

الجمهورية الجزائرية الديمقراطية الشعبية  
République Algérienne Démocratique et Populaire

وزارة التعليم العالي والبحث العلمي  
Ministère de L'Enseignement Supérieur et de la Recherche Scientifique

جامعة سعد دحلب البليدة-1  
Université Saad Dahlab Blida -1-



**Mémoire de fin D'études**

En vue de l'obtention du diplôme Master

**Faculté de sciences**  
**Département : Mathématiques**

**Spécialité : Modélisation stochastique et statistique**

**THEME**

Modélisation temporelle et prévision par des séries chronologiques.

Simulation et application sur des modèles SARIMA et des données réelles portent le prix du  
baril de Pétrole.

**Présenté par :**  
KERIREM Hayat.

**Soutenu le : 25/07/2019**

TOUAIT Naziha.

Devant le jury composé de :

**Président :** Mr. O. TAMI      MAA      Université de Blida 1

**Promoteur :** Mr. A. RASSOUL      MCA      ENSH de Blida

**Examineurs :** Mr R. FRIHI      MAA      Université de Blida 1

## DÉDICACE

*C'est avec un très grand plaisir que je dédie ce modeste travail aux personnes les plus chères au monde :*

*A ma chère mère qui m'ont permis de continuer mes études dans les meilleures conditions et qui m'ont appris à ne jamais baissé les bras.*

*Je dédie aussi cette modeste réalisation à :*

*Mes très chères soeurs ; mes frères. Je n'oublie pas mes tantes ; mes oncles et à toutes la familles : Kerirem, Lekbedj.*

*A mon encadreur RASSOUL Abdelaziz, en espérant qu'il trouve dans ce travail le témoignage de ma profonde gratitude.*

*A tout mes enseignants de l'université de Saad Dahleb sans exception.*

*A tous mes collègues et mes chers amis.*

*A tout ceux qui m'aiment, m'aident, m'encouragent toujours pour continuer sur la bonne voie.*

*K. Hayat*

## DÉDICACE

*A mes chers parents, pour tous leurs sacrifices, leur soutien et leurs prières tout au long de mes études.*

*A toute ma famille pour leur soutien tout au long de mon parcours universitaire.*

*A toute mes amis.*

*T. Naziha*

## REMERCIÉMENTS

Tout d'abord, nous remercions *ALLAH* qui donne-nous accordée la volonté et le courage pour réaliser ce mémoire.

Nous aimerions exprimer nos profonde gratitude à notre promoteur RASSOUL Abdelaziz avec qui nous avons eu le plaisir de travailler sous sa direction pour ses conseils, et ses critiques constructives.

Nous remercions sincèrement, les membres de jury d'avoir bien voulu accepter de faire partie de la commission d'examen.

Nous remerciers aussi l'ensemble de nos camarades de master, nos amis, nos proches et nos familles (au sens large) qui nous ont soutenus durant ce travail, et spécialement à tous ce qui nous ont encouragé à finir ce travail et qui nous ont accompagnés dans tous les moments de joie et de tristesse.

Enfin, nous ne saurons terminer cette partie sans exprimer notre gratitude à nos parents, nos frères et nos soeurs qui nous ont toujours nous soutenus et encouragés pendant nos études.

*Merci à tous et à toutes*

# TABLE DES MATIÈRES

0.1	Résumé . . . . .	1
0.2	Abstract . . . . .	1
<b>1</b>	<b>Généralités sur les séries chronologiques</b>	<b>1</b>
1.1	Introduction . . . . .	1
1.2	Définition d'une série chronologique . . . . .	1
1.3	Domaines d'application . . . . .	2
1.4	Types de séries chronologiques . . . . .	3
1.4.1	Séries continues . . . . .	3
1.4.2	Séries discrètes . . . . .	3
1.5	Objectifs de l'analyse d'une chronique . . . . .	3
1.5.1	Modélisation . . . . .	3
1.5.2	Prévision . . . . .	3
1.6	Composantes d'une série chronologique . . . . .	3
1.7	Modèles de décomposition . . . . .	4
1.7.1	Modèle additif . . . . .	5
1.7.2	Modèle multiplicatif . . . . .	5
1.7.3	Modèle mixte . . . . .	5
1.8	Choix du modèle . . . . .	5
1.9	Opérateurs définis sur une série chronologique . . . . .	5
1.9.1	Opérateur de retard . . . . .	6
1.9.2	Opérateur de différence d'ordre $d$ . . . . .	6
1.10	Filtres linéaires et moyennes mobiles . . . . .	6
1.10.1	Filtres linéaires . . . . .	6
1.10.2	Moyennes mobiles . . . . .	7
1.10.3	Moyenne mobile et composante saisonnière . . . . .	7
1.11	Stationnarité . . . . .	8
1.11.1	Relation entre stationnarité faible et stricte . . . . .	8
1.11.2	Test de Dickey Fuller Augmenté . . . . .	9

1.12	Autocorrélations simple et partielle . . . . .	10
1.12.1	Fonction d'autocovariance et d'autocorrélation . . . . .	10
1.12.2	La fonction d'autocorrélation partielle . . . . .	11
1.13	Bruit blanc . . . . .	12
1.14	Séries linéaires . . . . .	13
1.15	Analyse de la tendance . . . . .	13
1.15.1	Rappels sur la régression linéaire . . . . .	13
1.15.2	La méthode des moindres carrés . . . . .	14
1.15.3	Ajustement tendanciel linéaire par les moindres carrés . . . . .	15
1.15.4	Ajustement tendanciel linéaire par points médians . . . . .	16
1.15.5	Ajustements tendanciels non linéaires . . . . .	16
1.15.6	Estimation non paramétrique . . . . .	16
1.16	Théorème de Wold . . . . .	17
<b>2</b>	<b>Méthode de Box-Jenkins pour l'Analyse et la Prévision des Séries Chrono-</b>	
	<b>logiques</b>	<b>19</b>
2.1	Introduction . . . . .	19
2.2	Processus Aléatoires Stationnaires . . . . .	19
2.2.1	Processus Autorégressif d'ordre $p$ : AR( $p$ ) . . . . .	20
2.2.1.1	Représentation stationnaire . . . . .	21
2.2.1.2	Représentation inversible . . . . .	21
2.2.1.3	Représentation causale . . . . .	22
2.2.1.4	Principales caractéristiques d'un modèle AR( $p$ ) . . . . .	22
2.2.1.5	Estimation des paramètres . . . . .	23
2.2.1.6	Cas d'un Processus AR(1) . . . . .	24
2.2.2	Processus moyenne mobile d'ordre $q$ : MA( $q$ ) . . . . .	30
2.2.2.1	Représentation stationnaire . . . . .	30
2.2.2.2	Représentation causale . . . . .	30
2.2.2.3	Principales caractéristiques d'un modèle MA( $q$ ) . . . . .	30
2.2.2.4	Processus moyenne mobile MA(1) . . . . .	31
2.2.3	Processus ARMA( $p,q$ ) : Modèle autorégressif - Moyenne mobile . . . . .	35
2.2.3.1	Expression d'un ARMA( $p,q$ ) . . . . .	36
2.2.3.2	Principales caractéristiques d'un modèle ARMA( $p;q$ ) . . . . .	37
2.2.4	Processus SARMA . . . . .	39
2.3	Modèles linéaires non-stationnaires . . . . .	39
2.3.1	Processus TS . . . . .	39
2.3.2	Processus DS . . . . .	40
2.3.3	Modèle ARIMA(processus autorégressif moyenne mobile intégré) . . . . .	40
2.3.4	Modèle SARIMA . . . . .	41
2.4	METHODE DE BOX-JENKINS . . . . .	41
2.4.1	Identification . . . . .	42

2.4.2	Estimation . . . . .	43
2.4.2.1	Critères de choix des modèles . . . . .	43
2.4.3	Adéquation et validation du modèle . . . . .	44
2.4.4	La prévision . . . . .	46
2.4.4.1	Prévision d'un modèle AR(p) . . . . .	47
2.4.4.2	Prévision d'un modèle MA(q) . . . . .	47
2.4.4.3	Prévision d'un modèle ARMA(p,q) . . . . .	48
2.4.4.4	Prévisions dans le cas d'un processus ARIMA(p,d,q) . . . . .	48
<b>3</b>	<b>Simulations et Applications</b>	<b>50</b>
3.1	Introduction . . . . .	50
3.2	Simulation . . . . .	50
3.2.1	Simuler une série chronologique . . . . .	50
3.2.2	Transformation . . . . .	51
3.2.3	Identification . . . . .	52
3.2.4	Estimation des paramètres . . . . .	54
3.2.5	Tests d'adéquation et choix du modèle (validation) . . . . .	54
3.2.6	Prévision . . . . .	54
3.3	Application . . . . .	54
3.3.1	Application sur les données simulées . . . . .	54
3.3.1.1	Analyse graphique . . . . .	54
3.3.1.2	Tests de stationnarité . . . . .	56
3.3.1.3	Prévision . . . . .	57
3.3.2	Vérification de la stationnarité . . . . .	59
3.3.3	Stationnarisation de la série . . . . .	59
3.3.4	Application sur des données réelles . . . . .	63
3.3.4.1	Analyse graphique de la série "Prix du baril de pétrole" . . . . .	63
3.3.4.2	Transformation de la séries . . . . .	66
3.3.4.3	Identification . . . . .	68
3.3.4.4	Estimation des paramètres . . . . .	69
3.3.4.5	Validation . . . . .	70
3.3.4.6	Prévision . . . . .	70

## TABLE DES FIGURES

1.1	Le graphe de la série Airpassengers. . . . .	2
1.2	Les composantes d'une série chronologique. . . . .	4
2.1	Le graphe du modèle $AR(1) : \phi = 0.8$ . . . . .	27
2.2	Le corrélogramme simple du modèle $AR(1)$ . . . . .	27
2.3	Le corrélogramme partiel du modèle $AR(1)$ . . . . .	28
2.4	Le graphe du modèle $AR(1) : \phi = -0.9$ . . . . .	28
2.5	Le corrélogramme simple du modèle $AR(1) : \phi = -0.9$ . . . . .	29
2.6	Le corrélogramme partiel du modèle $AR(1) : \phi = -0.9$ . . . . .	29
2.7	Le graphe du modèle $MA(2)$ . . . . .	33
2.8	Représentation graphique d'Autocorrélation simple du modèle $MA(2)$ . . . . .	34
2.9	Le corrélogramme partiel du modèle $MA(2)$ . . . . .	34
3.1	Le graphe d'une série Bruit-blanc. . . . .	51
3.2	Représentation graphique d'autocorrélation simple de la série Bruit blanc. . . . .	53
3.3	Le corrélogramme partiel de la série Bruit blanc. . . . .	53
3.4	Le graphe du modèle $ARIMA(2, 0, 1)$ . . . . .	55
3.5	Le corrélogramme simple du modèle $ARIMA(2, 0, 1)$ . . . . .	55
3.6	Représentation graphique d'autocorrélation partielle du modèle $ARIMA(2, 0, 1)$ . . . . .	56
3.7	Représentation graphique des valeurs prédictives. . . . .	57
3.8	Représentation graphique du modèle SARIMA. . . . .	58
3.9	Le corrélogramme simple. . . . .	58
3.10	Représentation graphique d'autocorrélation partielle. . . . .	59
3.11	Représentation graphique du modèle diff_SARIMA. . . . .	60
3.12	Le corrélogramme simple du modèle fiff_SARIMA. . . . .	60
3.13	Représentation graphique de PACF. . . . .	61
3.14	Le graphe qqnorm. . . . .	62
3.15	Représentation graphique des valeurs prédictives. . . . .	63
3.16	Représentation graphique de la série des Prix du pétrole. . . . .	64

3.17 Représentation graphique d'autocorrélation simple de la série Prix du baril de pétrole. . . . .	64
3.18 Le corrélogramme partiel de la série Prix du baril de pétrole. . . . .	65
3.19 Les composantes de la série Prix du baril de pétrole. . . . .	66
3.20 La représentation graphique de la série différencier. . . . .	67
3.21 Le corrélogramme simple de la série diff. . . . .	68
3.22 Le corrélogramme partiel de la série diff. . . . .	68
3.23 Le graphe qqnorm du modèle choisis. . . . .	70
3.24 . . . . .	71
3.25 Graphe représente les valeurs prédictives de prix du baril de pétrole. . . . .	71

## LISTE DES TABLEAUX

2.1	Les propriétés des fonctions d'autocorrélation simple et partielle du modèle <i>ARMA</i> . . . . .	39
2.2	Les propriétés des fonctions d'autocorrélation simple et partielle des différence modèles . . . . .	42
3.1	Coefficients estimé du modèle. . . . .	62
3.2	Tableau représente le prix du baril de pétrole en dollar à l'Arabie Saoudite. . . . .	63
3.3	Tableau représente les données de la fonction diff du prix du baril de pétrole. . . . .	66
3.4	Tableau represente les différents modèles ajustés. . . . .	69

**ملخص**

الهدف من هذا العمل هو دراسة تقنيات تحليل ونمذجة السلاسل الزمنية. نهتم من خلال هذه الدراسة بدراسة نماذج أريما باستخدام تقنيات بوكس-جنكينز، التي تهدف إلى صياغة وقائع لنمذجة ضواهر بهدف التنبؤ بالقيم المستقبلية؛ نقوم بتطبيق هذا النموذج من خلال مبدأ المحاكاة في السلسلة الزمنية في برنامج R؛ كما نقوم بتطبيق هذه التقنية على معطيات حقيقية تتمثل في دراسة سعر البرميل الخام بترول برنت.  
**كلمات مفتاحية :** سلسلة زمنية؛ بوكس-جنكينز نماذج أريما؛ المحاكاة والتنبؤ .

**0.1 Résumé**

L'objectif de ce travail est d'étudier les techniques d'analyse et de modélisation des séries chronologiques. On s'intéresse essentiellement aux modèles *ARIMA*; en utilisant les techniques de Box et Jenkins. Cette méthode vise à formuler un modèle permettant de représenter une chronique avec comme finalité de prévoir des valeurs futures. On termine ce travail par le principe de simulation sur les séries chronologiques avec *R*; en attachant quelques exemples. A la fin on traite une application sur des données réelles.  
**Mots clés :** Série chronologique; Box-Jenkins; Modèles *ARIMA*; Simulation et Prévision.

**0.2 Abstract**

The objective of this work is to study the technics of analysis and modeling of time series. We are mainly interested in *ARIMA* models; using Box and Jenkins technics. This method is intended to formulate a model for representing a chronicle with the purpose of predicting future values. This work is completed by the simulation principle on the time series with *R*; we attach a few examples. At the end we have to treat an application on real data.

**Key words :** Time series; Box - Jenkins; *ARIMA* models; Simulation and Forecast.

## INTRODUCTION GÉNÉRALE

Une série chronologique est une suite formée d'observations au cours du temps. L'analyse des séries chronologiques est un outil couramment utilisé de nos jours pour la prédiction de données futures. Ce domaine possède beaucoup d'applications en finance, en médecine, en économétrie et en météorologie et dans bien d'autres domaines. L'idée est de prendre un échantillon de données et de construire le meilleur modèle qui ajuste ces données. Ce modèle nous permet de tirer certaines conclusions sur la série.

