulle stipule que les résidus sont nuls a été réalisé pour vérifier la nullité des résidus. Les résultats de cet test nous donne une p-value = 0.9975. Nous pouvons donc conclure au seuil de 5% que les résidus de notre modèle sont nuls.

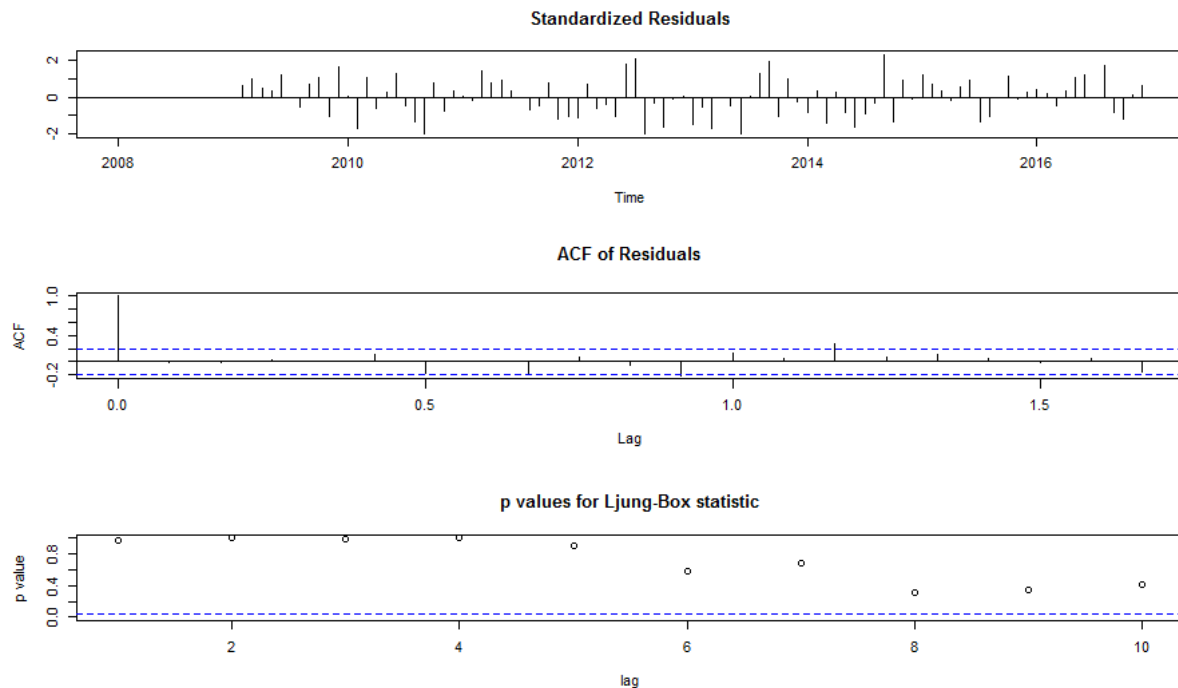


Figure 11 - Fonction d'autocorrélation et Statistique de Ljung-Box des résidus

La fonction d'autocorrélation présente des pics tous contenus à l'intérieur de l'intervalle de confiance de même que les p-values de la statistique de Ljung-Box sont toutes en dehors de l'intervalle de confiance. Pour pouvoir confirmer cette idée, des tests de normalité seront appliqués sur ces derniers.

Il existe deux familles de test pour tester la normalité des résidus :

- Les tests paramétriques (Skewness, Kurtosis, Jarque-Bera,...)
- Les tests non paramétriques (Shapiro-wilk, Kolmogorov-Smirnov,...).

Dans notre cas, nous allons utiliser le test de Jarque-Bera dont l'hypothèse nulle stipule que les résidus sont normaux. Les résultats de cet test nous donne une p-value = 0,7885. Nous pouvons donc conclure au seuil de 5% que les résidus de la modélisation suivent une loi normale.

Pour vérifier que la variance de nos résidus reste constante au cours du temps, un test d'homoscédasticité a été réalisé. Nous avons utilisé le test de White dont l'hypothèse nulle stipule l'homoscédasticité des résidus. Au seuil de 5% avec une p-value de 0,07991, le test de White ne rejette pas l'hypothèse d'homoscédasticité des résidus.

Pour tester l'autocorrélation, c'est le test de Portemanteau de Box Pierce qui sera utilisé. L'hypothèse nulle de ce test stipule que les fonctions d'autocorrélation des résidus de retard k sont nulles. Ainsi la statistique de Box-Pierce, sous H_0 : non corrélation, suit asymptotiquement

une distribution khi-deux. Avec une p-value de 0,9742, le test de Portemanteau de Box valide l'hypothèse de non autocorrélation des résidus au seuil de 5%. Le modèle SARIMA(0,1,1)(0,1,1)₁₂ sera finalement retenu pour effectuer les prévisions futures de l'IPI du Sénégal.

2.4 Prévisions et interprétations des résultats

Après avoir vérifié les hypothèses fondamentales de l'économétrie des séries temporelles sur le modèle retenu, les valeurs futures de la série vont être prédites. Soient X_t, X_{t-1}, \dots, X_1 les observations d'une série donnée. La meilleure prévision de X_t à l'horizon h conditionnellement à l'ensemble d'informations disponibles est définie par : $E(X_t/X_{t-1}, X_{t-2}, \dots, X_1)$. Puisque la série a été tronquée des valeurs de l'année 2017 avant la modélisation, nous allons donc prédire les valeurs de cette année pour avoir une idée de la fiabilité des prévisions de notre modèle.

Tableau 3- des prévisions de l'IPI de 2017

Mois de l'année 2017	Série IPI	Prévisions	Résidus
Janvier	125,786541429335	131,3583	7,024637
Février	126,140655582721	125,2353	7,220264
Mars	145,987488066375	139,8997	7,410730
Avril	129,350880663289	131,1125	7,596421
Mai	138,438668253346	134,0679	7,777681
Juin	115,79113330451	125,4529	7,954811
Juillet	111,569692640793	123,8427	8,128082
Août	120,104160151884	114,3801	8,297735
Septembre	110,318497850213	110,6424	8,463989
octobre	123,449011850511	116,4611	8,627039
Novembre	111,477424432313	116,4367	8,787065
Décembre	135,296147314566	134,0751	8,944228

Nous remarquons que les écarts entre les vraies valeurs de la série IPI en 2017 et les prévisions de notre modèle ne sont pas très grands.

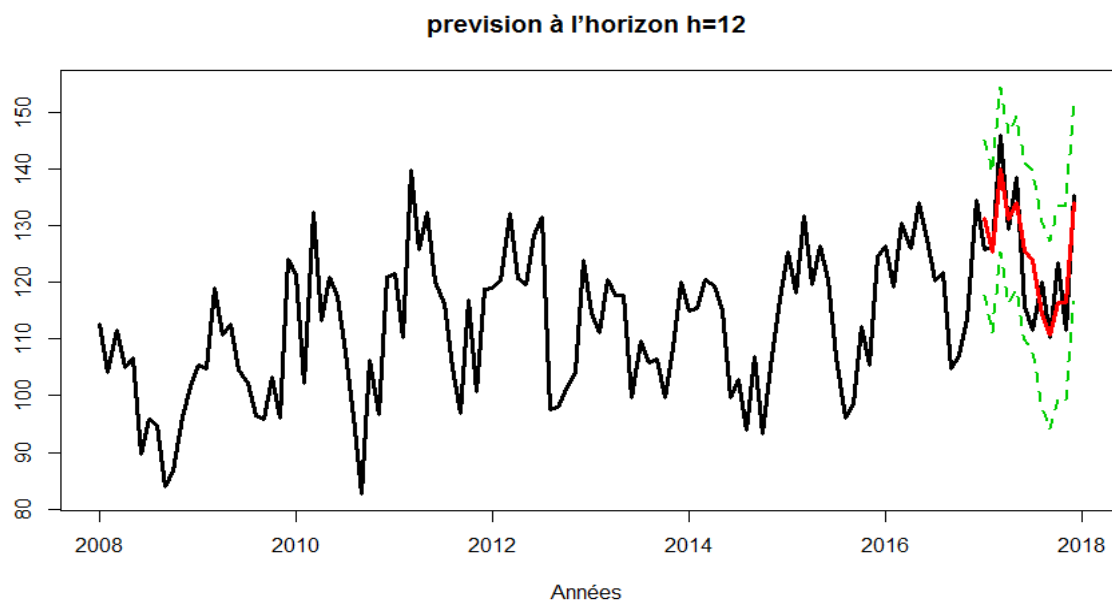


Figure 12- Prévisions de l'année 2017

Les prévisions l'IPI du Sénégal en 2017 semblent suivre la tendance de la série. Aussi la courbe se situe assez bien dans l'intervalle de confiance.