Notre travail est structuré de la façon suivante : le premier chapitre est consacré pour les préliminaires où nous mentionnons quelques rappels sur les séries temporelles et leurs caractéristiques, outils très importants; ensuite, dans le deuxième chapitre, nous présentons différentes classes de modèles : *AR*, *MA*, *ARMA* et *ARIMA* en étudiant leurs propriétés (la stationnarité, les fonctions d'autocorrélation simples et partielles,...) et la deuxième partie de ce chapitre traite la méthode de Box-Jenkins et leur quatre étapes, notons que l'ouvrage de Box-Jenkins "Time series analysis, forecasting and control" publié en 1970 a proposé une démarche de prévision pour les séries univariées, fondée sur l'utilisation de modèles *ARIMA*.

Finalement, dans le troisième chapitre nous avons faite une simulation et une application de la méthodologie de Box-Jenkins sur des données réelles "Le prix du baril de pétrole" pour obtenir le meilleure modèle qui représente cette série afin de faire la prévision.

# CHAPITRE 1

## GÉNÉRALITÉS SUR LES SÉRIES CHRONOLOGIQUES

### 1.1 Introduction

La théorie des séries chronologiques (ou temporelles) est appliquée de nos jours dans des domaines aussi variés que l'économétrie, la médecine ou la démographie. Dans ce chapitre, plusieurs concepts importants liés à l'analyse des séries chronologique seront abordés. Parmi ceux-ci, on retrouve les notions de bruit blanc, d'autocorrélation, de la stationnarité et les opérateurs définis sur une série chronologique.

### 1.2 Définition d'une série chronologique

**Définition 1.1** On appelle série chronologique une suite d'observations numériques d'une grandeur effectuées à intervalles réguliers au cours du temps.

**Définition 1.2** Une série chronologique, ou chronique ou série temporelle, est une suite finie de données indexée par le temps. Si  $t_1, t_2, \dots, t_n$  sont les  $n$  instants successifs d'observation et si  $X_{t_i}$  est la valeur mesurée à l'instant  $t_i$ , on notera la série chronologique  $(X_t)_{t \in T}$  où  $T$  est l'ensemble ordonné  $T = \{t_1, t_2, \dots, t_n\}$ .

**Exemple 1.1** Série : nombre mensuel de passagers aériens, de janvier 1949 à décembre 1960.

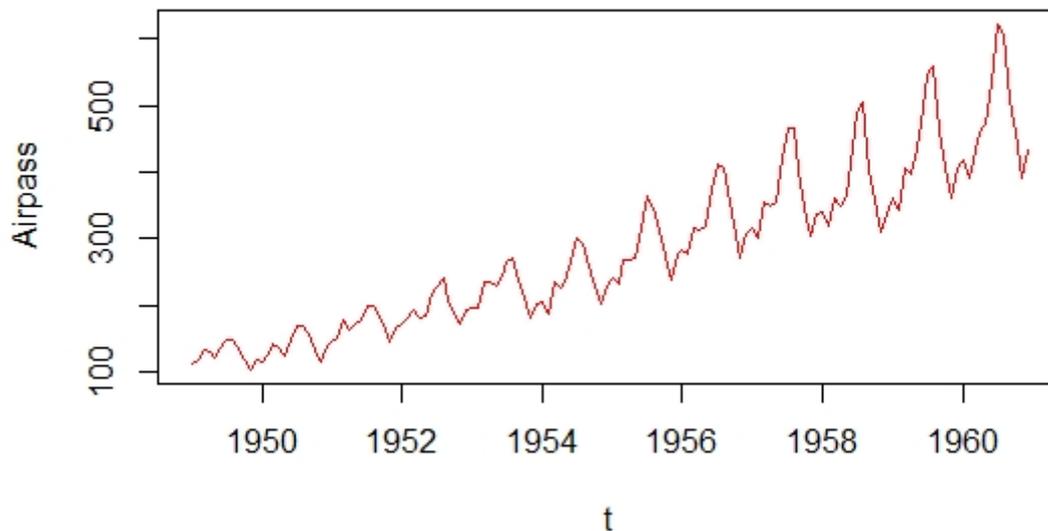


FIGURE 1.1 – Le graphe de la série Airpassengers.

### 1.3 Domaines d'application

On trouve des exemples de séries chronologiques dans de très nombreux domaines. La liste suivante n'est qu'un échantillon : (voir la référence [?])

1. **Finance et économétrie** : évolution des indices boursiers, des prix, des données économiques des entreprises, des ventes et achats de biens, des productions agricoles ou industrielles.
2. **Médecine** : suivi des évolutions des pathologies ; analyse d'électrocardiogrammes.
3. **Sciences de la terre et de l'espace** : indices de marées, variations des phénomènes physiques (météorologie), évolution des taches solaires.
4. **Traitement du signal** : signaux de communications, de radars, de sonars, analyse de la parole.
5. **Traitement des données** : mesures successives de position ou de direction d'un objet mobile (trajectographie).
6. **Traitement d'images** : analyse d'images satellites.
7. **Démographie** : analyse de l'évolution d'une population.
8. **Météorologie** : analyse de données climatiques.
9. **Energie** : prévision de la consommation d'électricité.

## 1.4 Types de séries chronologiques

On distingue deux types de séries chronologiques :

### 1.4.1 Séries continues

C'est une série où l'observation est faite d'une manière continue dans le temps. Exemple : résultat d'un électrocardiogramme.

### 1.4.2 Séries discrètes

C'est une série où l'observation est faite sur des intervalles de temps fixés à priori. Dans ce cas,  $T \in \mathbb{Z}$  et les dates d'observations sont le plus souvent équidistantes par exemple relevés mensuels, trimestriels, ....

**Exemple :** Evolution du nombre de voyageurs utilisant le train.

## 1.5 Objectifs de l'analyse d'une chronique

Les objectifs de l'analyse d'une chronique sont :

### 1.5.1 Modélisation

Elle consiste à :

- Développer des modèles permettant de décrire le comportement d'une ou plusieurs séries chronologiques.
- Mettre au point une méthodologie pour spécifier, estimer, valider (juger) un modèle approprié pour des données particulières.

### 1.5.2 Prévision

Étant données des observations  $X_1, \dots, X_t$ , la prévision consiste à évaluer une valeur non observée,  $X_{t+h}$ . La prévision peut être ponctuelle, ou prendre la forme d'un intervalle de prévision.

## 1.6 Composantes d'une série chronologique

La première étape dans l'étude des séries chronologiques consiste à représenter graphiquement l'événement qui permet d'entrevoir les quatre composantes fondamentales d'une chronique. (voir la référence [10])

Composante irrégulière.

1. **La tendance  $F_t$  (the trend)** : La tendance ou trend  $F_t$  capte l'orientation à long terme de la série.

2. **La composante saisonnière  $S_t$  (saisonnalité)** : Correspond à un phénomène qui se répète à intervalles de temps réguliers (périodiques); en général c'est un phénomène saisonnier d'où le terme de variations saisonnières (toutes les 12 périodes pour des données mensuelles, toutes les 7 périodes pour des données quotidiennes...).
3. **Les variations accidentelles ou résiduelles  $\varepsilon_t$**  : C'est une composante d'erreur  $\varepsilon_t$ . Idéalement, elle est de faible variabilité par rapport aux autres composantes. Elles proviennent de circonstances non prévisibles : catastrophes naturelles, crise boursière, grèves, A ces trois composantes, on ajoute parfois un cycle.
4. **Cycle  $C_t$**  : On appelle cycle un comportement qui se répète assez régulièrement mais avec une périodicité inconnue et changeante (Ces phases durent généralement plusieurs années, mais n'ont pas de durée).

Le graphique suivant présente l'ensemble des composantes précitées.

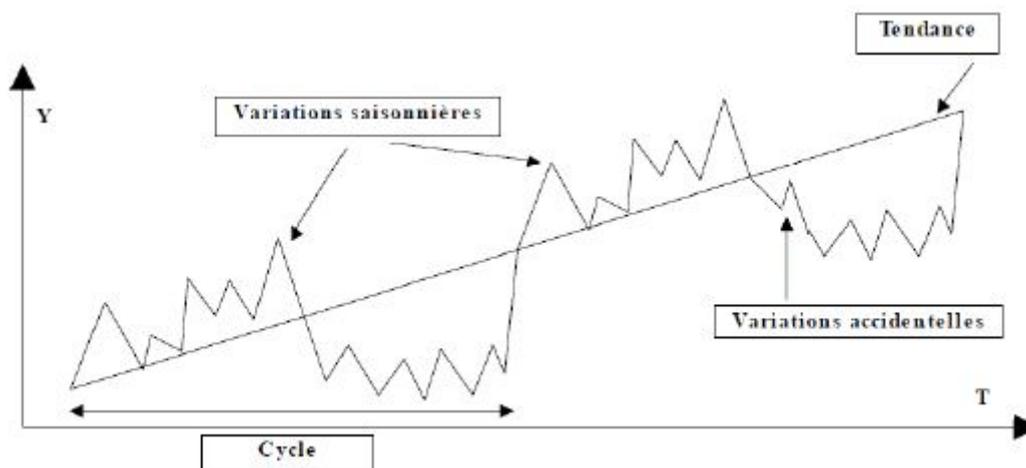


FIGURE 1.2 – Les composantes d'une série chronologique.

**Remarque 1.1** La tendance  $F_t$  prend différentes formes :

- a. **Linéaire** :  $F_t = a + bt$ .
- b. **Quadratique** :  $F_t = a + bt + Ct^2$ .
- c. **Exponentielle** :  $F_t = ae^{bt}$ .

## 1.7 Modèles de décomposition

Pour pouvoir séparer les composantes servant à décrire la série observée, il est nécessaire de préciser leur mode d'interaction.

La plupart des séries chronologiques entrent dans l'un des modèles suivants : (voir les références [8], [16])

### 1.7.1 Modèle additif

Dans ce modèle, toutes les composantes ont un effet additif, la série chronologique s'écrit de la façon suivante :

$$X_t = F_t + S_t + \varepsilon_t + C_t.$$

### 1.7.2 Modèle multiplicatif

Dans ce modèle, toutes les composantes ont un effet multiplicatif sur la taille de la série chronologique :

$$X_t = F_t \times S_t \times \varepsilon_t \times C_t.$$

### 1.7.3 Modèle mixte

Il s'agit là des modèles où addition et multiplication sont utilisées. On peut supposer, par exemple, que la tendance, la composante saisonnière, et cyclique ont la forme multiplicative, alors que les fluctuations irrégulières sont additives :

$$X_t = F_t \times S_t \times C_t + \varepsilon_t.$$

## 1.8 Choix du modèle

**Graphiquement** : Constatation graphique du parallélisme ou non des droites des minima et des maxima :

- Droites parallèles : modèle additif.
- Droites non parallèles : modèle multiplicatif.

**Algébriquement** : On recherche le coefficient directeur de ces deux droites par une méthode analytique (deux points).

- Coefficients directeurs très proches : le modèle choisi sera additif.
- Coefficients directeurs très différents : le modèle choisi sera multiplicatif.

## 1.9 Opérateurs définis sur une série chronologique

Si la série  $X_t$  n'est pas stationnaire l'opérateur suivant rendra cette série stationnaire. (voir les références[6], [13])

### 1.9.1 Opérateur de retard

**Définition 1.3** On note indifféremment  $B$ , l'opérateur qui fait passer de  $X_t$  à  $X_{t-1}$  :

$$BX_t = X_{t-1}. \quad (1.1)$$

**Remarque 1.2** L'opérateur  $B$  est linéaire et inversible . Son inverse  $B^{-1} = F$  est défini par :

$$\forall t \in \mathbb{Z}; FX_t = X_{t+1}.$$

L'opérateur  $F$  est appelé opérateur d'avance.

Si on compose  $B$  avec lui-même on obtient  $B^2 = B \circ B$  tel que :

$$\forall t \in \mathbb{Z}; B^2 X_t = B(BX_t) = BX_{t-1} = X_{t-2}.$$

On peut itérer cette application et définir par récurrence :

$$B^k X_t = X_{t-k}; k \in \mathbb{N}.$$

Par convention ,  $B^0$  est l'opérateur identité  $I$ .

### 1.9.2 Opérateur de différence d'ordre $d$

**Définition 1.4** On définit l'opérateur  $\Delta^d$  de différence d'ordre  $d$  comme l'opérateur linéaire tel que :

$$\Delta^d X_t = (1 - B)^d X_t. \quad (1.2)$$

On peut aussi prendre l'opérateur d'ordre 1 est :

$$\Delta X_t = (1 - B)X_t = X_t - BX_t = X_t - X_{t-1},$$

C'est la série des accroissements, alors que la différence seconde donne la série  $\Delta^2$  des «accroissements des accroissements». On a :

$$\Delta^2 X_t = \Delta(\Delta X_t) = (1 - B)^2 X_t = (1 - 2B + B^2)X_t = X_t - 2X_{t-1} + X_{t-2}.$$

## 1.10 Filtres linéaires et moyennes mobiles

Nous donnons ici deux définitions importantes, pour l'étude des séries chronologiques : (voir la référence [21])

### 1.10.1 Filtres linéaires

Un filtre linéaire d'ordre  $m = p_1 + p_2 + 1$  est défini par :

$$\begin{aligned}
 FL &= \sum_{j=-p_1}^{p_2} \omega_j B^{-j} \\
 &= \omega_{-p_1} B^{p_1} + \omega_{-p_1+1} B^{p_1-1} + \dots + \omega_{-1} B + \omega_0 I + \omega_1 F + \dots + \omega_{p_2-1} F^{p_2-1} + \omega_{p_2} F^{p_2}.
 \end{aligned} \tag{1.3}$$

où  $p_1, p_2 \in \mathbb{N}$  et  $\omega_j \in \mathbb{R}$ .

### 1.10.2 Moyennes mobiles

Une moyenne mobile d'ordre  $m = p_1 + p_2 + 1$  est un filtre linéaire tel que :

$$\sum_{j=-p_1}^{p_2} \omega_j = 1, \text{ pour tout } j = -p_1, \dots, p_2. \tag{1.4}$$

- Beaucoup de moyennes mobiles ont des poids  $\omega_j$  positifs, mais pas toutes.
- Une moyenne mobile est symétrique si  $p_1 = p_2 = p$  et  $\omega_j = \omega_{-j}$ , pour tout  $j = 1, \dots, p$ .
- Une moyenne mobile symétrique est dite non-pondérée si  $\omega_j = cst$  pour tout  $j = -p_1, \dots, p_2$ .

### 1.10.3 Moyenne mobile et composante saisonnière

Une moyenne mobile est un outil intéressant pour lisser une série temporelle et donc pour enlever une composante saisonnière. On utilise de préférence des moyennes mobiles non-pondérées d'ordre égal à la période, par exemple d'ordre 7 pour des données journalières, d'ordre 12 pour des données mensuelles. Par exemple, pour enlever la composante saisonnière due au jour de la semaine, on peut appliquer une moyenne mobile non-pondérée d'ordre 7.

$$MM_7 = \frac{1}{7}(B^3 + B^2 + B + I + F + F^2 + F^3).$$

Cette moyenne mobile accorde le même poids à chaque jour de la semaine. En effet :

$$MM_7(X_t) = \frac{1}{7}(X_{t-3} + X_{t-2} + X_{t-1} + X_t + X_{t+1} + X_{t+2} + X_{t+3}).$$

**Remarque 1.3** — Pour les composantes saisonnières d'une période paire, il n'existe pas de moyennes mobiles centrées non-pondérées.