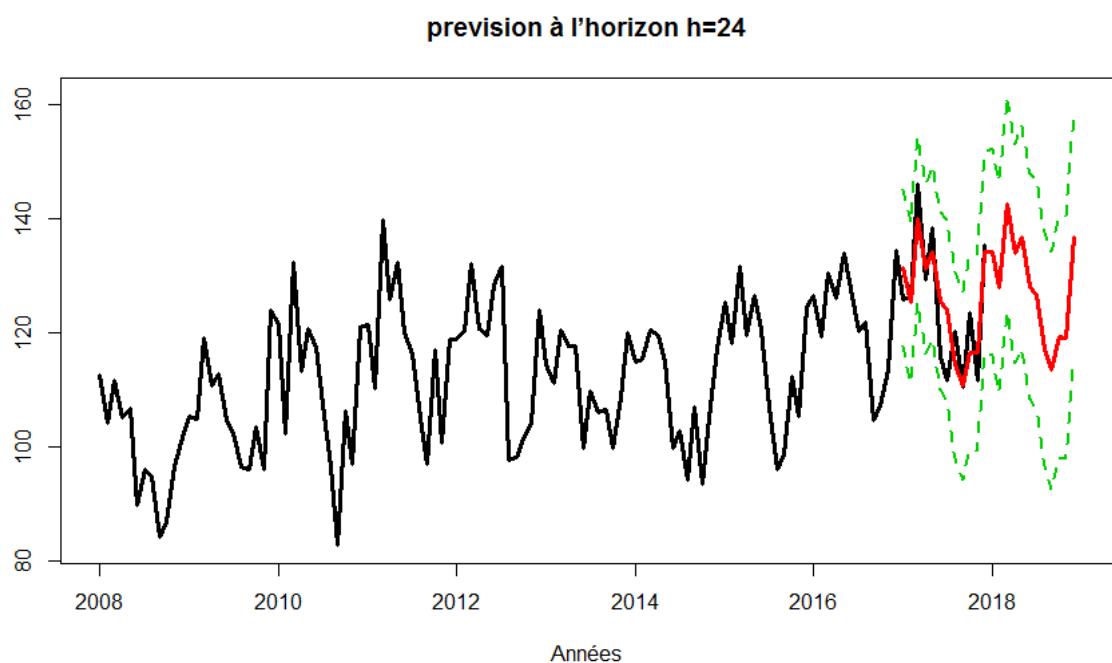


Figure 13 - Prévisions de l'année 2017 et 2018

Chapitre 3 : Modélisation multivariée de l'indice de Production Industrielle du Sénégal

L'objectif de la modélisation multivariée est d'expliquer l'évolution de l'IPI à travers les relations de causalité avec d'autres variables. La revue de littérature a permis de retenir une variable susceptible de déterminer l'indice de production industrielle du Sénégal : l'Indice de Production du Consommateur (IPC). Toute au long de cette section, on ne considérera que les valeurs de la série IPC allant de janvier 2008 à décembre 2017.

3.1 Analyse descriptive de l'IPC

Il ressort principalement qu'en moyenne le Sénégal a un IPC de 102,07 entre janvier 2008 et décembre 2016. Ainsi, on remarque que la série a une évolution comprise entre 95,64 et 107,49.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
95,64	100,87	102,31	102,07	103,91	107,49

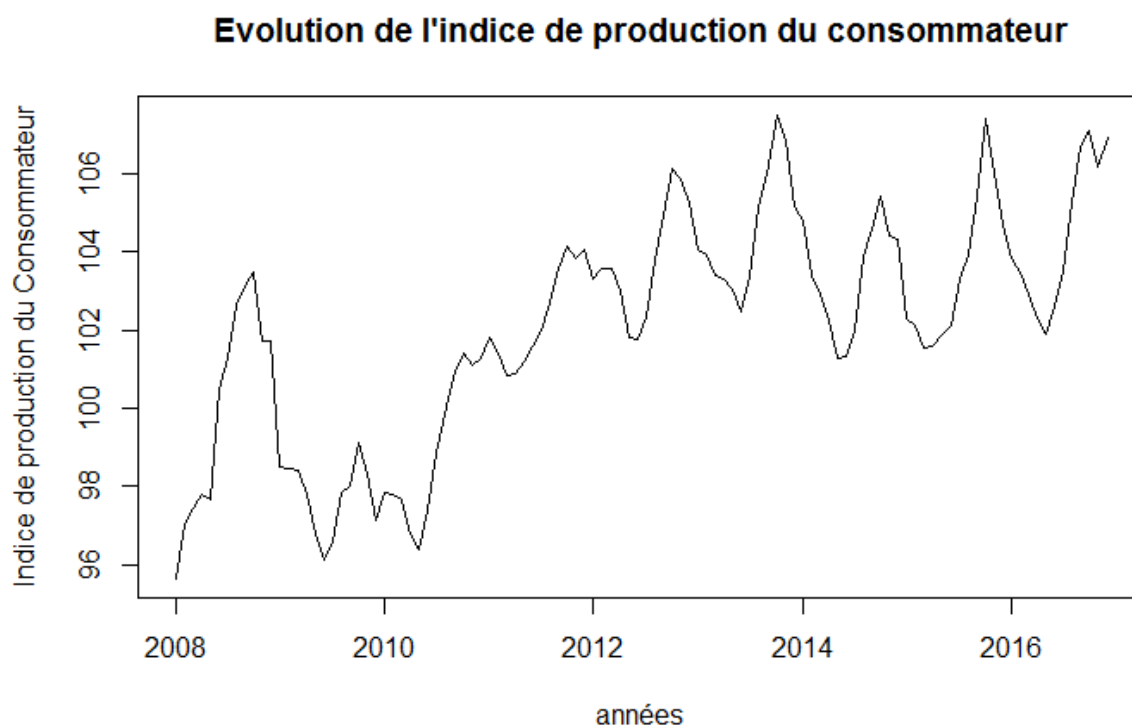


Figure 14 - Visualisation de la série IPC

Le graphique nous montre une évolution en dents de scie tout au long de la trajectoire. En plus, on remarque que notre série présente une forte dispersion avec un écartype de 2,920283.

3.2 Tests de stationnarité de variable IPC et ordres d'intégration

Les variabilités observées sur la série IPC nous préconisent d'appliquer une transformation logarithmique pour stabiliser la variance et centrée notre série. De plus ce procédé a pour avantage de faciliter l'interprétation des coefficients puisqu'ils seront dès lors assimilables à des élasticités. Les séries économiques sont généralement affectées d'une tendance de long terme qui peut être déterministe ou stochastique. Il est donc primordial avant toute modélisation multivariée d'identifier la nature de la stationnarité pour pouvoir en prendre compte dans les prédictions.

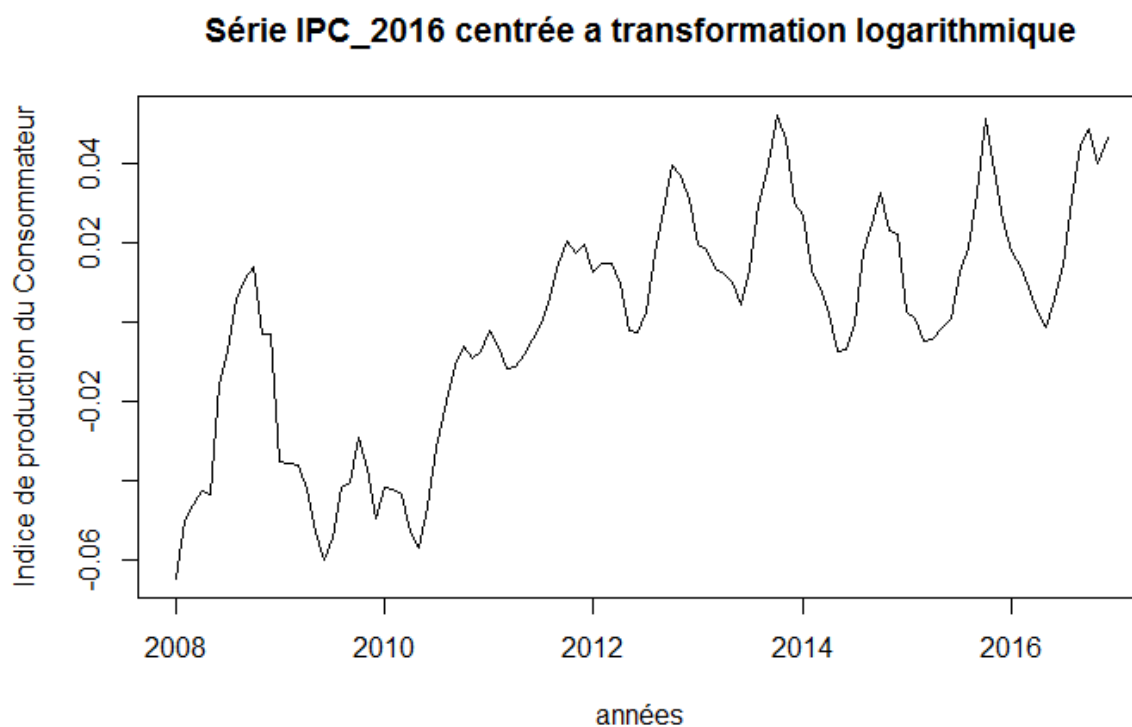


Figure 15 - Visualisation de la série \log (IPC) centrée

Une analyse minutieuse du graphique de la série permet à priori de postuler sur la présence d'une tendance. On remarque en effet une certaine tendance croissante.

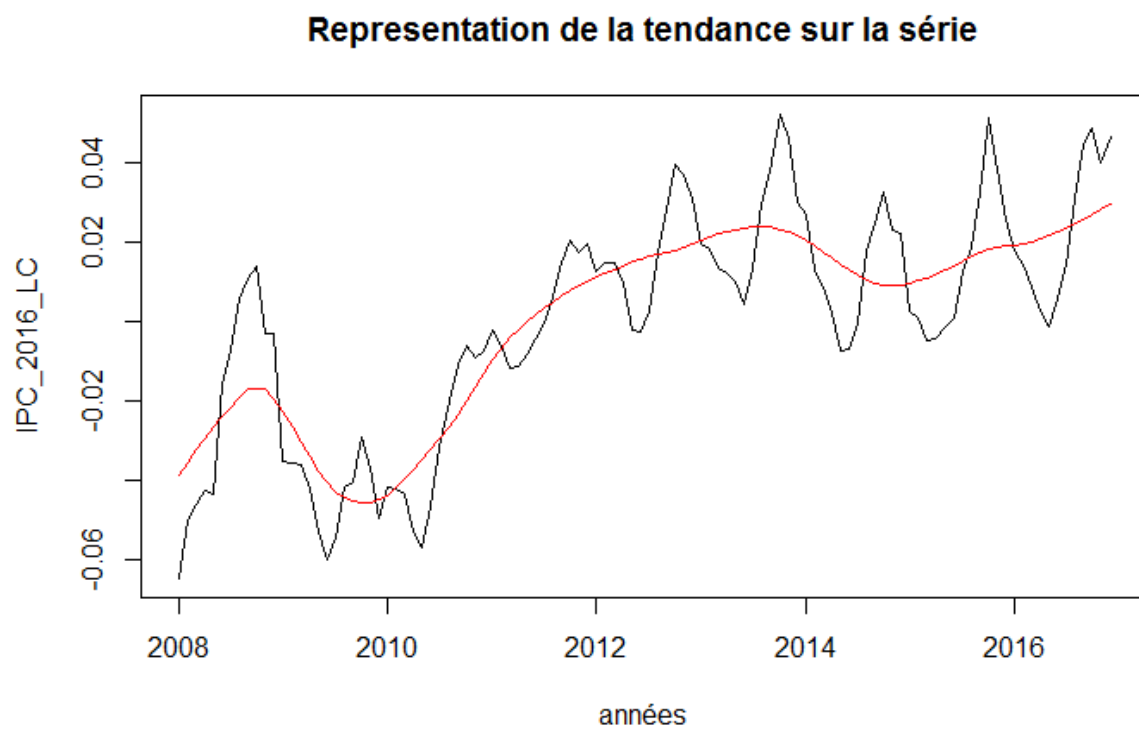


Figure 16 - Visualisation de la tendance du log (IPC) centrée

L'ACF de la série présente une décroissance lente vers 0, ce qui traduit un problème de non-stationnarité.

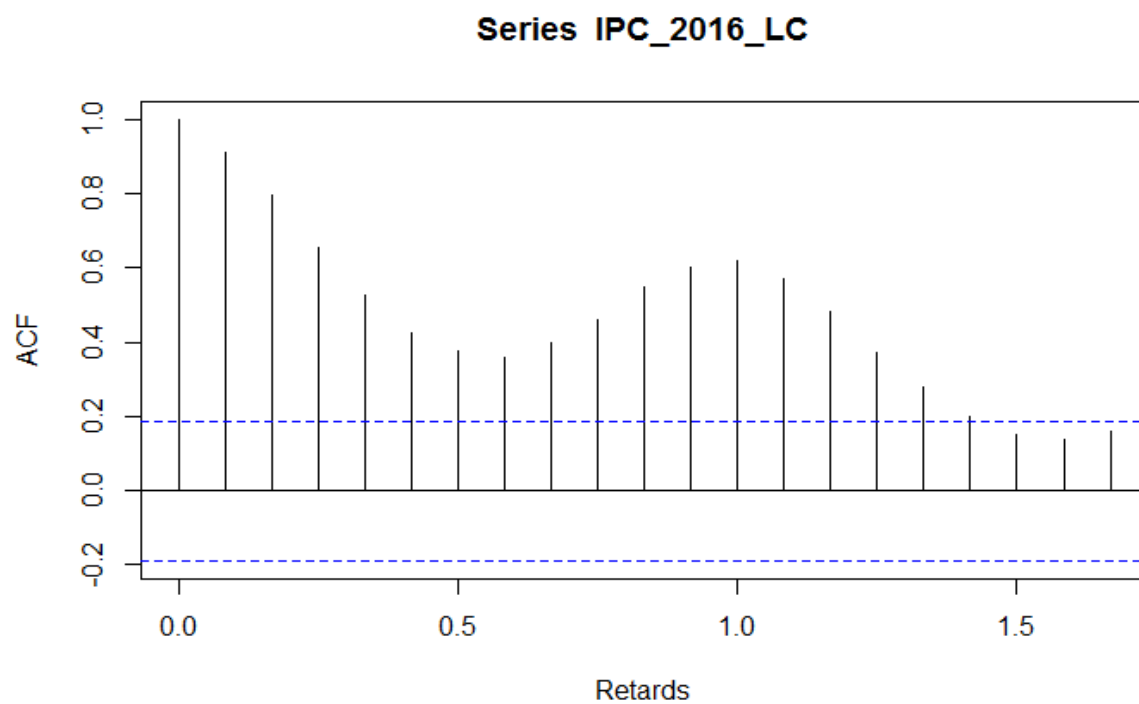


Figure 17- fonction d'autocorrélation de la série log(IPC) centrée

Pour avoir une idée plus précise de la présence ou de l'absence de stationnarité, les tests de Dickey-Fuller augmenté, Phillips Perron et KPSS seront effectués sur la série.

- **Test augmenté de Dickey Fuller (ADF)** : Le test de Dickey-Fuller augmenté suppose au départ que la série soit non stationnaire. Les résultats de ce test nous donnent une p-value égale à 0,0286. On rejette alors l'hypothèse H_0 . On peut donc affirmer au seuil de 5% que la série est stationnaire.
- **Test de Phillips Perron** : Tout comme le test de Dickey-Fuller augmenté, le test de Phillips Perron admet pour hypothèse nulle la présence d'une tendance dans la série. Les résultats nous donne une p-value de $0,05603 > 0,05$. On ne peut donc pas rejeter la non stationnarité.
- **Test KPSS** : Contrairement aux deux tests précédents, le test KPSS admet comme hypothèse nulle que la série soit stationnaire. Ici la p-value de 0.01 traduit donc la non stationnarité de la série.

Les deux derniers tests ci-dessus nous montrent que la série est non stationnaire. Cependant, le test augmenté de Dickey Fuller postule en faveur d'une stationnarité de la série. Pour uniformiser tous les tests, on effectue donc une différenciation de type $(I-B)$. La reprise de ses trois tests nous donne respectivement des p-values de 0,01, 0,01 et 0,1. Nous pouvons donc conclure que la série est stationnaire. Ainsi, la moyenne et la variance de la série ne dépendent pas du temps. Le graphique nous permet de confirmer visuellement cette stationnarité.

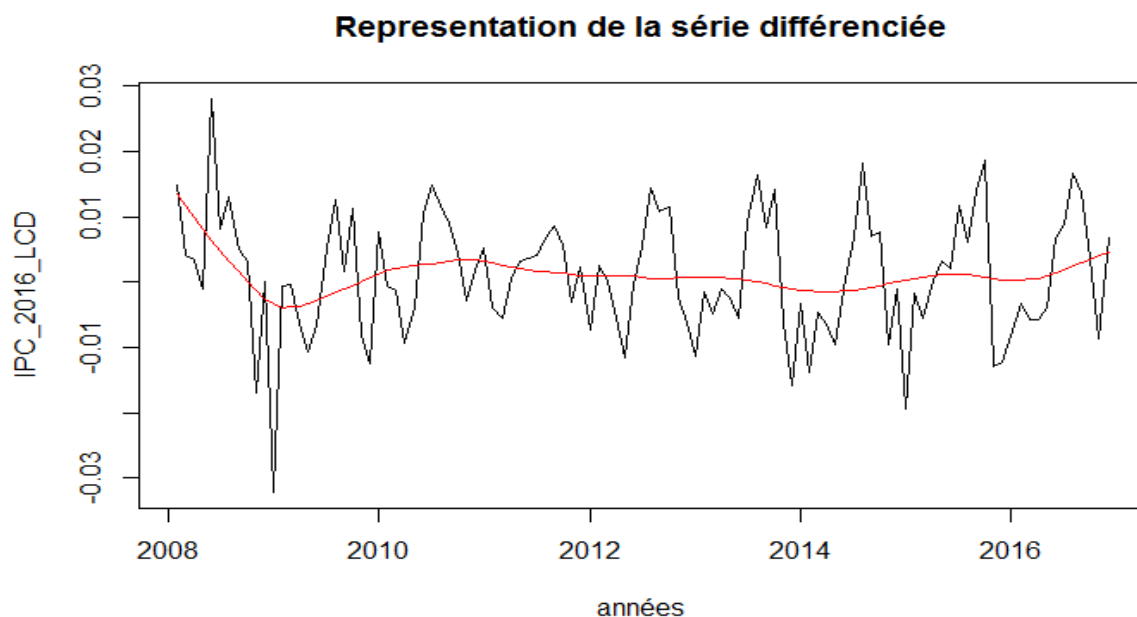


Figure 18 - Visualisation de la tendance de la série log (IPC) centrée et différenciée

3.3 Etude multivariée

L'observation de nos deux séries simultanée nous donne :

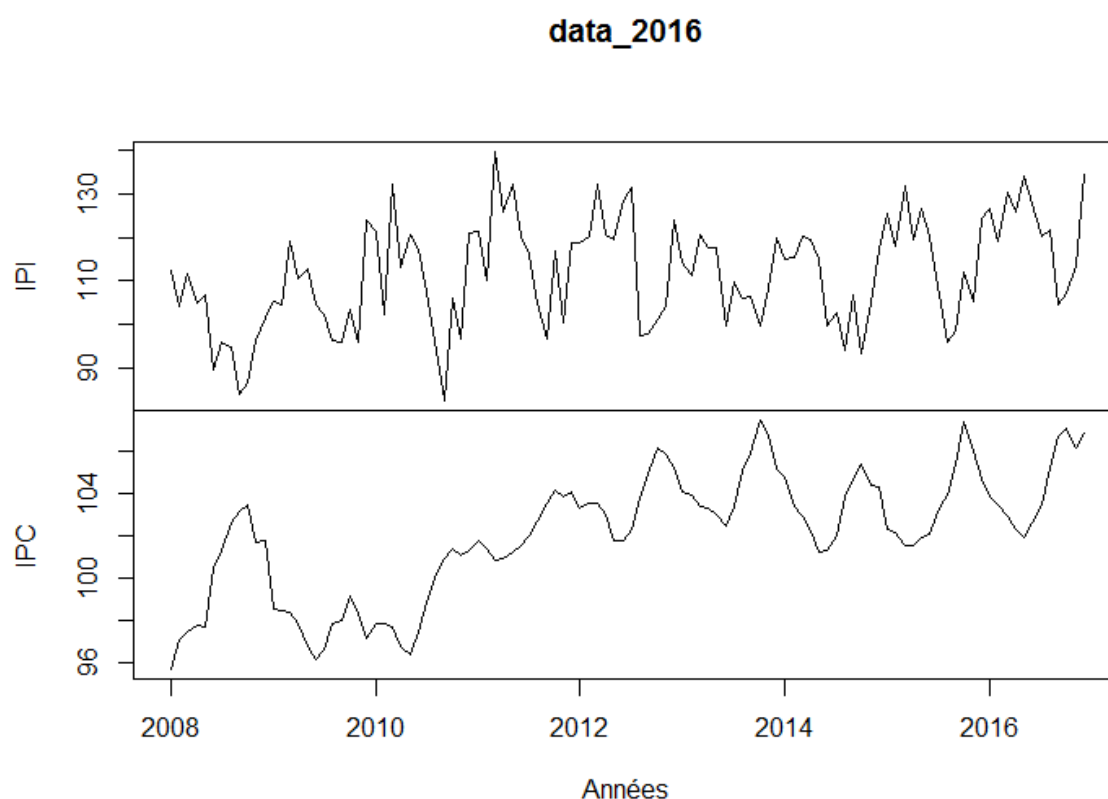


Figure 19 - Observation simultanée de IPI et IPC

Dans la suite de notre étude, nous allons considérer uniquement les séries stationnaires.