- Si la période est paire et égale à  $m$ , on utilise une moyenne mobile accordant un demi-poids aux deux extrémités.
- Par exemple, pour des données trimestrielles ( $m = 4$ ), la moyenne mobile est définie par :

$$MM_4 = \frac{1}{4}\left(\frac{1}{2}B^2 + B + I + F + \frac{1}{2}F^2\right),$$

En effet

$$MM_4(X_t) = \frac{1}{4} \left( \frac{1}{2} X_{t-2} + X_{t-1} + X_t + X_{t+1} + \frac{1}{2} X_{t+2} \right).$$

## 1.11 Stationnarité

La stationnarité est une caractéristique d'une série chronologique qui implique que le comportement de la série ne dépend pas du temps. En particulier, on dit qu'une série  $X_t$  est stable si elle ne comporte pas de tendance à la hausse ou à la baisse.

Plus formellement, on distingue deux types de stationnarité, à savoir stricte et faible. (voir les références [5],[13])

**Définition 1.5 (Stationnarité faible)** *Un processus  $(X_t)_{t \in \mathbb{Z}}$  est dit stationnaire ou faiblement stationnaire ou stationnaire au second ordre si :*

- i)  $E(X_t^2) < \infty$ ;
- ii)  $E(X_t) = \mu; \forall t \in \mathbb{Z}$  (constante indépendante de  $t$ );
- iii)  $Cov(X_t; X_{t-h}) = E(X_t X_{t-h}) - E(X_t)E(X_{t-h}) = \gamma(h), \forall t \in \mathbb{Z}, \forall h \in \mathbb{Z}$  (ne dépend pas du temps).

En résumé, un processus  $X_t$  est dit stationnaire du second ordre si sa moyenne, sa variance et sa covariance sont indépendantes du temps et si sa variance est finie.

**Définition 1.6 (Stationnarité stricte)** *Un processus  $(X_t)_{t \in \mathbb{Z}}$  est dit strictement stationnaire ou stationnaire au sens strict si les lois jointes de  $(X_{t_1}, \dots, X_{t_k})$  et de  $(X_{t_1+h}, \dots, X_{t_k+h})$  sont identiques pour tout entier positif  $k$  et pour tous  $t_1, \dots, t_k \quad \forall h \in \mathbb{Z}$ .*

Intuitivement, une série chronologique strictement stationnaire doit avoir le même comportement statistique sur des intervalles de temps égaux.

### 1.11.1 Relation entre stationnarité faible et stricte

Un processus strictement stationnaire du second ordre est faiblement stationnaire.

La réciproque n'est pas vraie en général.

**Contre-exemple :**

Soit  $(X_t)$  une suite de variables aléatoires indépendantes telle que :

- $X_t \sim \exp(1)$  lorsque  $t$  est pair,
- $X_t \sim \mathcal{N}(1, 1)$  lorsque  $t$  est impair,

Alors  $(X_t)$  est stationnaire avec  $\gamma_X(0) = 1$  et  $\gamma_X(h) = 0$  lorsque  $h \neq 0$ . Cependant  $X_1$  et  $X_2$  n'ont pas la même loi donc  $(X_t)$  n'est pas strictement stationnaire.

### 1.11.2 Test de Dickey Fuller Augmenté

Les tests de stationnarité permettent de vérifier si une série est stationnaire ou non. Le test de Dickey-Fuller, le test augmenté de Dickey-Fuller (*ADF*), ou encore le test de Phillips-Perron (*PP*) (voir la référence [19]), sont des tests de racine unitaire, pour lesquels l'hypothèse nulle est que la série a été générée par un processus présentant une racine unitaire, et donc, qu'elle n'est pas stationnaire. On donne ici le test de *ADF*. (voir la référence [7])

Les hypothèses du test de *ADF* se définissent de la façon suivante ; Un processus non stationnaire, il correspond à une de ces formes de non stationnarité :

$$\Delta X_t = \phi X_{t-1} + \sum_{i=1}^p \alpha_i \Delta X_{t-i} + \varepsilon_t \text{ avec } \phi = \rho - 1. \quad \text{Modèle[1]}$$

$$\Delta X_t = \alpha + \phi X_{t-1} + \sum_{i=1}^p \alpha_i \Delta X_{t-i} + \varepsilon_t. \quad \text{Modèle[2]}$$

$$\Delta X_t = \alpha + \beta t + \sum_{i=1}^p \alpha_i \Delta X_{t-i} + \varepsilon_t. \quad \text{Modèle[3]}$$

1. **Test de Dickey et Fuller augmenté** : Robuste à l'autocorrélation par rapport au test de Dickey-Fuller.

$$\begin{cases} H_0 : X_t \sim I(1) \\ H_1 : X_t \text{ n'est pas } I(1) \end{cases} \text{ tel que : } I(1) \text{ est un racine unitaire.}$$

Et

$$ADF = \frac{T \hat{\phi}}{(1 - \hat{a}_1 - \dots - \hat{a}_p)}. \quad (1.5)$$

2. **Test de Phillips-Perron** : Robuste à l'hétéroscédasticité

$$PP = T \hat{\phi} - \frac{1}{2} T^2 \frac{SE(\hat{\phi})}{\hat{\sigma}^2} (\hat{\lambda}^2 - \hat{\sigma}^2), \quad (1.6)$$

$$\hat{\sigma}^2 = \frac{1}{T-k} \sum_{t=1}^T (\hat{\varepsilon}_t^2 - \hat{\varepsilon}_t) : \text{résidus de la régression,}$$

$$SE(\hat{\phi}) = \text{écart-type de } \phi,$$

$$\hat{\lambda}^2 = \hat{\sigma}^2 + 2 \sum_{j=1}^q \left(1 - \frac{j}{q+1}\right) \hat{\gamma}(j),$$

$$\hat{\gamma}(j) = \frac{1}{T} \sum_{t=j+1}^T \hat{\varepsilon}_t \hat{\varepsilon}_{t-j}.$$

**Valeurs critiques de ADF et PP** :  $P(ADF < c) = \alpha$

—  $d_t = 0$  (processus centré)

$\alpha$	<b>T</b>	25	50	100	250	500	$\infty$
0.05		-7.3	-7.7	-7.9	-8	-8	-8.1

—  $d_t = 1$  (processus avec drift)

$\alpha$	<b>T</b>	25	50	100	250	500	$\infty$
0.05		-12.5	-13.3	-13.7	-14	-14	-14.1

## 1.12 Autocorrélations simple et partielle

Les principales caractéristiques temporelles sont données par l'autocorrélation (simple) et l'autocorrélation partielle. (voir la référence [5])

### 1.12.1 Fonction d'autocovariance et d'autocorrélation

Soit  $(X_t)$  une série temporelle, stationnaire.

**Définition 1.7** La fonction d'autocovariance  $\{\gamma(h)\}_{h \in \mathbb{Z}}$  mesure la covariance entre une variable et cette même variable à des dates différentes, pour un horizon  $h$  :

$$\gamma(h) = \text{cov}(X_t, X_{t-h}) = E[(X_t - E(X_t))(X_{t-h} - E(X_{t-h}))]. \quad (1.7)$$

Ainsi

$$\gamma(0) = \text{Var}(X_t) = E[(X_t - E(X_t))^2] = \sigma_X^2.$$

Elle fournit une information sur la variabilité de la série et sur les liaisons temporelles qui existent entre les diverses composantes de la série  $X_t$ .

**Proposition 1.1** La fonction d'autocovariance vérifie les propriétés suivantes :

- $\gamma(0) = \text{var}(X_t) \geq 0$ ;
- $|\gamma(h)| \leq \gamma_0 \forall h$ ; la fonction autocovariance est donc bornée .
- $\gamma(h) = \gamma(-h) \forall h$ ; autrement dit l'autocovariance est une fonction paire.

Cette fonction étant paire, on ne la représente que pour  $h = 0, 1, 2, \dots$ . On a également : La fonction d'autocovariance d'une série  $(X_t)$ , stationnaire est de type positif.

Cette propriété exprime le fait que la variance d'une combinaison linéaire de  $n$  v.a.  $X_{t_1}, \dots, X_{t_n}$  est positive.

**La fonction d'autocorrélation est définie par :**

$$\rho(h) = \frac{\text{cov}(X_t, X_{t-h})}{\sqrt{\text{var}(X_t)\text{var}(X_{t-h})}} = \frac{\text{cov}(X_t, X_{t-h})}{\text{var}(X_t)} = \frac{\gamma(h)}{\gamma(0)}, h \in \mathbb{Z}. \quad (1.8)$$

Avec  $\gamma(0) = 1$  et  $|\gamma(h)| < 1$  :

La dernière égalité tient car

$$\text{var}(X_{t-h}) = \text{var}(X_t) = \gamma(0).$$

Enfin, en notant que par la stationnarité  $E(X_t) = \mu$ , indépendant de  $t$ , on a en terme d'espérance mathématique :

$$\rho(h) = \frac{E[(X_t - \mu)(X_{t-h} - \mu)]}{E[(X_t - \mu)^2]}.$$

L'équivalent empirique de la fonction d'autocorrélation, notée  $\hat{\rho}(h)$ ; est obtenu à partir de l'estimateur suivant pour l'autocovariance  $\hat{\gamma}(h)$  à l'ordre  $h$  : Etant donné une série observée  $X_t, t = 1, \dots, T$ .

Notons

$$\bar{X} = \frac{1}{T} \sum_{t=1}^T X_t.$$

L'autocovariance empirique d'ordre  $h$  est :

$$\hat{\gamma}(h) = \frac{1}{T} \sum_{t=h+1}^T (X_t - \bar{X})(X_{t-h} - \bar{X}), 0 \leq h \leq T-1. \quad (1.9)$$

Le coefficient d'autocorrélation empirique d'ordre  $h$  est :

$$\hat{\rho}(h) = \frac{\sum_{t=h+1}^T (X_t - \bar{X})(X_{t-h} - \bar{X})}{\sum_{t=1}^T (X_t - \bar{X})^2} = \frac{\hat{\gamma}(h)}{S_T^2}, 0 \leq h \leq T-1. \quad (1.10)$$

On définit la matrice de corrélation (de dimension  $m$ ) de la manière suivante :

$$R(m) = \begin{pmatrix} 1 & \rho(1) & \rho(2) & \dots & \rho(m-1) \\ \rho(1) & 1 & \rho(1) & \dots & \rho(m-2) \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \rho(m-1) & \rho(m-2) & \dots & \rho(1) & 1 \end{pmatrix}.$$

### 1.12.2 La fonction d'autocorrélation partielle

Elle mesure la liaison (linéaire) entre  $X_t$  et  $X_{t-h}$  une fois retiré les liens transitant par les variables intermédiaires  $X_{t-1}, \dots, X_{t-h+1}$ .

**Définition 1.8** *Le coefficient d'autocorrélation partielle d'ordre  $h$ , noté  $r(h)$ , est définie par :*

$$r(h) = \text{corr}(X_t, X_{t-h}/X_{t-1}, \dots, X_{t-h+1}). \quad (1.11)$$

Le coefficient d'autocorrélation partielle d'ordre  $h$  d'un processus stationnaire se calcule de la manière suivante :

$$r(h) = \frac{|R(h)^*|}{|R(h)|}$$

Avec

$$R(h) = \begin{pmatrix} 1 & \rho(1) & \dots & \rho(h-1) \\ \rho(1) & 1 & \dots & \rho(h-2) \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \rho(h-1) & \rho(h-2) & \dots & 1 \end{pmatrix}$$

Et  $R(h)^*$  la matrice  $R(h)$  dans laquelle on a remplacé la colonne  $h$  par :

$$\left( \rho(1) \quad \rho(2) \quad \dots \quad \rho(h) \right)'$$

Soit :

$$R(h)^* = \begin{pmatrix} 1 & \rho(1) & \dots & \rho(1) \\ \rho(1) & 1 & \dots & \rho(2) \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \rho(h-1) & \rho(h-2) & \dots & \rho(h) \end{pmatrix}$$

Ainsi,

$$r(1) = \rho(1), r(2) = \frac{\rho(2) - \rho(1)^2}{1 - \rho(1)^2}, \dots$$

De manière empirique, les autocorrélations partielles s'estiment par l'estimation des autocorrélations simples et en calculant  $\hat{r}(h)$  à partir de la formule ci-dessus.

### 1.13 Bruit blanc

**Définition 1.9** On dit que la suite de variables aléatoires  $(\varepsilon_t)$  constitue un bruit blanc faible si elle possède les propriétés suivantes : (voir la référence [11])

$$\left\{ \begin{array}{l} E(\varepsilon_t) = 0, \text{ pour tout } t \in \mathbb{Z}. \\ E(\varepsilon_t^2) = \sigma_\varepsilon^2 \neq 0. \\ cov(\varepsilon_s; \varepsilon_t) = 0, \text{ si } t \neq s. \end{array} \right.$$

En d'autres termes, les variables aléatoires  $\varepsilon_t$  sont de moyenne nulle, de variance constante et non corrélées.

**Remarque 1.4** On dit que  $(\varepsilon_t)$  est un bruit blanc fort s'il est un bruit blanc faible et que les variables aléatoires  $\varepsilon_t$  sont i.i.d (indépendantes et identiquement distribuées).

Un bruit blanc gaussien  $(\varepsilon_t)$  est une suite de v.a. i.i.d.  $\mathcal{N}(0, \sigma_\varepsilon^2)$ , on note :

$$\varepsilon_t \sim \text{BBN}(0, \sigma_\varepsilon^2).$$

## 1.14 Séries linéaires

**Définition 1.10** Une série  $(X_t)$  est dite linéaire si elle peut s'écrire :

$$X_t = \mu + \sum_{i=-\infty}^{+\infty} \psi_i \varepsilon_{t-i}. \quad (1.12)$$

Où  $\varepsilon_t \sim \text{BBN}(0, \sigma_\varepsilon^2)$ ,  $\psi_0 = 1$  et la suite  $(\psi_i)$  est absolument sommable, c'est-à-dire

$$\sum_i |\psi_i| < \infty.$$

**Définition 1.11** Une série  $(X_t)$  est dite linéaire et causale si elle est linéaire avec  $\psi_i = 0, i < 0$  :

$$X_t = \mu + \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}. \quad (1.13)$$

On admettra qu'une série linéaire est stationnaire. L'étude des séries non causales conduit à des résultats non intuitifs difficilement utilisables, aussi nous ne considérerons parmi les séries linéaires que des séries causales. L'écriture de  $X_t$  comme somme de v.a. non corrélées permet d'obtenir facilement :

$$E(X_t) = \mu, \text{var}(X_t) = \sigma_\varepsilon^2 \left(1 + \sum_{i=1}^{\infty} \psi_i^2\right).$$

$$\gamma(h) = \text{cov}(X_t, X_{t-h}) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}.$$

## 1.15 Analyse de la tendance

Dans ce section , nous nous plaçons dans le cadre d'un modèle composé uniquement d'une tendance et de fluctuations irrégulières et donnons différentes méthodes permettant d'estimer la tendance.