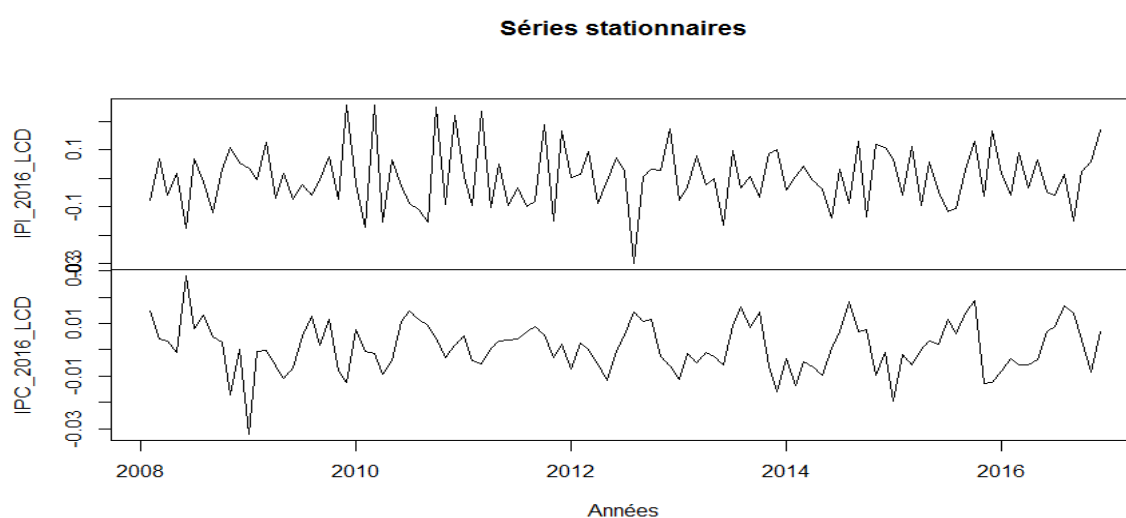


Figure 20 - Figure 19 - Observation simultanée de $\log(IPI)$ centrée différenciée et $\log(IPC)$ centrée différenciée

3.3.1 Identification du nombre de retards du modèle

L'estimation du paramètre p du modèle VAR est obtenue grâce à la fonction VARselect(). Le nombre de retards optimaux d'après chaque critère d'information sont consignées dans le tableau 4 :

Tableau 4 - Retards optimaux selon chaque critère d'information

AIC	HQ	SC	FPE
9	8	3	9

D'après le critère d'AIC et de FPE, le nombre de retard optimal est $p=9$ tandis que $p=3$ selon le critère de SC et 8 selon HQ. Nous allons donc retenir le critère de SC en vertu du principe de parcimonie qui exige à ce qu'on retienne le minimum des minima.

Tableau 5 - liste des différentes valeurs de chaque critère en fonction des retards

Retards	AIC	HQ	SC	FPE
1	-1,422309E+01	-1,415870E+01	-1,406383E+01	6,652832E-07
2	-1,433084E+01	-1,422351E+01	-1,406541E+01	5,974122E-07
3	-1,447234E+01	-1,432208E+01	-1,410073E+01	5,187538E-07
4	-1,445626E+01	-1,426306E+01	-1,397847E+01	5,274643E-07
5	-1,450992E+01	-1,427380E+01	-1,392597E+01	5,003484E-07
6	-1,451468E+01	-1,423563E+01	-1,382455E+01	4,986125E-07
7	-1,459410E+01	-1,427211E+01	-1,379779E+01	4,613567E-07
8	-1,472213E+01	-1,435721E+01	-1,381965E+01	4,068521E-07
9	-1,475495E+01	-1,434710E+01	-1,374630E+01	3,948763E-07
10	-1,474375E+01	-1,429297E+01	-1,362892E+01	4,007913E-07

3.3.2 Estimations des paramètres et validation du modèle multivarié

Comme dans le cas univarié, la validation du modèle passe par la vérification de la significativité des paramètres et de l'hypothèse de bruits blancs des résidus. L'estimation des paramètres de notre processus VAR se fait par la fonction VAR().

a) Significativité des paramètres

La significativité des paramètres est vérifiée en faisant appel à la fonction summary(). Les résultats sont affichés ci-dessous :

```

VAR Estimation Results:
=====
Endogenous variables: IPI_2016_LCD, IPC_2016_LCD
Deterministic variables: const
Sample size: 104
Log Likelihood: 465.122
Roots of the characteristic polynomial:
0.7615 0.7615 0.576 0.576 0.5228 0.5228
Call:
VAR(y = varmat, p = 3)

Estimation results for equation IPI_2016_LCD:
=====
IPI_2016_LCD = IPI_2016_LCD.l1 + IPC_2016_LCD.l1 + IPI_2016_LCD.l2 + IPC_2016_LCD.l2 + IPI_2016_LCD.l3 + IPC_2016_LCD.l3 + const

      Estimate Std. Error t value Pr(>|t|)
IPI_2016_LCD.l1 -0.506808   0.104830  -4.835 5.01e-06 ***
IPC_2016_LCD.l1 -3.834736   1.194179  -3.211 0.001794 **
IPI_2016_LCD.l2 -0.077320   0.120542  -0.641 0.522748
IPC_2016_LCD.l2  0.167743   1.158578   0.145 0.885182
IPI_2016_LCD.l3  0.044884   0.109566   0.410 0.682967
IPC_2016_LCD.l3  3.715387   1.070559   3.471 0.000777 ***
const           0.001607   0.009140   0.176 0.860838
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.09195 on 97 degrees of freedom
Multiple R-squared: 0.2903,    Adjusted R-squared: 0.2464
F-statistic: 6.612 on 6 and 97 DF,  p-value: 6.964e-06

```

Tout d'abord, les résultats de l'estimation des paramètres de l'équation de l'IPI_2016_LCD nous donnent un R^2 ajusté de 0,2464. Ensuite, on remarque que seuls les coefficients IPI_2016_LCD.l1, IPC_2016_LCD.l1 et IPC_2016_LCD.l3 sont significatifs. Enfin, IPI_2016_LCD.l1, IPC_2016_LCD.l1 ont un impact négatif sur IPI_2016_LCD et seul l'IPC_2016_LCD.l3 a un impact positif. Enfin tous ces trois variables ont un impact négatif sur IPC_2016_LCD.

Estimation results for equation IPC_2016_LCD:

IPC_2016_LCD = IPI_2016_LCD.l1 + IPC_2016_LCD.l1 + IPI_2016_LCD.l2 + IPC_2016_LCD.l2 + IPI_2016_LCD.l3 + IPC_2016_LCD.l3 + const

	Estimate	Std. Error	t value	Pr(> t)
IPI_2016_LCD.l1	-0.0240603	0.0093172	-2.582	0.011307 *
IPC_2016_LCD.l1	0.1819715	0.1061373	1.714	0.089632 .
IPI_2016_LCD.l2	-0.0389985	0.0107136	-3.640	0.000439 ***
IPC_2016_LCD.l2	0.0677578	0.1029732	0.658	0.512088
IPI_2016_LCD.l3	-0.0170728	0.0097381	-1.753	0.082727 .
IPC_2016_LCD.l3	-0.2415961	0.0951501	-2.539	0.012704 *
const	0.0008997	0.0008124	1.108	0.270799

signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.008172 on 97 degrees of freedom
Multiple R-Squared: 0.2961, Adjusted R-squared: 0.2525
F-statistic: 6.8 on 6 and 97 DF, p-value: 4.853e-06

Covariance matrix of residuals:

	IPI_2016_LCD	IPC_2016_LCD
IPI_2016_LCD	0.0084539	-2.249e-04
IPC_2016_LCD	-0.0002249	6.678e-05

Correlation matrix of residuals:

	IPI_2016_LCD	IPC_2016_LCD
IPI_2016_LCD	1.0000	-0.2993
IPC_2016_LCD	-0.2993	1.0000

En ce qui concerne les résultats de l'estimation des paramètres de l'équation de l'IPC_2016_LCD, nous avons un R^2 ajusté de 0,2525. Ensuite, on remarque que seuls les coefficients de l'IPI_2016_LCD.l1, l'IPI_2016_LCD.l2 et l'IPC_2016_LCD.l3 sont significatifs. La figure suivante nous représente c'est deux séries avec leur résidus.

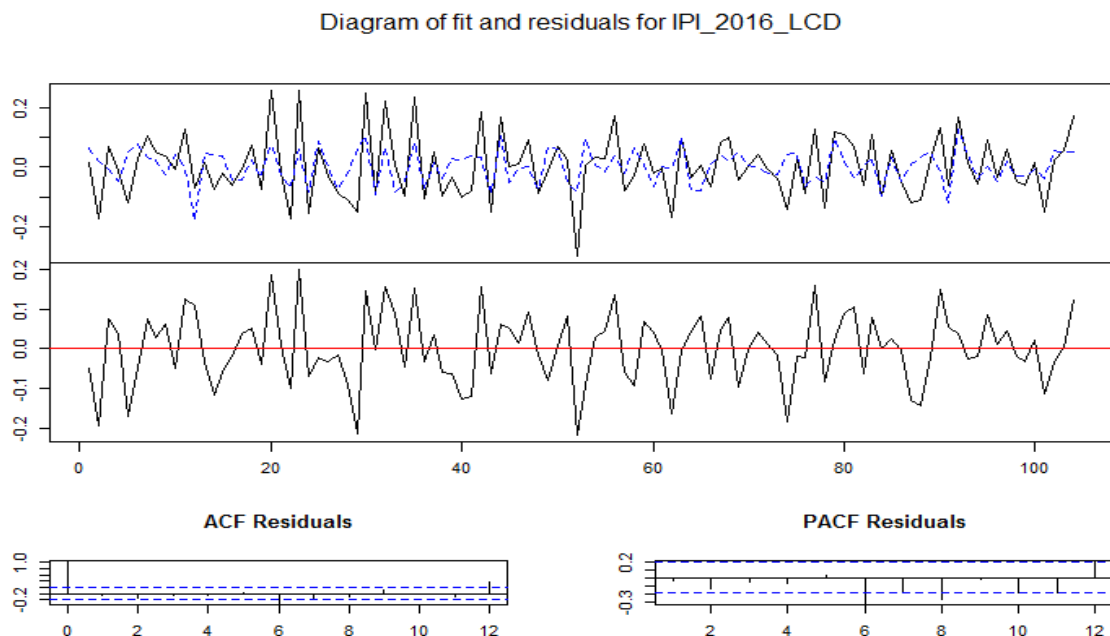


Figure 21 - Représentation de la série $\log(IPI)$ centrée différenciée avec ses résidus

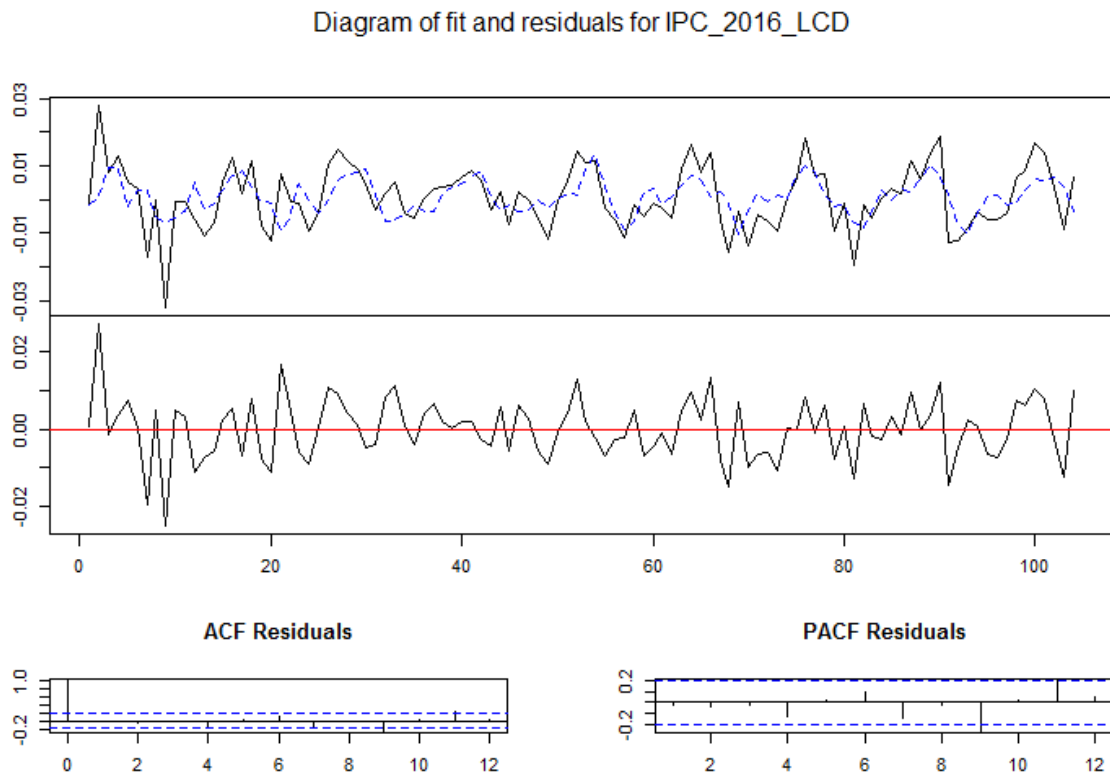


Figure 22 - Représentation de la série $\log(\text{IPC})$ centrée différenciée avec ses résidus

b) Stabilité

Un processus VAR (p) est stable, si son polynôme caractéristique inverse n'a pas de racines dans ou sur le cercle complexe. Cela équivaut à la condition que toutes les valeurs propres de la matrice compagnon A aient un module inférieur à 1. La fonction `roots()` calcule les valeurs propres de la matrice compagnon A et renvoie par défaut leurs modules. Les valeurs obtenues sont: 0,7614602, 0,7614602, 0,5760118, 0,5760118, 0,5228368, 0,5228368. Pour une représentation visuelle du test de stabilité nous allons utiliser la fonction `stability()`.

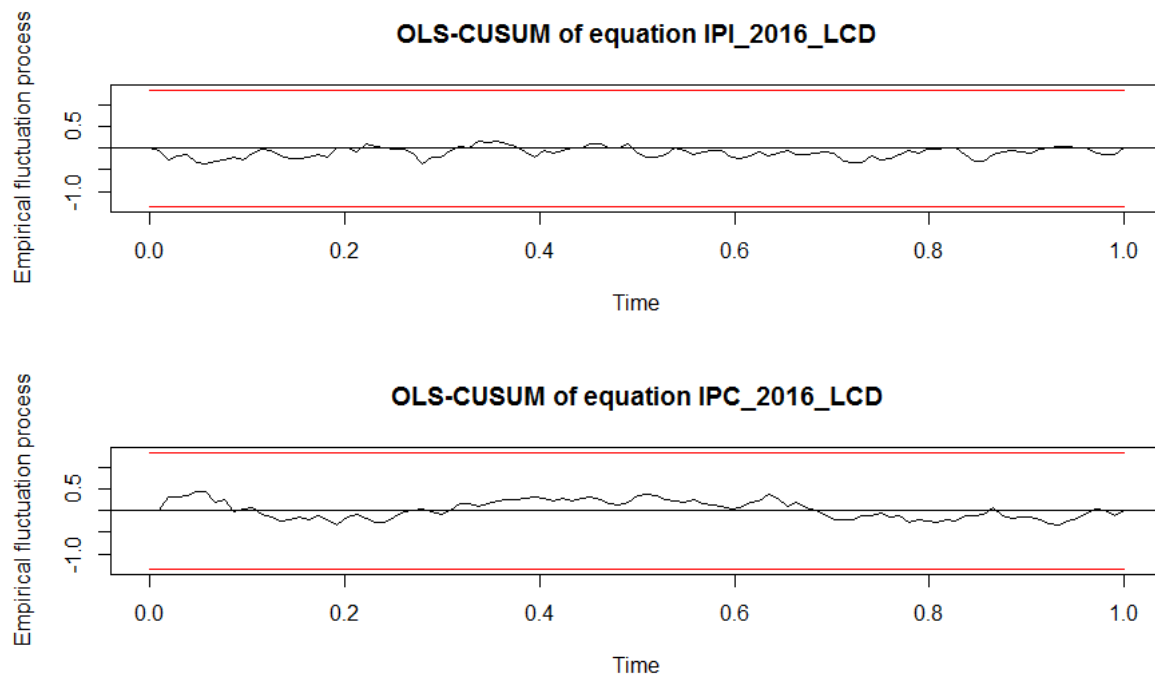


Figure 23 - Stabilité du modèle

Nous observons donc que les fluctuations sont entre les bornes de l'intervalle de confiance. Nous pouvons donc conclure à la stabilité du modèle VAR(3).

c) Analyse des résidus

- **Normalité** : La représentation graphique de l'histogramme des résidus et de la courbe des quantiles théoriques nous pousse à penser que les résidus sont normaux.

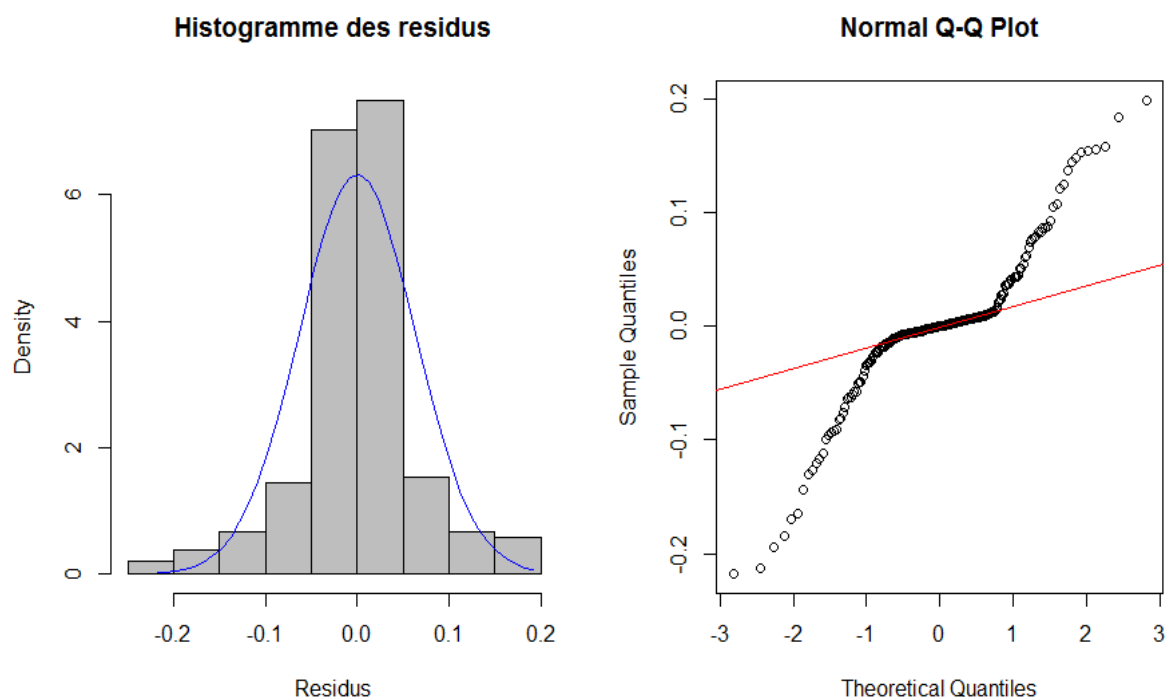


Figure 24 - Visualisation des résidus

Pour vérifier cette normalité, nous avons effectué le test de Jarque-Bera. Les résultats de cet test nous donne une $p\text{-value} = 0,7737$. Nous pouvons donc conclure au seuil de 5% que les résidus de la modélisation suivent une loi normale.