### 1.15.1 Rappels sur la régression linéaire

Lorsqu'une liaison linéaire forte entre deux variables  $X$  et  $Y$  semble raisonnable au vu du nuage de points, on a alors une relation du type :

$$Y \simeq aX + b.$$

Où les coefficients  $a$  et  $b$  sont inconnus.

Le problème est que les points du nuage sont rarement (parfaitement) alignés : ils sont proches d'une droite.

Nous cherchons maintenant la droite qui passe au plus près des points du nuage. Pour cela, il faut donc mesurer l'éloignement des points du nuage par rapport à une droite  $D$  d'équation  $y = ax + b$  puis minimiser un critère d'erreur donné . On peut envisager de minimiser.

— La somme des erreurs en valeur absolue :

$$\min_{a,b} \sum_{i=1}^n |Y_i - aX_i - b|.$$

— La somme des erreurs au carré :

$$\min_{a,b} \sum_{i=1}^n (Y_i - aX_i - b)^2.$$

La méthode des moindres carrés minimisant le second critère est la plus usité.

### 1.15.2 La méthode des moindres carrés

On démontre en minimisant la fonction de deux variables

$$g(a; b) = \sum_{i=1}^n (Y_i - aX_i - b)^2.$$

Que le couple solution  $(\widehat{a}; \widehat{b})$  est donné par :

$$\widehat{a} = \frac{\frac{1}{n} \sum_{i=1}^n Y_i X_i - (\frac{1}{n} \sum_{i=1}^n Y_i)(\frac{1}{n} \sum_{i=1}^n X_i)}{\frac{1}{n} \sum_{i=1}^n X_i^2 - (\frac{1}{n} \sum_{i=1}^n X_i)^2} = \frac{cov(X; Y)}{var(X)}.$$

Et

$$\widehat{b} = \frac{1}{n} \sum_{i=1}^n Y_i - \widehat{a} \frac{1}{n} \sum_{i=1}^n X_i = \bar{Y} - \widehat{a} \bar{X}.$$

La droite d'équation  $y = \widehat{a}x + \widehat{b}$  est appelée droite de régression de  $Y$  en  $X$  et est notée :  $\Delta_{Y/X}$ .

1. Cette droite passe par le point moyen  $M(\bar{X}; \bar{Y})$ .
2. Le coefficient directeur  $a$  de  $\Delta_{Y/X}$ ;  $Cov(X; Y)$  et  $r(X; Y)$  sont de même signe :

- Lorsqu'ils sont positifs, on parle de corrélation positive ( $y$  augmente quand  $x$  augmente).
- Lorsqu'ils sont négatifs, on parle de corrélation négative ( $y$  diminue quand  $x$  augmente).

**Propriétés et interprétation du coefficient de corrélation linéaire**

Afin de confirmer qu'il est raisonnable d'approximer le nuage de points par une droite, on calcule le coefficient de corrélation linéaire :

$$r(X; Y) = \frac{Cov(X; Y)}{\sigma_X \sigma_Y}.$$

a. Le coefficient de corrélation linéaire est symétrique :

$$r(X; Y) = r(Y; X).$$

b. L'inégalité de **Cauchy-Schwarz** donne :

$$-1 \leq r(X; Y) \leq 1.$$

En pratique, il faut commencer par tracer le nuage de points puis calculer  $r(X; Y)$  et ce n'est que si la corrélation linéaire est assez forte que l'on cherchera la droite de régression de  $Y$  en  $X$ .

**1.15.3 Ajustement tendanciel linéaire par les moindres carrés**

Supposons que l'on observe  $T$  valeurs d'une série dont la tendance semble être linéaire. La méthode des moindres carrés décrite au paragraphe précédent consiste à estimer la tendance par une fonction linéaire

$$\widehat{Z}_t = \widehat{a}t + \widehat{b}.$$

Le couple solution  $(\widehat{a}; \widehat{b})$  est donné par :

$$\widehat{a} = \frac{Cov(t; Z_t)}{Var(t)} \text{ et } \widehat{b} = \overline{Z} - \widehat{a}\bar{t}.$$

En posant  $\bar{t} = \frac{1}{T} \sum_{t=1}^T t$ ,  $\overline{Z} = \frac{1}{T} \sum_{t=1}^T Z_t$  et

$$\left\{ \begin{array}{l} Cov(t; Z_t) = \frac{1}{T} \sum_{t=1}^T (t - \bar{t})(Z_t - \overline{Z}). \\ Var(t) = \frac{1}{T} \sum_{t=1}^T (t - \bar{t})^2. \end{array} \right.$$

**Remarque 1.5** *Le coefficient de corrélation linéaire est défini par :*

$$r = \frac{Cov(t; Z_t)}{\sqrt{Var(t)Var(Z_t)}}.$$

La corrélation linéaire entre la date  $t$  et la variable  $Z_t$  est d'autant plus importante que  $|r|$  est proche de 1.

#### 1.15.4 Ajustement tendanciel linéaire par points médians

On suppose ici aussi que la tendance est linéaire. Cette méthode est empirique et ne repose sur aucun critère d'erreur à minimiser. Elle peut cependant s'avérer efficace en présence de valeurs aberrantes. On choisit deux points de coordonnées  $(t_\alpha; Z_\alpha)$  et  $(t_\beta; Z_\beta)$  et on fait passer la droite par ces deux points.

Les coefficients  $(\tilde{a}; \tilde{b})$  vérifient :

$$\begin{cases} Z_\alpha = \tilde{a}t_\alpha + \tilde{b}. \\ Z_\beta = \tilde{a}t_\beta + \tilde{b}. \end{cases} \text{ soit } \begin{cases} \tilde{a} = \frac{Z_\alpha - Z_\beta}{t_\alpha - t_\beta}. \\ \tilde{b} = Z_\alpha - \frac{Z_\alpha - Z_\beta}{t_\alpha - t_\beta} t_\alpha. \end{cases}$$

Pour choisir les deux points, on constitue deux sous-séries d'observations en général d'effectifs égaux. Puis on prend les points médians de chaque sous-série.

#### 1.15.5 Ajustements tendanciels non linéaires

Lorsque la tendance n'est pas linéaire, une technique simple consiste à se ramener à un ajustement linéaire après un changement de variable approprié.

**Exemple 1.2** Si  $Z_t = at^2 + b$ , en posant  $Y_t = t^2$ , on se ramène à  $Z_t = aY_t + b$  et on peut faire un ajustement linéaire entre  $Y_t$  et  $Z_t$ .

Si  $Z_t = b \exp(at)$ , en posant  $Y_t = \ln(Z_t)$ , on se ramène à  $Y_t = at + \ln(b)$  et on peut faire un ajustement linéaire entre  $Y_t$  et  $t$ .

#### 1.15.6 Estimation non paramétrique

Dans certaines situations, il n'est pas facile de trouver le degré du polynôme d'ajustement pour  $Z_t$  ou de changement de variable adéquat. On pourrait utiliser un polynôme avec un degré élevé mais le nombre de paramètres à estimer serait important et rendrait les calculs fastidieux. Par ailleurs, on ne sait pas non plus déterminer l'allure de cette fonction.

Dans cette situation, on a recours à la théorie non paramétrique de l'estimation de la tendance qui ne suppose rien sur celle-ci a priori et on approxime la tendance par la moyenne mobile arithmétique d'ordre  $k$ .

$$Z_t^* = M_k(Z_t).$$

Alors la tendance à la date  $t$  peut être estimée par la moyenne mobile (centrée) d'ordre  $k$  à la date  $t$ .

## 1.16 Théorème de Wold

Le théorème de Wold est fondamental pour l'analyse de séries temporelles stationnaires. (voir la référence [11])

**Théorème 1.1** *Tout processus  $(X_t)_{t \in \mathbb{Z}}$  faiblement stationnaire peut s'écrire sous la forme :*

$$X_t = \sum_{j=0}^{+\infty} \psi_j \varepsilon_{t-j} + k_t. \quad (1.14)$$

Où les paramètres  $\psi_1, \psi_2, \dots$  sont des réels tels que  $\psi_0 = 1$  et  $\sum_{j=0}^{\infty} \psi_j^2 < +\infty$ ,  $(\varepsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc et  $k_t$  est une composante linéaire telle que :

$$\text{Cov}(k_t, \varepsilon_t) = 0 \text{ pour tout } j \in \mathbb{Z}.$$

Selon le théorème de Wold, tout processus stationnaire d'ordre 2 peut être représenté comme une somme pondérée infinie de chocs passés, caractérisés par un bruit blanc.

Dans ce qui suit, deux exemples de séries chronologiques stationnaires sont présentés. On obtiendra, dans chaque cas, leur représentation selon la formule du théorème de Wold.

**Exemple 1.3** *Soit le processus :*

$$X_t = \sum_{j=0}^{+\infty} \left(\frac{1}{2}\right)^j v_{t-j+1}.$$

Où  $(v_t)_{t \in \mathbb{Z}}$  est un bruit blanc gaussien de variance unitaire. Afin que la condition  $\psi_0 = 1$  du théorème de Wold soit satisfaite, il suffit de poser :

$$X_t = \frac{1}{2} v_t.$$

On a alors :

$$X_t = \sum_{j=0}^{+\infty} \left(\frac{1}{2}\right)^{j-1} \varepsilon_{t-j+1} = \varepsilon_t + \sum_{j=0}^{+\infty} \left(\frac{1}{2}\right)^j \varepsilon_{t-j}.$$

Où  $(\varepsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc de variance

$$\text{Var}(\varepsilon_t) = \text{Var}\left(\frac{v_t}{2}\right) = \frac{1}{4}.$$

Ainsi,  $\psi_0 = 1$  et  $\psi_j = \left(\frac{1}{2}\right)^j$ . On note que la condition

$$\sum_{j=0}^{+\infty} \psi_j^2 < +\infty.$$

Est également satisfaite car :

$$\sum_{j=0}^{\infty} \psi_j^2 = \sum_{j=0}^{+\infty} \left(\frac{1}{2}\right)^{2j} = \sum_{j=0}^{+\infty} \left(\frac{1}{4}\right)^j = \frac{4}{3}.$$

Où le résultat :

$$\sum_{j=0}^{+\infty} \alpha^j = \frac{1}{1-\alpha}, |r| < 1.$$

a été utilisé.

## CHAPITRE 2

# MÉTHODE DE BOX-JENKINS POUR L'ANALYSE ET LA PRÉVISION DES SÉRIES CHRONOLOGIQUES

### 2.1 Introduction

La méthodologie de Box Jenkins vise à formuler un modèle permettant de représenter une chronique avec comme finalité de prévoir des valeurs futures. De ce fait l'objet de cette méthode est de modéliser une série temporelle en fonction de ses valeurs passées en utilisant des modèles *ARMA* et *ARIMA*. (voir la référence [17])

### 2.2 Processus Aléatoires Stationnaires

Les économètres ont mis au point toute une galerie de profils théoriques de processus temporels qui permettent de modéliser une gamme étendue de séries chronologiques.

Les processus *AR* et *MA* sont basés sur l'hypothèse suivante :

**Hypothèse :** On suppose que notre série  $X_t$  est générée à partir d'une autre série  $\varepsilon_t$  qui suit une loi de Gauss de moyenne 0 et d'écart type  $\sigma_\varepsilon$ , sous la forme :

$$X_t = \mu + \theta_t \varepsilon_{t-1} + \theta_{t-1} \varepsilon_{t-2} + \dots$$

$\varepsilon_t$  vérifie donc :

$$\left\{ \begin{array}{l} \varepsilon \sim BBN(0, \sigma_\varepsilon^2). \\ \gamma(h) = cov(\varepsilon_t, \varepsilon_{t+h}) = \begin{cases} \sigma_\varepsilon^2 & \text{si } h = 0. \\ 0 & \text{si } h \neq 0. \end{cases} \end{array} \right.$$

La série  $\varepsilon_t$  est un bruit blanc. c'est l'exemple parfait d'une série stationnaire puisque son espérance est toujours nulle, que sa dispersion est stable et qu'il y a indépendance entre les observations , donc la valeur  $\mu$  est la moyenne de la série.

## 2.2.1 Processus Autorégressif d'ordre $p$ : AR( $p$ )

Dans l'étude d'une série chronologique, il est naturel de penser que la valeur de la série à la date  $t$  peut dépendre des valeurs prises aux dates précédentes :

$$X_t = f(X_{t-1}, X_{t-2}, \dots). \quad (2.1)$$

Il n'est généralement pas nécessaire de prendre en compte tout le passé de la série et on peut le plus souvent se limiter à  $p$  valeurs :

$$X_t = f(X_{t-1}, X_{t-2}, \dots, X_{t-p}) + \varepsilon_t.$$

où  $\{\varepsilon_t\}$  est un bruit blanc  $\varepsilon_t \sim \text{BBN}(0, \sigma_\varepsilon^2)$ .

**Définition 2.1** Dans un modèle Autorégressif d'ordre  $p$  noté AR( $p$ ), la valeur future d'une variable est supposée être une combinaison linéaire de  $p$  observations passées et d'une erreur aléatoire associée à un terme constant, le modèle AR( $p$ ) peut être exprimé par : (voir la référence [22])

$$\begin{aligned} X_t &= c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t, \\ &= c + \sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t. \end{aligned} \quad (2.2)$$

Parfois le terme constant  $c$  est omis pour des raisons de simplicité d'où :

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t,$$

Alors :

$$\begin{aligned} X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} &= \varepsilon_t, \\ (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t &= \varepsilon_t, \\ \Phi(B) X_t &= \varepsilon_t. \end{aligned}$$

Où :

$X_t$  et  $\varepsilon_t$  sont respectivement la valeur réelle et l'erreur aléatoire (le choc aléatoire) à la période  $t$ .

$\phi_i$  : des coefficients réels sachant que  $i = 1, \dots, p$ .

$\Phi(B)$  : est une équation caractéristique de  $B$  (L'opérateur retard) .

### Condition de la stationnarité

**Théorème 2.1** Un processus AR( $p$ ) est stationnaire si le module des solutions (les racines) de son équation caractéristique est à chaque fois strictement supérieur à 1 en valeur absolue.

- Le polynôme des retards d'un processus  $AR(1)$  qui est  $X_t = \phi X_{t-1} + \varepsilon_t$  s'écrit  $(1 - \phi B)X_t = \varepsilon_t$ . Sa résolution (en remplaçant l'opérateur retard  $B$  par la simple valeur  $x$ ) donne  $1 - \phi x = 0 \implies x = \frac{1}{\phi}$ . D'après le théorème pour que  $AR(1)$  soit stationnaire, il faut que  $\left| \frac{1}{\phi} \right| > 1 \implies |\phi| < 1$ .
  - Le polynôme des retards d'un processus  $AR(2)$  qui est  $X_t = \phi X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t$  s'écrit  $(1 - \phi B - \phi_2 B^2)X_t = \varepsilon_t$  pour que soit  $AR(2)$  stationnaire il faut que la solution de l'équation du second degré  $(1 - \phi_1 x - \phi_2 x^2)$  sera supérieur à 1 en valeur absolue.
- Exemple 2.1** Soit le modèle  $AR(2)$  :

$$X_t = -0.1X_{t-1} + 0.5X_{t-2} + \varepsilon_t.$$

Où

$$\Phi(B) = 1 + 0.1B - 0.5B^2.$$

Ses racines sont  $-1.31$  et  $1.51$

Comme  $\begin{cases} |-1.31| > 1 \\ |1.51| > 1 \end{cases}$  Alors ce processus est stationnaire.