- **Autocorrélation des résidus:** elle est également testée ici par le test de Portemanteau. Nous avons utilisé la fonction `serial.test()` en donnant 30 à l'argument `lags.pt`. On obtient alors une $p\text{-value} = 0,1002 > 0.05$. L'hypothèse H_0 ne peut être rejeter. Il n'y a donc pas d'autocorrélation des erreurs.
- **Homoscédasticité des résidus:** le test d'Arch est utilisé ici pour vérifier l'hypothèse d'homoscédasticité des résidus. On obtient donc une $p\text{-value} = 0,7386 > 0.05$. L'hypothèse H_0 ne peut être rejeter. Il n'y a donc pas d'homoscédasticité des résidus.

d) Causalité entre variables

On dit qu'une variable X cause au sens de Granger la variable Y si et seulement si la connaissance du passé de X améliore la prévision de Y en tout horizon h . L'analyse de la causalité permet de connaître les influences statistiquement significatives des variables du modèle entre elles. Pour étudier le lien causal entre les variables du modèle, nous avons utilisé

le test de causalité de GRANGER à l'aide de la fonction `causality()`. Ce test suppose au départ que la série IPC ne cause pas la série IPI. Les résultats de ce test nous donnent une p-value égale à 0.0001289. On rejette alors l'hypothèse H_0 . On peut donc affirmer au seuil de 5% que la série IPC cause la série IPI au sens de GRANGER.

e) Les fonctions de demande impulsionnelles

Les fonctions impulsives permettent de ranger les variables de la moins dépendante à la plus dépendante. Cependant, dans notre cas on a une seule variable explicative.

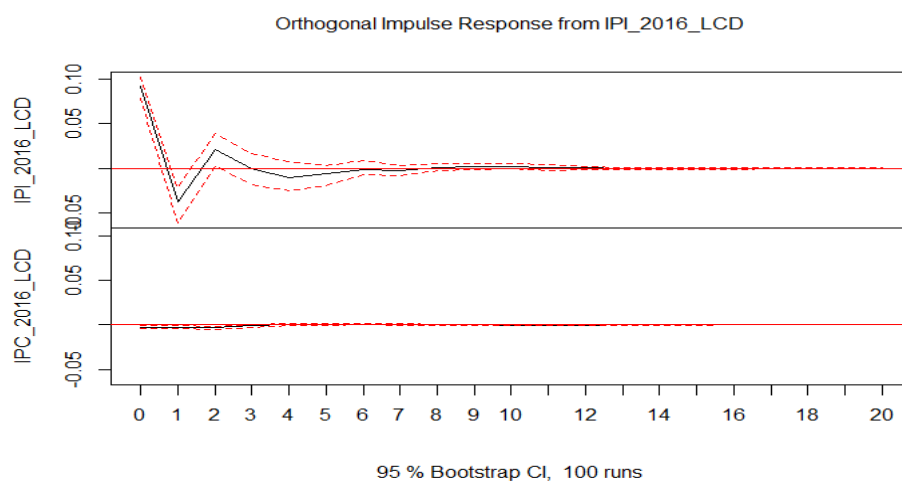


Figure 25 - Fonction d'impulsion

f) Décomposition de la variance

La décomposition de la variance de l'erreur de prévision a pour objectif de calculer pour chacune des innovations sa contribution à la variance de l'erreur. Les résultats relatifs à l'étude de la décomposition de la variance obtenus à l'aide de la commande `fevd()`.

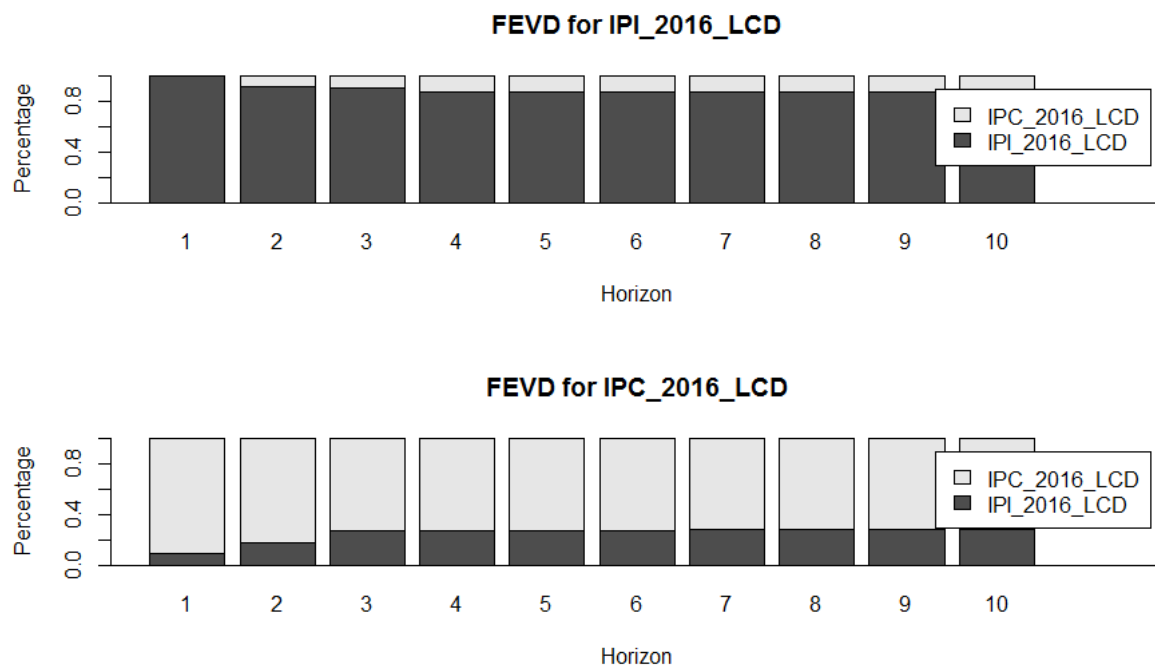


Figure 26 - Décomposition de la variance des erreurs des prévisions

e) Prévisions

Les prévisions du modèle VAR sont obtenues à l'aide de la fonction predict().

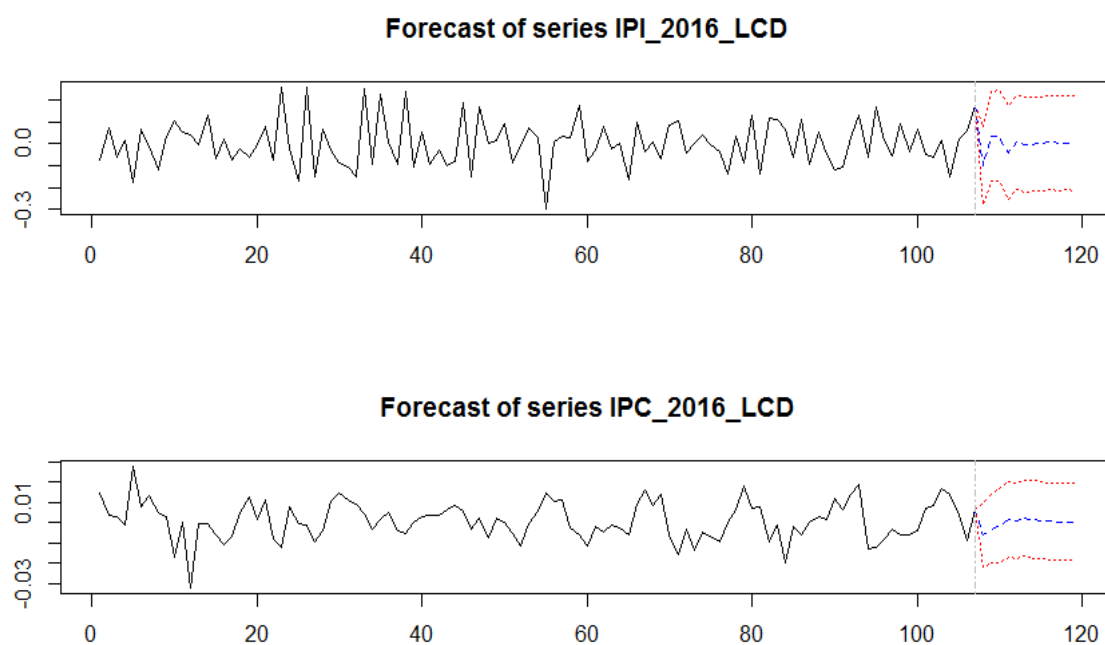


Figure 27 - Prévision pour 12 mois

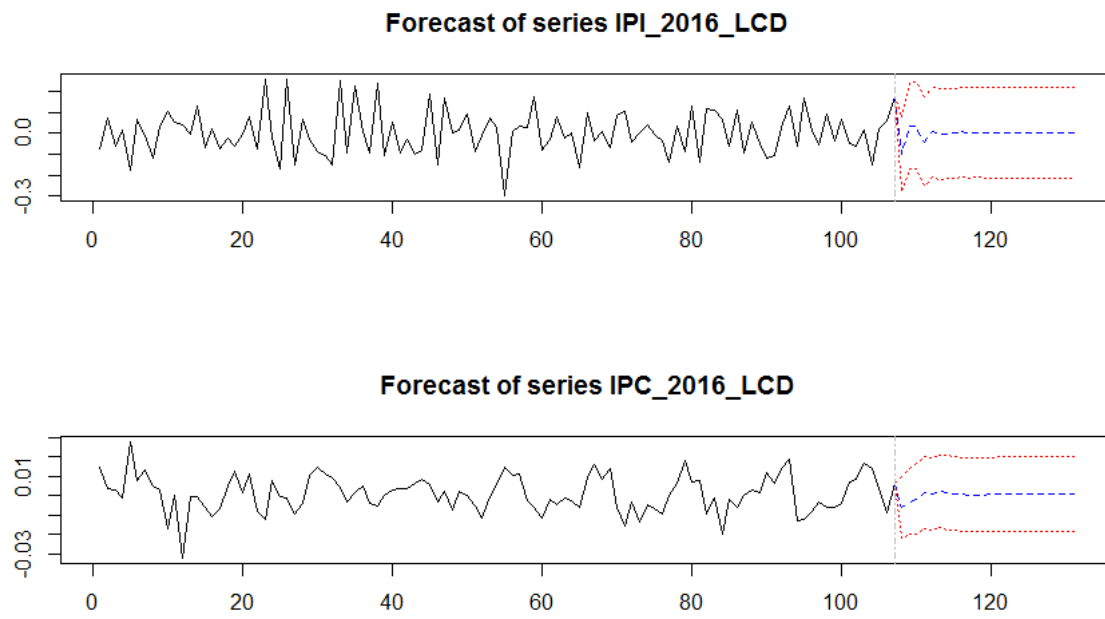


Figure 28 - Pr vision pour 24 mois

Conclusion

En somme, ce travail nous a permis d'appliquer les notions de modélisation vues en cours de séries temporelles. Les deux bases de données utilisées sont la base d'Indice de Production Industrielle (IPI) et la base d'Indice des Prix à la Consommation (IPC). En premier lieu, nous avons eu à faire une modélisation univariée en suivant la méthodologie de Box et Jenkins avec la variable IPI du Sénégal entre 2008 et 2016. Cette étape commence par l'analyse préliminaire, en passant par la détermination des ordres, l'identification des modèles, l'estimation des paramètres, l'analyse des résidus, le choix du modèle jusqu'à la prévision. La seconde partie de notre travail a été consacrée à l'analyse multivariée où une autre variable a été introduite : la variable IPC. Dans cette partie nous avons eu à commencer par une analyse préliminaire de la variable IPC. Ensuite, nous avons identifié le nombre de retards optimaux de notre modèle. Enfin, l'estimation des paramètres et la validation de notre modèle a été effectué. Cette étude m'a permis d'apprendre de nouveaux concepts (causalité et analyse de choc), de comprendre les objectifs de l'étude d'une série chronologique et de mieux maîtriser les modèles SARIMA et les modèles VAR. Ce travail m'a permis aussi d'assembler mes connaissances et d'améliorer mes compétences sur l'utilisation du logiciel R.

Bibliographie

- [1] Dr Souleymane Fofana, cours : Séries Temporelles Théorie et Pratique, Université Alioune DIOP de Bambey, 2017-2018
- [2] Dr Souleymane Fofana, cours : Econométrie des Séries Temporelles : Modélisation VAR, Université Alioune DIOP de Bambey, 2017-2018
- [3] Hamissou ALAJI BOUHARI, Aurelie MANAA GONTA, Mohamed NIANG, Abdoul Razack YONABA, Déterminants des importations de riz au Sénégal, ENSAE Dakar, juillet 2018
- [4] Yves Aragon, Séries temporelles avec R: Méthodes et cas, Springer-Verlag France, 2011
- [5] Robert H. Shumway, David S. Stoffer, Time Series Analysis and Its Applications With R Examples, Fourth Edition, Springer International Publishing AG 1999, 2012, 2016, 2017
- [6] Chafiaà Ayhar, mémoire de master : la modélisation VAR, Université de Saida - Dr Moulay Tahar, 2017-2018

Codes R utilisés

#####

Code Projet Modelisation de l'indice de production au senegal

#####

Importation des library nécessaire a l'étude

library(tseries)

library(forecast)

library(lmtest)

library(vars)

library("astsa")

library(urca)

library(forecast)

#####

Analyse descriptive

#####

setwd("C:/Users/Sid Ousmane/Documents/texte/MASTER_SID_UADB/M1SID/Semestre 2/Séries
Temporelles/Master_SID_Séries_Temp")

Importation de la base de donnée

IPI<- read.csv("IPI.csv", sep=";", header=T)

Transformation de la base en type série temporelle

IPI <- ts(IPI[,2], start=c(2008,1), end=c(2017,12), frequency=12)

IPI

Restriction de la serie à l'année 2016

IPI_2016 <- window(IPI, end=c(2016,12))

IPI_2016

```
## Description de la base de donnée
```

```
summary(IPI_2016)
```

```
sd(IPI_2016)
```

```
## Representation de la série
```

```
ts.plot(IPI_2016, xlab ="année", ylab ="Indice de production Industrielle")
```

```
## Calcul de la Variance
```

```
var(IPI_2016)
```

```
## Reduction de la dispersion de la série en la stabilisant la variance
```

```
IPI_2016_L <- log(IPI_2016)
```

```
ts.plot(IPI_2016_L,main ="Série IPI_2016", xlab ="année", ylab ="Indice de production Industrielle")
```

```
## Centrage de la serie transformé logarithmiquement
```

```
IPI_2016_LC <- IPI_2016_L -mean(IPI_2016_L)
```

```
ts.plot(IPI_2016_LC,main ="Série IPI_2016 centré a transformation logarithmique", xlab ="année",  
ylab ="Indice de production Industrielle")
```

```
acf(IPI_2016_LC, xlab= "Retards")
```

```
##### TEST DE STATIONNARITE
```

```
# affichage de la tendance de la série
```

```
tend =stl(IPI_2016_LC, "per")
```

```
ts.plot(IPI_2016_LC, main ="Representation de la tendance sur la série", xlab ="années")
```

```
lines(tend$time.series[,2], col = "red")
```

```
# Affichage de l'aCF de la serie
```



```
acf(IPI_2016_LC, xlab= "Rétards")
```

```
# Test augmenter de Dickey-Fuller
```

```
adf.test(IPI_2016_LC)
```

```
# Test de Phillips Perron
```

```
PP.test(IPI_2016_LC)
```

```
# Test KPSS
```

```
kpss.test(IPI_2016_LC)
```

```
### Differenciation de la série et reprise des tests de stationnaires
```

```
# Differenciation de la série
```

```
IPI_2016_LCD <- diff(IPI_2016_LC)
```

```
# Test augmenter de Dickey-Fuller
```

```
adf.test(IPI_2016_LCD)
```

```
# Test de Phillips Perron
```

```
PP.test(IPI_2016_LCD)
```

```
# Test KPSS
```

```
kpss.test(IPI_2016_LCD)
```

```
# confirmation visuelle
```

```
tendD =stl(IPI_2016_LCD, "per")
```

```
ts.plot(IPI_2016_LCD, main ="Representation de la série différenciée", xlab ="années")
```

```
lines(tendD$time.series[,2], col = "red")
```

```
# Saisonnalité
```

```
acf(IPI_2016_LCD, lag.max=36, xlab= "Retards")
```

```
IPI_2016_LCDS <- diff(IPI_2016_LCD, lag = 12)
```

```
ts.plot(IPI_2016_LCDS, xlab ="année", main ="Serie desaisonnalisée")
```

```
##### Identification du modèle
```

```
acf(IPI_2016_LCDS, xlab="Retards", lag.max =36)
```

```
pacf(IPI_2016_LCDS, xlab ="Retards", lag.max = 36)
```

```
# Estimation des paramètres
```

```
model1 = arima(IPI_2016, order = c(0,1,1), seasonal=list(order=c(0,1,1),period=12))
```

```
model2 = arima(IPI_2016, order = c(1,1,1), seasonal=list(order=c(0,1,1),period=12))
```

```
model3 = arima(IPI_2016, order = c(2,1,1), seasonal=list(order=c(0,1,1),period=12))
```

```
model4 = arima(IPI_2016, order = c(1,1,0), seasonal=list(order=c(0,1,1),period=12))
```

```
model5 = arima(IPI_2016, order = c(2,1,0), seasonal=list(order=c(0,1,1),period=12))
```

```
model6 = arima(IPI_2016, order = c(0,1,1), seasonal=list(order=c(1,1,1),period=12))
```

```
model7 = arima(IPI_2016, order = c(1,1,1), seasonal=list(order=c(1,1,1),period=12))
```

```
model8 = arima(IPI_2016, order = c(2,1,1), seasonal=list(order=c(1,1,1),period=12))
```

```
model9 = arima(IPI_2016, order = c(1,1,0), seasonal=list(order=c(1,1,1),period=12))
```

```
model10 = arima(IPI_2016, order = c(2,1,0), seasonal=list(order=c(1,1,1),period=12))
```

```
model11 = arima(IPI_2016, order = c(0,1,1), seasonal=list(order=c(2,1,1),period=12))
```

```
model12 = arima(IPI_2016, order = c(1,1,1), seasonal=list(order=c(2,1,1),period=12))
```

```
model13 = arima(IPI_2016, order = c(2,1,1), seasonal=list(order=c(2,1,1),period=12))
```

```
model14 = arima(IPI_2016, order = c(1,1,0), seasonal=list(order=c(2,1,1),period=12))
```

```
model15 = arima(IPI_2016, order = c(2,1,0), seasonal=list(order=c(2,1,1),period=12))
```

```
model16 = arima(IPI_2016, order = c(0,1,1), seasonal=list(order=c(1,1,0),period=12))  
model17 = arima(IPI_2016, order = c(1,1,1), seasonal=list(order=c(1,1,0),period=12))  
model18 = arima(IPI_2016, order = c(2,1,1), seasonal=list(order=c(1,1,0),period=12))  
model19 = arima(IPI_2016, order = c(1,1,0), seasonal=list(order=c(1,1,0),period=12))  
model20 = arima(IPI_2016, order = c(2,1,0), seasonal=list(order=c(1,1,0),period=12))
```

```
model21 = arima(IPI_2016, order = c(0,1,1), seasonal=list(order=c(2,1,0),period=12))  
model22 = arima(IPI_2016, order = c(1,1,1), seasonal=list(order=c(2,1,0),period=12))  
model23 = arima(IPI_2016, order = c(2,1,1), seasonal=list(order=c(2,1,0),period=12))  
model24 = arima(IPI_2016, order = c(1,1,0), seasonal=list(order=c(2,1,0),period=12))  
model25 = arima(IPI_2016, order = c(2,1,0), seasonal=list(order=c(2,1,0),period=12))
```