### 2.2.1.1 Représentation stationnaire

Ce processus est pour l'instant défini sous forme implicite et en particulier il n'est pas certain que cette dernière équation admette toujours une solution stationnaire.

Si le polynôme  $\Phi$  a toutes ses racines de module différentes de 1, on peut inverser l'opérateur  $\Phi(B)$ . On en déduit que l'équation admet une solution unique, avec l'écriture :

$$X_t = \Phi(B)^{-1} \varepsilon_t = \sum_{i=-\infty}^{+\infty} h_i \varepsilon_{t-1}. \quad (2.3)$$

On peut alors montrer que l'on a  $\sum_{i=-\infty}^{+\infty} |h_i| < \infty$  et donc que la représentation est stationnaire.

### 2.2.1.2 Représentation inversible

**Définition 2.2** Si  $(\xi_t)_t$  peut être exprimé en fonction de  $\{X_s, s \leq t\}$  seulement, i.e.

$$\xi_t = \sum_{i=0}^{+\infty} \phi_i X_{t-i}. \quad (2.4)$$

Alors  $(X_t)_t$  est inversible.

La représentation  $AR(p)$  est inversible par définition.

### 2.2.1.3 Représentation causale

Si  $(X_t)_t$  peut être exprimé en fonction de  $\{\xi_s, s \leq t\}$  seulement, i.e.

$$X_t = \sum_{i=0}^{+\infty} \phi_i \xi_{t-i}. \quad (2.5)$$

Alors  $(X_t)_t$  est causale.

### 2.2.1.4 Principales caractéristiques d'un modèle AR(p)

**La fonction d'autocorrélation :**

**Définition 2.3** La fonction **d'autocovariance** d'un processus autorégressif AR(p) est donnée par l'équation de récurrence suivante pour  $h > 0$  :

$$\gamma(h) = \text{cov}(X_t, X_{t-h}) = E(X_t X_{t-h}).$$

Or

$$X_t X_{t-h} = \phi_1 X_{t-1} X_{t-h} + \phi_2 X_{t-2} X_{t-h} + \dots + \phi_p X_{t-p} X_{t-h} + \varepsilon_t X_{t-h}.$$

D'où

$$\begin{aligned} \gamma(h) &= \phi_1 \gamma_{h-1} + \phi_2 \gamma_{h-2} + \dots + \phi_p \gamma_{h-p} + \underbrace{E(\varepsilon_t X_{t-h})}_{=0 \text{ car } \varepsilon_t \perp X_{t-h}} \quad h > 0, \\ &= \sum_{i=1}^p \phi_i \gamma(h-i). \end{aligned} \quad (2.6)$$

**Remarque 2.1** L'équation  $\gamma(h) = \sum_{i=1}^p \phi_i \gamma(h-i)$  est une équation de **Yule-Walker**.

— Lorsque l'on inclut également l'autocovariance d'ordre 0 il faut également rajouter la variance des résidus pour la première équation. D'où l'équation de **Yule-Walker** donne : (voir la référence [9])

$$\gamma(h) = \sum_{i=1}^p \phi_i \gamma(h-i) + \sigma_\varepsilon^2.$$

Ce terme supplémentaire ne se trouve que dans la première équation car  $X_{t-h}$  est indépendant de  $\varepsilon_t$  où  $h$  est plus grand que zéro (pour  $h > 0$ ,  $E(X_{t-h} \varepsilon_t) = 0$ ).

En divisant l'équation  $\gamma(h)$  par  $\gamma(0)$  :

Où :

$$\gamma(0) = \phi_1 \gamma(1) + \phi_2 \gamma(2) + \dots + \phi_p \gamma(p) + \sigma_\varepsilon^2.$$

On obtient la fonction d'autocorrélation satisfaisant à l'équation de récurrence suivante :

$$\begin{aligned}\rho(h) &= \frac{\gamma(h)}{\gamma(0)} = \phi_1\rho(h-1) + \phi_2\rho(h-2) + \dots + \phi_p\rho(h-p) \quad \forall h \geq 1, \\ &= \sum_{i=1}^p \phi_i\rho(h-i).\end{aligned}\tag{2.7}$$

L'équation peut être exprimée alors telle que :

$$\Phi(B)\rho(h) = 0.$$

**Remarque 2.2** Nous constatons que la fonction d'autocorrélation  $\rho(h)$  d'un processus  $AR(p)$  décroît exponentiellement vers zéro. Cette décroissance exponentielle est lisse si  $\rho(h)$  est positive et oscillante si le signe de  $\rho(h)$  alterne.

**La fonction d'autocorrélation partielle :**

Dans un processus autorégressif d'ordre  $p$  il est possible de montrer que :

$$\Gamma(h) = \begin{cases} \rho(0) = 1 & \text{si } h = 0, \\ \rho(1) = \phi & \text{si } h = 1, \\ \phi_p & \text{si } h = p, \\ 0 & \text{si } h > p. \end{cases}\tag{2.8}$$

**Remarque 2.3** Nous constatons que la fonction d'autocorrélation partielle  $\Gamma(h)$  d'un processus  $AR(p)$  est égale à zéro quand  $h > p$ . Autrement dit la fonction d'autocorrélation partielle d'un processus autorégressif chute à zéro au delà de  $p$ .

### 2.2.1.5 Estimation des paramètres

On dispose d'une observation  $\{X_1, \dots, X_T\}$  de longueur  $T$  d'un processus stationnaire  $X_t$  supposé suivre un modèle  $AR(p)$ , c'est à dire vérifiant :

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t.$$

Avec  $t \in \mathbb{Z}$  et  $\phi_1, \dots, \phi_p$ , des paramètres inconnus. On cherche alors à estimer ces paramètres à l'aide des observations disponibles.

**Estimateurs de Yule-Walker :**

La méthode consiste à reprendre les équations de **Yule-Walker** en inversant les relations : on exprime les coefficients en fonction des autocorrélations, on trouve les paramètres estimés d'après les autocorrélations estimées.

On a vu précédemment que :

$$\begin{aligned}\rho(h) &= \sum_{i=1}^p \phi_i\rho(h-i) \\ &= \phi_1\rho(h-1) + \phi_2\rho(h-2) + \dots + \phi_p\rho(h-p), \forall h \geq 0.\end{aligned}$$

En prenant l'équation sous sa forme matricielle :

$$\begin{cases} \rho(1) = \phi_1\rho(0) + \phi_2\rho(1) + \dots + \phi_p\rho(p-1), \\ \rho(2) = \phi_1\rho(1) + \phi_2\rho(0) + \dots + \phi_p\rho(p-2), \\ \vdots \\ \rho(p) = \phi_1\rho(p-1) + \phi_2\rho(p-2) + \dots + \phi_p\rho(0). \end{cases}$$

Que l'on écrira :

$$R_p\phi = \rho.$$

Alors :

$$\widehat{\phi} = \widehat{R}_p^{-1}\widehat{\rho}. \quad (2.9)$$

Avec :

$$\widehat{R}_p = \begin{pmatrix} 1 & \widehat{\rho}(1) & \widehat{\rho}(2) & \dots & \widehat{\rho}(p-1) \\ \widehat{\rho}(1) & 1 & \widehat{\rho}(1) & \dots & \widehat{\rho}(p-2) \\ \widehat{\rho}(2) & \widehat{\rho}(1) & 1 & \dots & \widehat{\rho}(p-3) \\ \vdots & \vdots & & & \vdots \\ \widehat{\rho}(p-1) & \widehat{\rho}(p-2) & \widehat{\rho}(p-3) & \dots & 1 \end{pmatrix}$$

$$\widehat{\rho} = (\widehat{\rho}(i))_{i=1,\dots,p} \text{ et } \widehat{\phi} \text{ le vecteur des paramètres } \widehat{\phi} = \begin{pmatrix} \widehat{\phi}_1 \\ \vdots \\ \widehat{\phi}_p \end{pmatrix}.$$

**Remarque 2.4** L'estimation d'un modèle  $AR(p)$  par la méthode du **maximum de vraisemblance** est délicate car la **fonction de vraisemblance** est très complexe et n'a pas de dérivée analytique. Cette difficulté provient de l'interdépendance des valeurs ainsi que du fait que les observations antérieures ne sont pas toutes disponibles pour les  $p$  premières valeurs.

### 2.2.1.6 Cas d'un Processus AR(1)

**Définition 2.4** Le processus autorégressif d'ordre 1, noté  $AR(1)$ , est stationnaire et vérifie l'équation :

$$\begin{aligned} X_t &= \phi_1 X_{t-1} + \varepsilon_t = \phi_1 B(X_t) + \varepsilon_t. \\ X_t - \phi_1 X_{t-1} &= \varepsilon_t. \\ (1 - \phi_1 B)X_t &= \varepsilon_t. \end{aligned} \quad (2.10)$$

Où  $\varepsilon_t$  est un bruit blanc de variance  $\sigma^2$  et  $\phi$  est un coefficient réel.

**Stationnarité :** La racine de l'équation polynômiale  $1 - \phi_1 B = 0$  est :

$$B = \frac{1}{\phi_1}$$

AR(1) est stationnaire si  $|B| = \left| \frac{1}{\phi_1} \right| > 1 \iff |\phi_1| < 1$ ,

Donc

$$\begin{aligned} X_t &= (1 - \phi_1 B)^{-1} \varepsilon_t = \sum_{i=0}^{+\infty} \phi_1^i B^i \varepsilon_t, \\ &= \sum_{i=0}^{+\infty} \phi^i \varepsilon_{t-i}. \end{aligned}$$

Ce qui donne :

$$\begin{aligned} E(X_t) &= E\left(\sum_{i=0}^{+\infty} \phi^i \varepsilon_{t-i}\right), \\ &= \sum_{i=0}^{+\infty} \phi^i E(\varepsilon_{t-i}) = 0. \end{aligned}$$

D'autre part, l'autocovariance est donnée par

$$\begin{aligned} \gamma(h) &= \text{cov}(X_t, X_{t+h}) = E\left[\left(\sum_{i=0}^{+\infty} \phi^i \varepsilon_{t-i}\right)\left(\sum_{i=0}^{+\infty} \phi^i \varepsilon_{t+h-i}\right)\right] & (2.11) \\ &= E[(\varepsilon_t + \phi \varepsilon_{t-1} + \phi^2 \varepsilon_{t-2} + \dots)(\varepsilon_{t+h} + \phi \varepsilon_{t+h-1} + \dots + \phi^h \varepsilon_t + \phi^{h+1} \varepsilon_{t-1} + \dots)] \\ &= E(\phi^h \varepsilon_t^2 + \phi^{h+2} \varepsilon_{t-1}^2 + \phi^{h+4} \varepsilon_{t-2}^2 + \dots) \\ &= \sigma^2 \phi^h \sum_{i=0}^{+\infty} \phi^{2i} \text{(série convergente car } |\phi| < 1) \\ &= \frac{\sigma^2 \phi^h}{1 - \phi^2}. \end{aligned}$$

**Variance de X :** pour  $h = 0$ , on peut montrer que :

$$\text{var}(X) = \gamma(0) = \frac{\sigma_\varepsilon^2}{1 - \phi^2} \text{ avec } |\phi| < 1 \quad (2.12)$$

**Autocorrélation de X :** De même, on peut montrer que :  $\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \phi^h$  pour  $h \geq 1$   
(Cette fonction décroît exponentiellement vers 0, en oscillant si  $\phi < 0$ ).

**Autocorrélation partielle de X :**

Dans un processus autorégressif d'ordre  $p$  il est possible de montrer que :

$$\left\{ \begin{array}{l} \Gamma(0) = \rho(0) = 1 \\ \Gamma(1) = \rho(1) = \phi \\ \Gamma(2) = \frac{\begin{vmatrix} 1 & \rho(1) \\ \rho(1) & \rho(2) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{vmatrix}} = \frac{\rho(2) - \rho(1)^2}{1 - \rho(1)^2} = \frac{\phi^2 - \phi^2}{1 - \phi^2} = 0 \\ \Gamma(3) = \frac{\begin{vmatrix} 1 & \rho(1) & \rho(1) \\ \rho(1) & 1 & \rho(2) \\ \rho(2) & \rho(1) & \rho(3) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) & \rho(2) \\ \rho(1) & 1 & \rho(1) \\ \rho(2) & \rho(1) & 1 \end{vmatrix}} = 0 \end{array} \right.$$

D'où

$$\Gamma(h) = \begin{cases} 1 & \text{si } h = 0 \\ \rho(1) = \phi & \text{si } h = 1 \\ 0 & \text{si } h \geq 2 \end{cases} \quad (2.13)$$

La condition de stationarité vise à autoriser le passage à l'écriture  $MA(\infty)$  :

$$\begin{aligned} X_t &= (1 - \phi B)^{-1} \varepsilon_t, \\ &= \sum_{i=0}^{+\infty} \phi^i \varepsilon_{t-i}. \end{aligned}$$

Est un processus  $MA(\infty)$  avec  $\phi^i = \theta_i$ .

— Si  $|B| = 1 \iff |\phi_1| = 1$ ;

Donc le processus s'écrit :

$$\begin{aligned} X_t &= X_{t-1} + \varepsilon_t = X_{t-n} + \sum_{i=0}^{n-1} \varepsilon_{t-i}. \\ \iff &\begin{cases} X_1 = X_0 + \varepsilon_1 \\ X_2 = X_1 + \varepsilon_2 = X_0 + \varepsilon_2 + \varepsilon_1 \\ \vdots \\ X_t = X_0 + \varepsilon_t + \dots + \varepsilon_1 \end{cases} \end{aligned}$$

De plus, en posant  $X_0 = 0$ , on voit tout de suite que la variance de la série  $X_t$  est donnée par :

$$\text{var}(X_t) = t\sigma^2.$$

Est une marche aléatoire (le processus est donc non stationnaire car la variance dépend de  $t$ ).

On peut obtenir deux sortes de corrélogramme suivant si  $\phi$  est positif ou négatif.

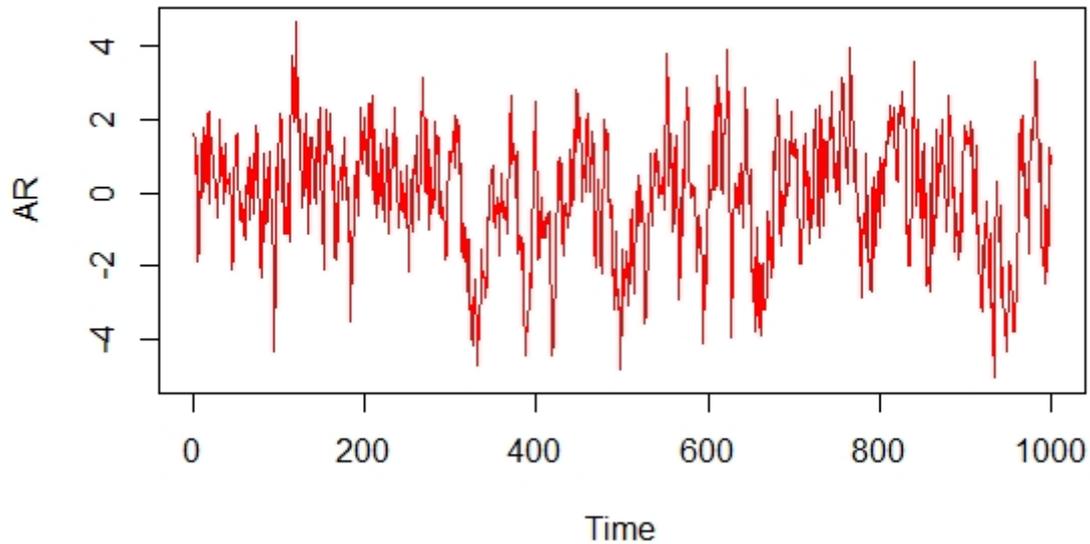


FIGURE 2.1 – Le graphe du modèle AR(1) :  $\phi = 0.8$ .

Son corrélogramme simple est :

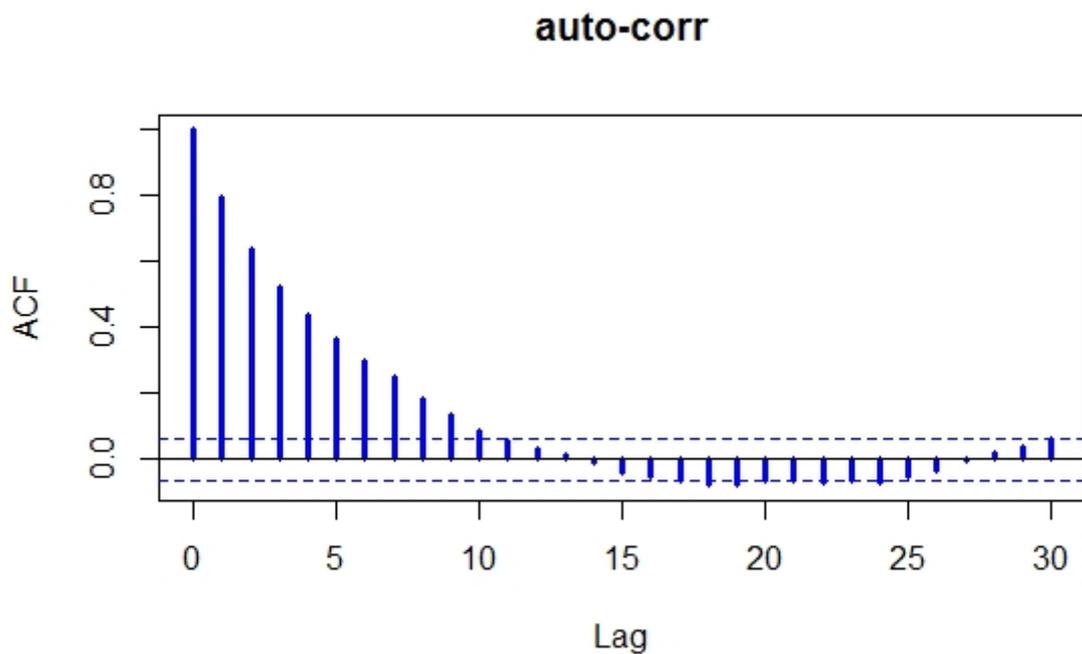


FIGURE 2.2 – Le corrélogramme simple du modèle AR(1).

Et le corrélogramme partiel est :

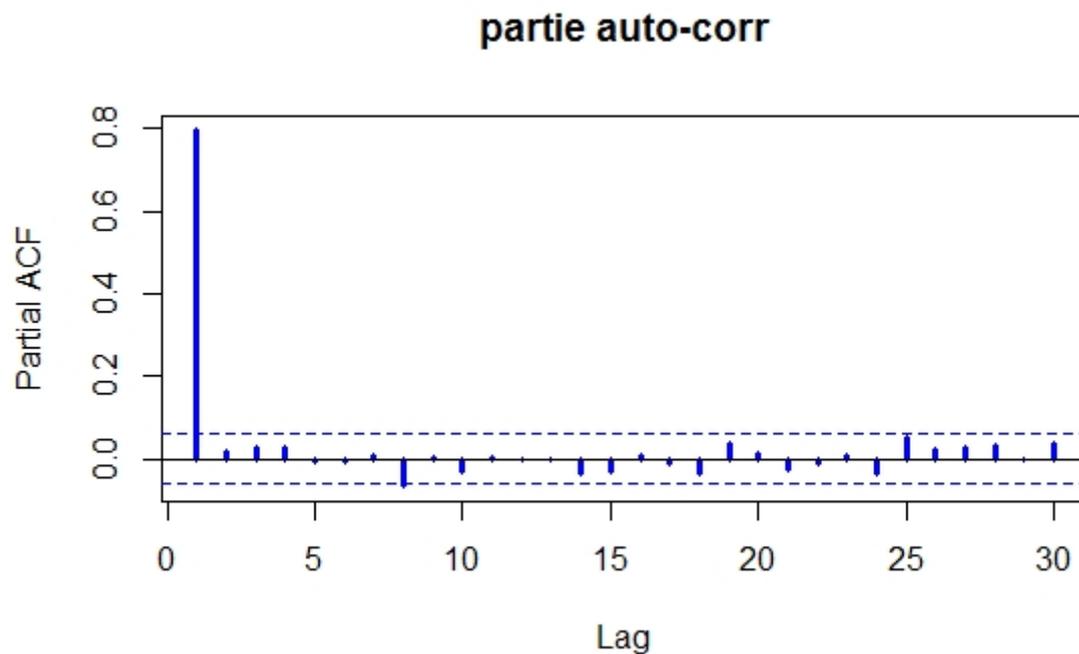


FIGURE 2.3 – Le corrélogramme partiel du modèle AR(1).

Aussi

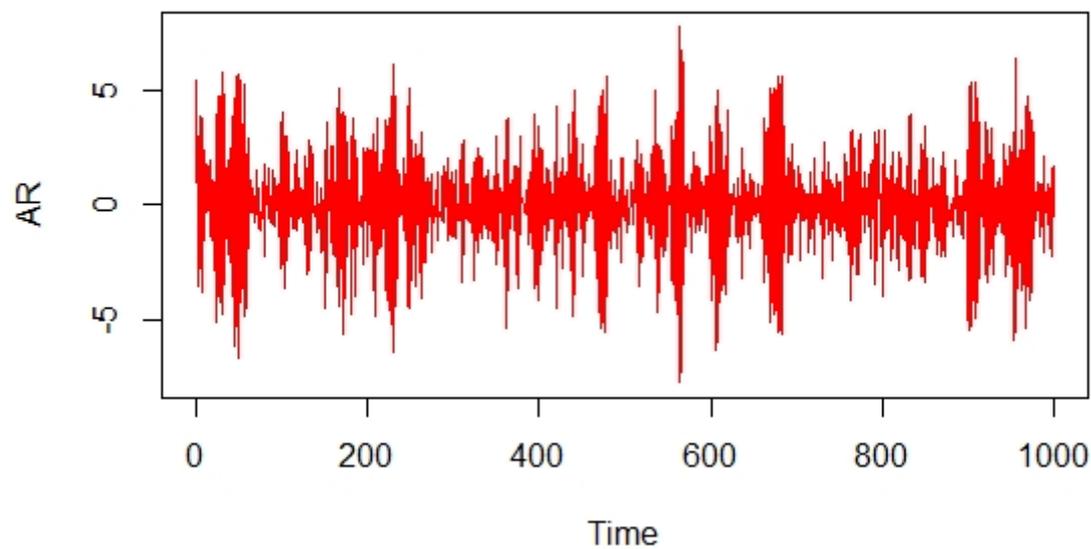


FIGURE 2.4 – Le graphe du modèle AR(1) :  $\phi = -0.9$ .

Le corrélogramme simple est :

---

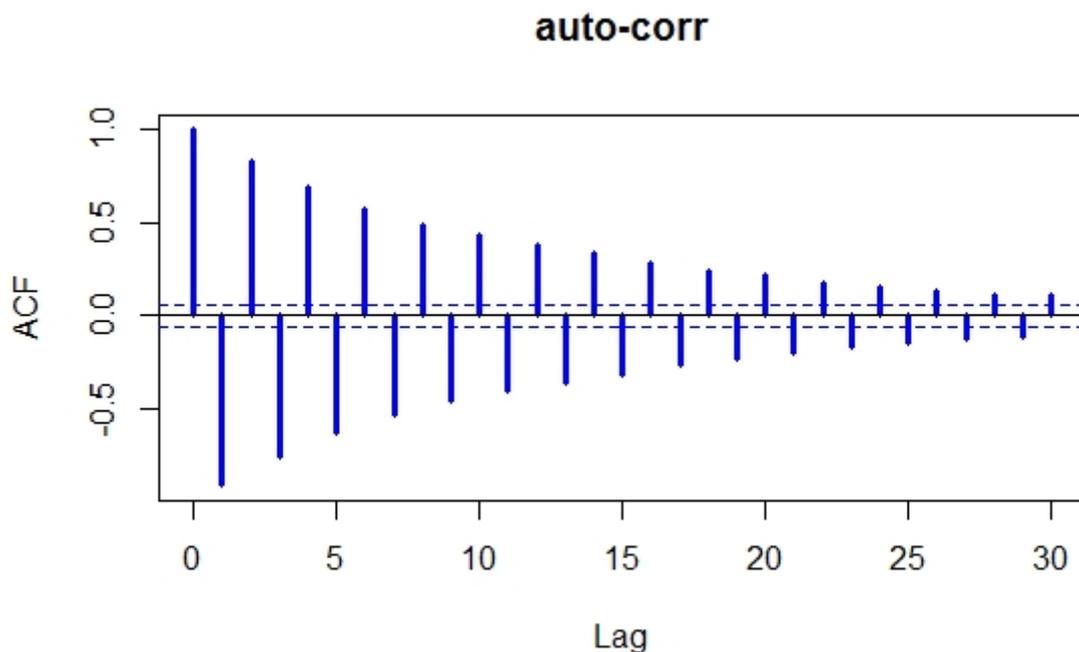


FIGURE 2.5 – Le corrélogramme simple du modèle AR(1) :  $\phi = -0.9$ .

Et son corrélogramme partiel est :

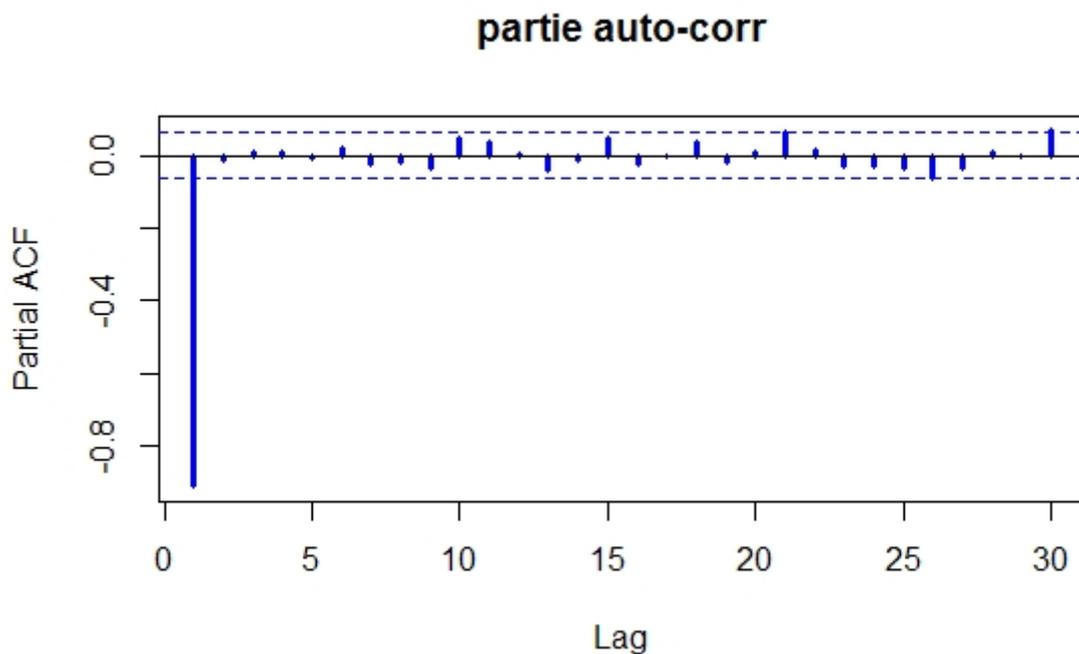


FIGURE 2.6 – Le corrélogramme partiel du modèle AR(1) :  $\phi = -0.9$ .

## 2.2.2 Processus moyenne mobile d'ordre $q$ : $MA(q)$

Le terme  $\varepsilon_t$  est souvent présenté comme un "choc" (une innovation pour les économètres), il rend compte d'un élément nouveau dans l'évolution du processus. On peut envisager que ces chocs (non-correlés et d'espérances nulles) aient des effets sur l'évolution du processus non seulement à la date à laquelle ils se produisent mais aussi à des dates ultérieures.

**Définition 2.5** Soit  $X_t$  un processus stochastique stationnaire, on dit que  $X_t$  est un processus moyenne mobile d'ordre  $q$  noté  $MA(q)$  le processus qui est générée par un bruit blanc vérifie l'équation suivante : (voir les références [2],[3])

$$\begin{aligned} X_t &= \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}, \\ &= \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i}. \end{aligned} \quad (2.14)$$

$\theta_i, i = 1, \dots, q$  sont les paramètres de modèle et  $\varepsilon_t$  un bruit blanc centrée de variance  $\sigma^2$ ,

On écrit habituellement :

$$X_t = \Theta(B)\varepsilon_t.$$

Où

$$\Theta(B) = \theta_0 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \text{ avec } \theta_0 = 1.$$

Et  $B$  est l'opérateur de décalage arrière (appelé aussi opérateur retard).

Le polynôme  $\Theta(B)$  est appelé polynôme caractéristique du processus moyenne mobile.

**Remarque 2.5** Certains textes et logiciels écrivent le modèle  $MA$  avec un coefficient positif qui est :

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}.$$

### 2.2.2.1 Représentation stationnaire

La définition d'un  $MA(q)$  est explicite et ne pose donc pas de problème : le processus  $X_t$  est parfaitement défini et est automatiquement stationnaire.

### 2.2.2.2 Représentation causale

La représentation est causale par définition.

### 2.2.2.3 Principales caractéristiques d'un modèle $MA(q)$

**La fonction d'autocorrélation :**

Le calcul de la fonction d'autocovariance d'un processus  $MA(q)$  est obtenu aussi :

$$\begin{aligned}\gamma_h &= E(X_t X_{t+h}) \\ &= E[(\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q})(\varepsilon_{t+h} - \theta_1 \varepsilon_{t+h-1} - \theta_2 \varepsilon_{t+h-2} - \dots - \theta_q \varepsilon_{t+h-q})].\end{aligned}$$

Et donc :

$$\gamma_h = \begin{cases} (-\theta_h + \theta_1 \theta_{h+1} + \dots + \theta_q \theta_{h+q}) \sigma_\varepsilon^2 & 0 < h \leq q. \\ 0 & \text{si } h > q. \end{cases} \quad (2.15)$$

La variance du processus  $X_t$  est donnée par la formule suivante :

$$\gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma_\varepsilon^2 = \text{var}(X_t).$$

Par conséquent, la fonction d'autocorrélation est définie telle que :

$$\rho_h = \frac{\gamma_h}{\gamma_0} = \begin{cases} 1 & \text{si } h = 0. \\ \frac{-\theta_h + \theta_1 \theta_{h+1} + \dots + \theta_q \theta_{q-h}}{1 + \theta_1^2 + \dots + \theta_q^2} & 1 \leq h \leq q. \\ 0 & h > q. \end{cases} \quad (2.16)$$

**Remarque 2.6** Nous constatons que la fonction d'autocorrélation  $\rho_h$  d'un processus  $MA(q)$  est égale à zéro quand  $h > q$ . Autrement dit la fonction d'autocorrélation d'un processus moyenne mobile chute à zéro au delà de  $q$ .