#Significativité des paramètres des modèles estimées

```
coeftest(model1)  
coeftest(model2)  
coeftest(model3)  
coeftest(model4)  
coeftest(model5)  
coeftest(model6)  
coeftest(model7)  
coeftest(model8)  
coeftest(model9)  
coeftest(model10)  
coeftest(model11)  
coeftest(model12)  
coeftest(model13)  
coeftest(model14)  
coeftest(model15)
```

```
coeftest(model16)
coeftest(model17)
coeftest(model18)
coeftest(model19)
coeftest(model20)
coeftest(model21)
coeftest(model22)
coeftest(model23)
coeftest(model24)
coeftest(model25)
```

```
# Critère de selection selon l'AIC
```

```
model1$aic
model5$aic
model16$aic
model19$aic
model20$aic
model21$aic
model24$aic
model25$aic
```

```
##### Analyse des résidus
```

```
## Représentation graphique des résidus
```

```
par(mfrow=c(1,2))
hist(model1$residuals, col="grey", freq=F, xlab="Résidus", main = "Histogramme des residus")
qqnorm(model1$residuals)
qqline(model1$residuals, col="red")
```

```
# Test du t de student
```

```
t.test(model1$residuals)
```

```
#Fonction d'autocorrélation et Statistique de Ljung-Box des résidus
```

```
tsdiag(model1)
```

```
# Test de normalité des résidus
```

```
jarque.bera.test(model1$residuals)
```

```
# Homoscédasticité des résidus
```

```
white.test(model1$residuals)
```

```
# Test d'autocorrélation
```

```
Box.test(model1$residuals, lag = 1, type = c("Ljung-Box"))
```

```
#####
```

```
##### Prévison #####
```

```
#####
```

```
# prevision pour 12mois
```

```
h <- 12
```

```
arima.prev12<-predict(model1, n.ahead=h)
```

```
arima.prev12
```

```
arima.inf12<-arima.prev12$pred-1.96*arima.prev12$se
```

```
arima.inf12
```

```
arima.sup12<-arima.prev12$pred+1.96*arima.prev12$se
```

```
arima.sup12
```

```
ts.plot(IPI, arima.prev12$pred, arima.inf12, arima.sup12, main="prevision à l'horizon h=12",  
col=c(1,2,3,3),lty=c(1,1,2,2),lwd=c(3,3,2,2), xlab = "Années")
```

```
# prévisions pour 24 mois
```

```

h <- 24
arima.prev24<-predict(model1, n.ahead=h)
arima.prev24
arima.inf24<-arima.prev24$pred-1.96*arima.prev24$se
arima.inf24
arima.sup24 <-arima.prev24$pred+1.96*arima.prev24$se
arima.sup24
ts.plot(IPI, arima.prev24$pred, arima.inf24, arima.sup24, main="prevision à l'horizon h=24",
col=c(1,2,3,3),lty=c(1,1,2,2),lwd=c(3,3,2,2), xlab ="Années")

```

```

#####
##### Modelisation multivariée #####
#####

```

```

###IMPORTATION DES DONNEES

```

```

IPC<- read.csv("IPC.csv", sep=";", header=T)
IPC <- ts(IPC[,2], start=c(2008,1), end=c(2017,12), frequency=12)
IPC

```

```

#Restriction de la variable IPC a l'année 2016

```

```

IPC_2016 <- window(IPC, end=c(2016,12))
IPC_2016

```

```

###STATISTIQUES DESCRIPTIVES de la variable IPC

```

```

summary(IPC_2016)
sd(IPC_2016)

```

```

### Representation de la série

```

```
ts.plot(IPC_2016, xlab="années", ylab="Indice de production du Consommateur", main="Evolution  
de l'indice de production du consommateur")
```

```
## Etudes transformation preliminaires de la serie
```

```
IPC_2016_L <- log(IPC_2016)
```

```
IPC_2016_LC <- IPC_2016_L - mean(IPC_2016_L)
```

```
ts.plot(IPC_2016_LC, xlab="années", ylab="Indice de production du Consommateur", main="Série  
IPC_2016 centrée a transformation logarithmique")
```

```
# affichage de la tendance de la série
```

```
tend_ipc = stl(IPC_2016_LC, "per")
```

```
ts.plot(IPC_2016_LC, main="Représentation de la tendance sur la série", xlab="années")
```

```
lines(tend_ipc$time.series[,2], col = "red")
```

```
# Affichage de l'aCF de la serie
```

```
acf(IPC_2016_LC, xlab="Retards")
```

```
# Test augmenter de Dickey-Fuller
```

```
adf.test(IPC_2016_LC)
```

```
# Test de Phillips Perron
```

```
PP.test(IPC_2016_LC)
```

```
# Test KPSS
```

```
kpss.test(IPC_2016_LC)
```

```
### Differenciation de la série et reprise des tests de stationnaires
```

```
# Differenciation de la série
```

```
IPC_2016_LCD <- diff(IPC_2016_LC)
```

```
# Test augmenter de Dickey-Fuller
```

```

adf.test(IPC_2016_LCD)

# Test de Phillips Perron
PP.test(IPC_2016_LCD)

# Test KPSS
kpss.test(IPC_2016_LCD)

# confirmation visuelle
tend_ipc_D = stl(IPC_2016_LCD, "per")
ts.plot(IPC_2016_LCD, main = "Représentation de la série différenciée", xlab = "années")
lines(tend_ipc_D$time.series[,2], col = "red")

##### Modelisation Multivariée

data_2016 <- cbind(IPI_2016, IPC_2016)

## Observation des deux séries simultanément
plot(data_2016, xlab="Années")

## Observation des deux séries après stationnarisées
varmat <- as.matrix(cbind(IPI_2016_LCD, IPC_2016_LCD))
plot(varmat, xlab="Années", main = "Séries stationnaires")

##### Identification du modèle

## ## Création de la matrice
serie3S = matrix(ncol=2, nrow=length(IPI_2016_LCD))

## Dans cette matrice on met les données stationnarisé de ipi et ipc (IPI_2016_LCD et IPC_2016_LCD)
serie3S[1:length(IPI_2016_LCD),1] = IPI_2016_LCD

```



```
serie3S[1:length(IPI_2016_LCD),2]=IPC_2016_LCD  
VARselect(serie3S)
```

```
## Estimation des paramètres du modèle VAR  
varfit <- VAR(varmat, p=3)  
summary(varfit)  
plot(varfit)
```

```
## Test de Stabilité du modèle  
roots(varfit)  
varfit.stable <- stability(varfit, type = "OLS-CUSUM")  
plot(varfit.stable)
```

```
##### Analyse des résidus
```

```
norm1 <- normality.test(varfit)
```

```
## Representation graphique  
par(mfrow=c(1,2))  
hist(norm1$resid, col="grey", freq=F, main = "Histogramme des residus", xlab="Residus")  
z=seq(min(norm1$resid), max(norm1$resid), 0.01)  
lines(z, dnorm(z,mean(norm1$resid),sd(norm1$resid)), col="blue")  
qqnorm(norm1$resid)  
qqline(norm1$resid, col="red")
```

```
## Test de normalité  
# L'hypothese nulle est la normalit?  
# Si p-value < 5%, on rejette H0.  
norm1$jb.mul  
# p-value = 0.7737 > 0.05 , on ne rejette pas H0.
```

```
##### Autocorrelation des erreurs
```

```
# Test de Portmanteu H0 = non autocorrelation.
```

```
serial.test(varfit, lags.pt = 30)
```

```
# p-value = 0,1002 > 0.05 donc on ne peut rejeter H0.
```

```
#homoscédasticité des résidus
```

```
# test d'arch H0 = non autocorrelation
```

```
arch.test(varfit)
```

```
#### Test de causalité
```

```
IPCcauseIPI=matrix(ncol=2,nrow=107)
```

```
#Remplissage de la matrice IPCcauseIPI
```

```
IPCcauseIPI[1:107,1]=IPI_2016_LCD
```

```
IPCcauseIPI[1:107,2]=IPC_2016_LCD
```

```
IPCcauseIPI=ts(IPCcauseIPI, start=c(2008,1), freq=12,names=c("IPI_2016_LCD","IPC_2016_LCD"))
```

```
causality(VAR(y=IPCcauseIPI,p=3), cause="IPC_2016_LCD")
```

```
#Fonctions de réponses impulsionnelles
```

```
impulresp <- irf(varfit, reponse="IPC_2016_LCD", n.ahead = 20, boot = TRUE)
```

```
plot(impulresp)
```

```
#Décomposition de la variance
```

```
vardecomp <- fevd(varfit)##$IPI_2016_LCD
```

```
plot(vardecomp)
```

```
# Prevision de 12 mois
```

```
prevision12 <- predict(varfit, n.ahead=12)
```

```
plot(prevision12)
```

```
# Prevision de 24 mois
```

```
prevision24 <- predict(varfit, n.ahead=24)
```

```
plot(prevision24)
```