#### 2.2.2.4 Processus moyenne mobile $MA(1)$

On dit que la série  $X_t$  suit un processus de Moyenne Mobile  $MA(1)$  si elle est générée par un bruit blanc  $\varepsilon_t$  sous la forme :

$$\begin{aligned}X_t &= \varepsilon_t - \theta_1 \varepsilon_{t-1}. \\ X_t &= (1 - \theta_1 B) \varepsilon_t.\end{aligned}$$

Pour que cette équation soit stable, il faut que la racine du polynôme caractéristique  $1 - \theta_1 B = 0$  soit en dehors du cercle unité, c'est-à-dire que :

$$|B| = \left| \frac{1}{\theta_1} \right| > 1 \Leftrightarrow |\theta_1| < 1$$

#### Autocovariance :

Si on multiplie les deux membres de ce modèle par  $X_{t+h}$ , et si on calcule les espérances mathématiques, on obtient :

$$\begin{aligned}\gamma(h) &= \text{cov}(X_t, X_{t+h}), \\ &= E(X_t X_{t+h}) = E((\varepsilon_t - \theta_1 \varepsilon_{t-1})(\varepsilon_{t+h} - \theta_1 \varepsilon_{t+h-1})).\end{aligned}$$

— On peut voir que pour  $h = 0$ , en tenant compte que  $E(\varepsilon_t \varepsilon_{t'}) = 0$  pour  $t \neq t'$ , on a :

$$\gamma(0) = E(X_t^2) = E(\varepsilon_t^2 - 2\theta_1 \varepsilon_t \varepsilon_{t-1} + \theta_1^2 \varepsilon_{t-1}^2) = (1 + \theta_1^2) \sigma_\varepsilon^2.$$

— Pour  $h = 1$ ,

$$\begin{aligned}\gamma(1) &= E(X_t X_{t+1}) = E((\varepsilon_t - \theta_1 \varepsilon_{t-1})(\varepsilon_{t+1} - \theta_1 \varepsilon_t)), \\ &= -\theta_1 \sigma_\varepsilon^2.\end{aligned}$$

— Pour les valeurs de  $h > 1$ , on déduit que :

$$\begin{aligned}\gamma(h) &= E(X_t X_{t+h}) = E((\varepsilon_t - \theta_1 \varepsilon_{t-1})(\varepsilon_{t+h} - \theta_1 \varepsilon_{t+h-1})), \\ &= 0 \text{ pour } h > 1.\end{aligned}$$

**Autocorrélation de X :** De même, on peut montrer que :

$$\rho(h) = \begin{cases} 1 & \text{si } h = 0 \\ \frac{-\theta}{1+\theta^2} & \text{si } h = 1 \\ 0 & \text{si } h > 1 \end{cases} \quad (2.17)$$

**Autocorrélation partielle de X :**

$$\left\{ \begin{array}{l} \Gamma(0) = \rho(0) = 1 \\ \Gamma(1) = \rho(1) = \frac{-\theta}{1+\theta^2} \\ \Gamma(2) = \frac{\begin{vmatrix} 1 & \rho(1) \\ \rho(1) & \rho(2) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{vmatrix}} = \frac{\rho(2) - \rho(1)^2}{1 - \rho(1)^2} = \frac{-\left(\frac{-\theta}{1+\theta^2}\right)^2}{1 - \left(\frac{-\theta}{1+\theta^2}\right)^2} = \frac{-\theta^2}{1+\theta^2+\theta^4} \\ \Gamma(3) = \frac{\begin{vmatrix} 1 & \rho(1) & \rho(1) \\ \rho(1) & 1 & \rho(2) \\ \rho(2) & \rho(1) & \rho(3) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) & \rho(2) \\ \rho(1) & 1 & \rho(1) \\ \rho(2) & \rho(1) & 1 \end{vmatrix}} = \frac{-\theta^3}{1+\theta^2+\theta^4+\theta^6} \end{array} \right.$$

D'où

$$\Gamma(h) = \frac{-\theta^h}{\sum_{i=0}^h \theta^{2i}}. \quad (2.18)$$

La représentation graphique de  $MA(2)$ , prend la forme suivante :

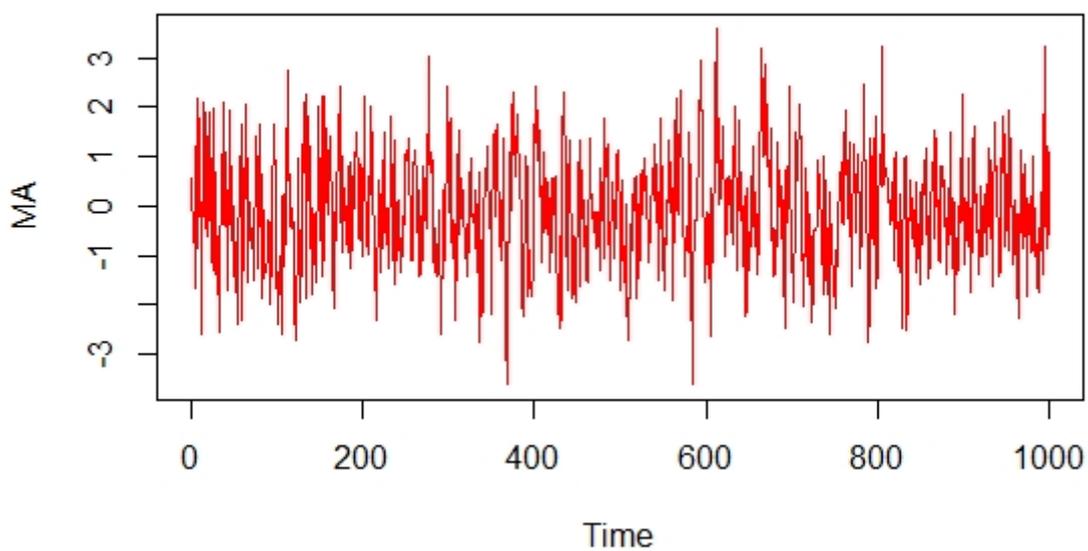


FIGURE 2.7 – Le graphe du modèle  $MA(2)$ .

Et

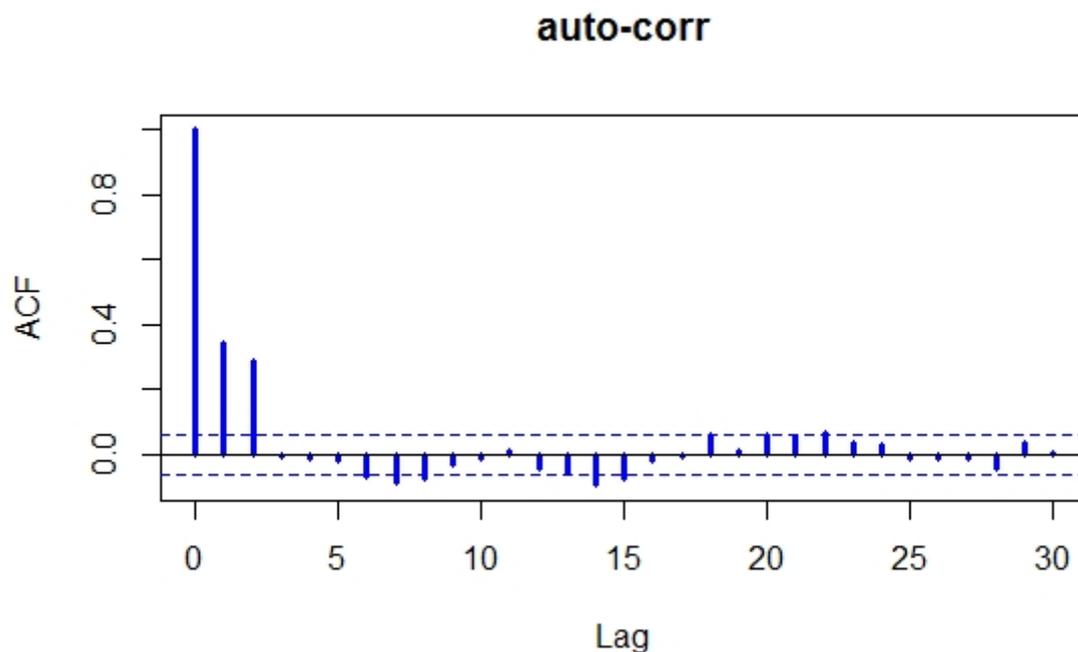


FIGURE 2.8 – Représentation graphique d'Autocorrélation simple du modèle  $MA(2)$ .

Et on a aussi

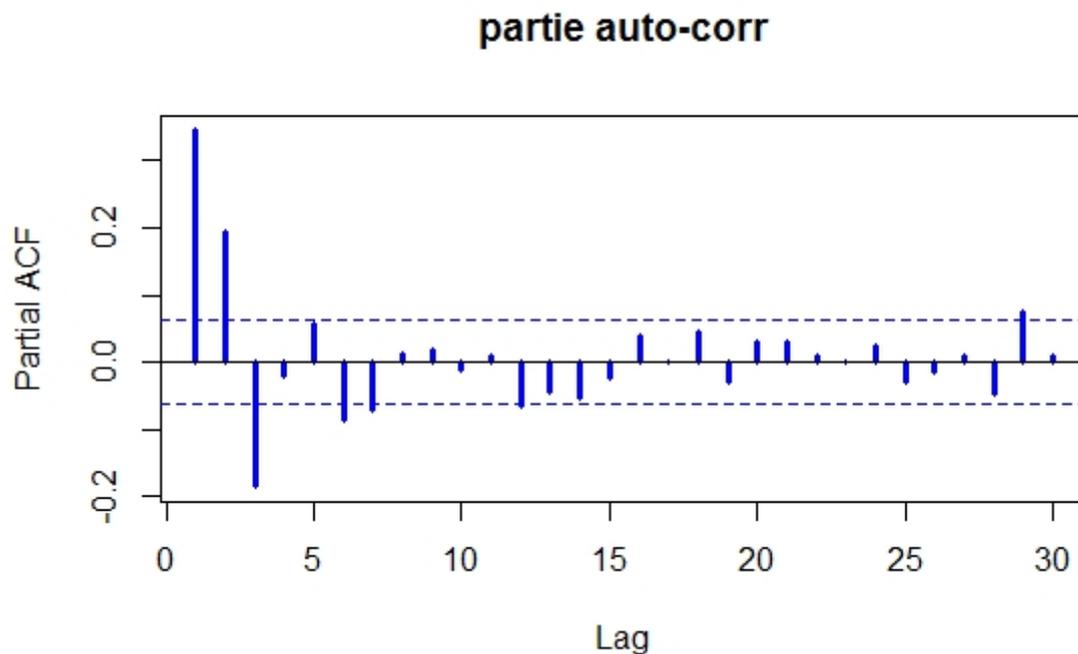


FIGURE 2.9 – Le corrélogramme partiel du modèle  $MA(2)$ .

On peut observer que la formule de l' $AR(\infty)$  découle de l'écriture :

$$X_t = (1 - \theta_1 B)\varepsilon_t \Leftrightarrow \varepsilon_t = X_t(1 - \theta_1 B)^{-1}.$$

On voit que si  $|\theta| < 1$ , on peut développer  $(1 - \theta_1 B)^{-1}$  en série entière :

$$(1 - \theta_1 B)^{-1} = 1 + \theta_1 B + \theta_1^2 B^2 + \dots$$

Donc :

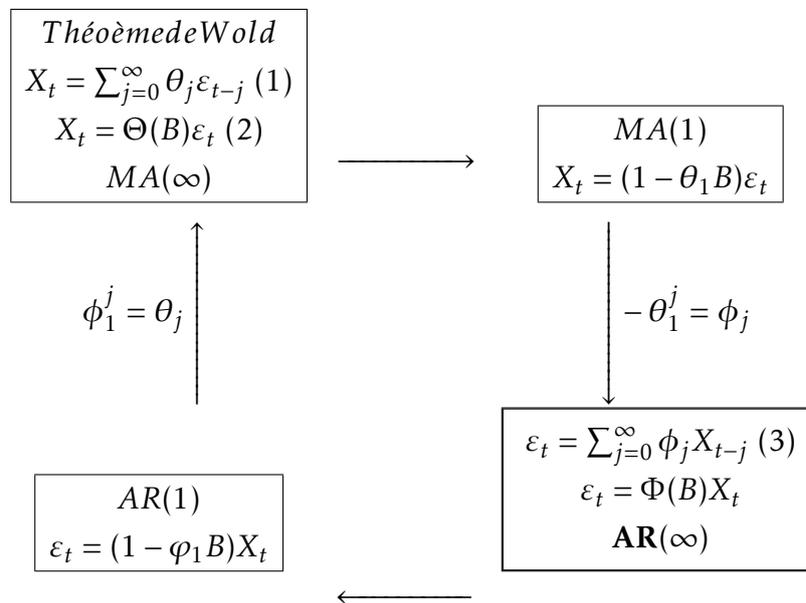
$$\varepsilon_t = X_t + \theta_1 X_{t-1} + \theta_1^2 X_{t-2} + \dots \Leftrightarrow X_t = \varepsilon_t - \theta_1 X_{t-1} - \theta_1^2 X_{t-2} - \dots$$

Est une Autorégression infinie avec  $\phi_j = -\theta_1^j$ .

D'après les résultats précédents on aura la correspondance :

$$AR(1) \Leftrightarrow MA(\infty)$$

$$MA(1) \Leftrightarrow AR(\infty)$$



**Remarque 2.7**

- Un processus  $AR$  est toujours inversible. Il est stationnaire lorsque les racines de  $\Phi(B)$  en module sont strictement supérieures à 1.
- Un processus  $MA$  est toujours stationnaire. Il est inversible si les racines de  $\Theta(B)$  sont en module strictement supérieures à 1.

**2.2.3 Processus ARMA(p,q) : Modèle autorégressif - Moyenne mobile**

On peut bien évidemment envisager de combiner les deux modèles précédents en introduisant :

- Une dépendance du processus vis-à-vis de son passé : modèle  $AR(p)$ ,
- Un effet retardé des chocs : modèle  $MA(q)$ .

Un tel modèle, appelé autorégressif - moyenne mobile ( $ARMA$ ), est caractérisé par le paramètre  $p$  de la partie autorégressive et le paramètre  $q$  de la partie moyenne mobile.

**Définition 2.6** Un modèle autorégressif moyenne mobile d'ordre  $(p, q)$  abrégé en  $ARMA(p, q)$  tout processus  $(X_t)_{t \in \mathbb{Z}}$  stationnaire vérifie l'équation : (voir la référence [22])

$$\begin{aligned} X_t &= +\varepsilon_t + \sum_{i=1}^p \phi_i X_{t-i} - \sum_{i=1}^q \theta_i \varepsilon_{t-i} & (2.19) \\ X_t - \sum_{i=1}^p \phi_i X_{t-i} &= \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i} \\ (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t &= (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \varepsilon_t \\ \Phi(B) X_t &= \Theta(B) \varepsilon_t. \end{aligned}$$

Où les  $\phi_i$  et  $\theta_i$  sont les paramètres du modèle et les  $\varepsilon_i$  les termes d'erreur.

**Remarque 2.8** Il est immédiat qu'un  $ARMA(p, 0)$  est un AR pur et qu'un  $ARMA(0, q)$  est un  $MA(q)$ . Les seuls processus admettant simultanément une représentation MA pure et une représentation AR pure correspondent au cas  $ARMA(0, 0)$ , c'est-à-dire aux bruits blancs.

### 2.2.3.1 Expression d'un $ARMA(p, q)$

#### 2.2.3.1.1 Représentation moyenne mobile infinie

i)- Si le polynôme  $\Phi$  a toutes ses racines de module différent de 1, l'opérateur  $\Phi(B)$  est inversible et la relation admet une solution stationnaire donnée par :

$$\forall t \in \mathbb{Z}, X_t = \frac{\Theta(B)}{\Phi(B)} \varepsilon_t = \sum_{i \in \mathbb{Z}} \psi_i \varepsilon_{t-i} \text{ avec } \sum_{i \in \mathbb{Z}} |\psi_i| < \infty. \quad (2.20)$$

C'est à dire sous une forme moyenne mobile infinie.

ii)- Si de plus les racines du polynôme  $\Phi$  sont de module strictement supérieur à 1, seules les valeurs présente et passées du bruit interviennent dans cette écriture  $MA(\infty)$ . Dans ce cas, les  $\psi_i$  de la représentation causale vérifient

$$\psi_k = \begin{cases} 1 & \text{si } k = 0. \\ -\theta_k + \sum_{i=1}^k \phi_i \psi_{k-i} & \text{pour } 1 \leq k \leq \max(p, q + 1). \\ \sum_{i=1}^p \phi_i \psi_{k-i} & \text{pour } k \geq \max(p, q + 1). \end{cases} \quad (2.21)$$

En posant  $\theta_0 = -1, \theta_i = 0$  si  $i > q$  et  $\phi_i = 0$  si  $i > p$ .

**Exemple 2.2** On considère  $X$  définie par :

$$X_t - X_{t-1} + \frac{1}{4} X_{t-2} = \varepsilon_t + \varepsilon_{t-1}.$$

Est un ARMA(2,1) avec :

$$\phi_1 = 1, \phi_2 = -\frac{1}{4} \text{ et } \theta_1 = -1.$$

En utilisant les équations précédentes, on obtient

$$\psi_0 = 1, \psi_1 = -\theta_1 + \phi_1 \psi_0 = 1 + 1 = 2.$$

Et

$$\psi_k = \sum_{i=1}^2 \phi_i \psi_{k-i} \Leftrightarrow \psi_k = \psi_{k-1} - \frac{1}{4} \psi_{k-2} \text{ pour } k \geq 2.$$

**2.2.3.1.2 Représentation autorégressive** Une démarche analogue peut-etre suivie pour le polynome  $\Theta(B)$  :

- i)- Supposons que le polynome  $\Theta(B)$  a toutes ses racines de module différent de 1, l'opérateur  $\Theta(B)$  est inversible et on obtient la forme autorégressive infinie :

$$\forall t \in \mathbb{Z}, \varepsilon_t = \frac{\Theta(B)}{\Phi(B)} X_t = \sum_{i \in \mathbb{Z}} \pi_i X_{t-i} \text{ avec } \sum_{i \in \mathbb{Z}} |\pi_i| < \infty. \quad (2.22)$$

- ii)- Si de plus les racines du polynôme  $\Theta$  sont de module strictement supérieur à 1, cette représentation  $AR(\infty)$  ne fait intervenir que les valeurs présentes et passées du processus. Dans ce cas, les  $\pi_i$  de la représentation causale vérifiant :

$$\pi_k = \begin{cases} 1 & \text{si } k = 0. \\ -\phi_k + \sum_{i=1}^k \theta_i \pi_{k-i} & \text{pour } 1 \leq k \leq \max(p+1, q). \\ \sum_{i=1}^q \theta_i \pi_{k-i} & \text{pour } k \geq \max(p+1, q). \end{cases} \quad (2.23)$$

**Remarque 2.9** Lorsque  $\Phi$  et  $\Theta$  ont des racines de module strictement supérieur à 1, les deux représentations ci-dessus sont vérifiées et  $\varepsilon_t$  est non corrélé avec  $X_{t-1}, X_{t-2}, \dots$  et représente l'innovation du processus  $X_t$  à la date  $t$ .

### 2.2.3.2 Principales caractéristiques d'un modèle ARMA(p;q)

**La fonction d'autocorrélation :**

La covariance vaut :

$$\begin{aligned}
 \gamma(h) &= E(X_t X_{t+h}), & (2.24) \\
 &= E\left[\left(\sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t - \sum_{j=1}^q \theta_j \varepsilon_{t-j}\right) X_{t+h}\right], \\
 &= \sum_{i=1}^p \phi_i E(X_{t-i} X_{t+h}) - \sum_{j=1}^q \theta_j E(\varepsilon_{t-j} X_{t+h}) + E(\varepsilon_t X_{t+h}).
 \end{aligned}$$

Si on note :

$$\gamma_{X_\varepsilon}(h) = E(\varepsilon_t X_{t+h}).$$

La covariance croisée entre  $\varepsilon_t$  et  $X_{t+h}$  On a :

$$\gamma(h) = \sum_{i=1}^p \phi_i \gamma(h-i) - \sum_{j=1}^q \theta_j \gamma_{X_\varepsilon}(h-j) + \gamma_{X_\varepsilon}(h).$$

Or

$$E(X_t \varepsilon_t) = E\left[\left(\sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t - \sum_{j=1}^q \theta_j \varepsilon_{t-j}\right) \varepsilon_t\right] = E(\varepsilon_t^2) = \sigma^2.$$

Et  $\varepsilon_t$  n'est pas corrélé avec le passé :

$$\gamma_{X_\varepsilon}(h) = \begin{cases} \sigma^2 & \text{si } h = 0 \\ 0 & \text{si } h > 0 \end{cases}, \text{ et } \gamma_{X_\varepsilon}(h) \neq 0 \text{ si } h < 0.$$

Si  $h > q$ , tous les  $\gamma_{X_\varepsilon}(h) = 0$  et on a :

$$\begin{cases} \gamma(h) = \sum_{i=1}^p \phi_i \gamma(h-i). \\ \rho(h) = \sum_{i=1}^p \phi_i \rho(h-i). \end{cases}$$

Ce qui est un système d'équation de type **Yule-Walker**. En effet, la partie *MA* ne joue que jusqu'au délai  $q$ , après elle ne joue plus. Il est donc normal de retrouver des équations ressemblant à celles trouvées dans les *AR(p)* à partir de  $h > q$ .

Pour un *ARMA(p, q)* la structure d'auto-corrélation ne suit pas un schéma connu jusqu'au délai  $q$  mais ensuite le comportement est le même que celui d'un *AR(p)*.

Il y a donc une infinité de  $\gamma(h)$  non-nuls, mais à partir de  $h > q$ ; la fonction d'autocorrélation a le même comportement que pour *AR(p)*

**Remarque 2.10** La fonction d'autocorrélation  $\rho_h$  d'un processus *ARMA(p, q)* décroît exponentiellement vers zéro.

#### La fonction d'autocorrélation partielle :

Pour  $h \leq p$ , il n'y a pas de schéma bien précis.

Pour  $h > p$ , on retrouve le même comportement que pour un *MA(q)*.

Le tableau suivant illustre les propriétés des fonctions d'autocorrélation simple et partielle du modèle *ARMA*

TABLE 2.1 – Les propriétés des fonctions d'autocorrélation simple et partielle du modèle *ARMA*

Type de modèle	Configuration typique d'une <i>ACF</i>	Configuration typique d'une <i>PACF</i>
<i>ARMA</i> (1, 1)	Décroissance géométrique à partir du premier retards, le signe est déterminé par $\theta_1 - \phi_1$	Décroissance exponentielle $\phi_1 > 0$ ou sinusoidale amortie $\phi_1 < 0$
<i>ARMA</i> ( <i>p</i> , <i>q</i> )	Décroissance exponentielle ou sinusoidale amortie tronquée après ( <i>q</i> - <i>p</i> )retards.	Décroissance exponentielle ou sinusoidale amortie tronquée après ( <i>q</i> - <i>p</i> )retards.

## 2.2.4 Processus SARMA

**Définition 2.7** Le processus  $(X_t)_{t \in \mathbb{Z}}$  satisfait une représentation d'un processus autorégressif moyenne mobile saisonnier, noté SARMA, si et seulement si :

$$\sum_{i=0}^p \phi_{is} X_{t-is} = c + \sum_{i=0}^q \theta_{is'} \varepsilon_{t-is'}. \quad (2.25)$$

Où  $c \in \mathbb{R}$  et  $(\varepsilon_t)_{t \in \mathbb{Z}}$  sont *i.i.d.*  $N(0, \sigma^2)$ . Ici, *s* désigne la période de saisonnalité de la composante AR, alors que *s'* représente la période de saisonnalité de la composante MA.

## 2.3 Modèles linéaires non-stationnaires

À la section précédente, quelques processus aléatoires stationnaires ont été présentés et étudiés. Toutefois, dans la réalité, il arrive souvent que des séries observées montrent un comportement non stationnaire. La classe des processus non stationnaire est relativement vaste, il existe différents types de non stationnarité, on présente deux classes des processus non stationnaires : les processus *TS* et *DS*.

### 2.3.1 Processus TS

Un processus  $(X_t)_{t \in \mathbb{Z}}$  est un processus non stationnaire de type (*TS*) s'il peut s'écrire sous la forme :

$$X_t = f(t) + \varepsilon_t. \quad (2.26)$$

*f* : Est une fonction polynômiale du temps, linéaire ou non linéaire.

$\varepsilon_t$  : Est un processus stationnaire.

Le processus *TS* est le plus simple (et le plus répandu) est représenté par une fonction

polynômiale de degré 1. Le processus  $TS$  porte alors le nom de linéaire et s'écrit :

$$X_t = a_0 + a_1 t + \varepsilon_t.$$

Si  $\varepsilon_t$  est un bruit blanc (gaussien ou non), les caractéristiques de ces processus sont alors :

$$\begin{cases} E(X_t) = a_0 + a_1 t + E(\varepsilon_t) = a_0 + a_1 t. \\ \text{var}(X_t) = \sigma_\varepsilon^2. \\ \text{cov}(X_t, X_s) = 0, \text{ pour } t \neq s. \end{cases}$$

Ce processus  $TS$  est non stationnaire car  $E(X_t)$  dépend du temps. La chronique retrouve son mouvement de long terme qui est ici la droite de la tendance.

Pour stationnariser un processus de type  $TS$  on estime  $f(t)$  par les moindres carrés ordinaire

### 2.3.2 Processus DS

$(X_t)$  est un processus non stationnaire de type  $DS$  ou intégré d'ordre  $d$ , noté  $I(d)$ , si la processus obtenu après " $d$ " différenciation est stationnaire.

$$Y_t = (1 - B)^d X_t = \Delta^d X_t. \quad (2.27)$$

Donc, la définition des processus  $DS$  repose sur la présence des racines unitaires.

**Définition 2.8** Les processus  $DS$  sont des processus que l'on peut rendre stationnaires par l'utilisation d'un filtre aux différences :

$$(1 - B)X_t = \beta + \varepsilon_t. \quad (2.28)$$

Où  $\varepsilon_t$  est un processus stationnaire, et  $\beta$  une constante réelle.

Le processus de premier ordre s'écrit :

$$(1 - B)X_t = \beta + \varepsilon_t \iff X_t = X_{t-1} + \beta + \varepsilon_t.$$

### 2.3.3 Modèle ARIMA(processus autorégressif moyenne mobile intégré)

La classe des modèles non stationnaires est très utile dans le domaine de l'économétrie et les mathématiques financières. Ce genre de modèle permet de représenter des processus présentant un certain type de non-stationnarité que l'on peut éliminer par différenciation d'ordre  $d$ .

On dit qu'un processus  $X_t$  est un *ARIMA* (Autoregressive Integrated Moving Average) d'ordre  $(p, d, q)$  si  $\nabla^d X_t$  est un processus *ARMA* $(p, q)$ . (voir les références [17],[3])

On modélise alors le processus  $X_t$  sous la forme :

$$\Phi(B)(1 - B)^d X_t = \Theta(B)\varepsilon_t \quad (2.29)$$

Où le polynôme  $\Phi(B)$  est de degré  $p$  et le polynôme  $\Theta(B)$  est de degré  $q$ .

On écrit que la série  $X_t$  suit un processus *ARIMA* $(p, d, q)$ .

**Remarque 2.11 :**

Les polynômes  $\Phi(B)$  et  $\Theta(B)$  sont de degrés  $p$  et  $q$  respectivement et  $\Phi(B) \neq 0$  pour  $|B| \leq 1$ , alors que le polynôme  $\Phi(B)$  admet une racine multiple  $B = 1$ .

### 2.3.4 Modèle SARIMA

Le processus *SARIMA* est un modèle *ARIMA* dans lequel une composante saisonnière est ajoutée. Ainsi,  $(X_t)_{t \geq 0}$  est un processus *SARIMA* $(p, d, q)$  de périodes  $s$  et  $s'$  si et seulement si :

$$\phi(B^s)(1 - B^s)^d X_t = \theta(B^{s'})\varepsilon_t. \quad (2.30)$$

Où

$$\phi(B) = \sum_{j=0}^p \phi_j B^j.$$

Et

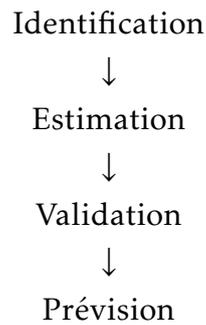
$$\theta(B) = \sum_{j=0}^q \theta_j B^j.$$

**Remarque 2.12** La période de saisonnalité  $s$  est associée à la composante *AR*, alors que la période de saisonnalité  $s'$  est associée à la composante *MA*. On suppose que toutes les racines des polynômes  $\Phi(B)$  et  $\Theta(B)$  sont situées à l'extérieur du cercle unité.

## 2.4 METHODE DE BOX-JENKINS

La méthodologie de Box-Jenkins peut se décomposer en quatre étapes :

Nous présenterons tout d'abord, l'étape de transformation des données afin de stabiliser la variance (log, sqrt,...) et différenciation des données pour les stationariser, puis l'étape de **l'identification**, ensuite nous jetterons la lumière sur la phase de **l'estimation**, puis nous présenterons des tests de diagnostic dans l'étape de **validation**, enfin la dernière étape consiste à utiliser le modèle *ARIMA* validé à des fins de **prévision**. (voir les références [20], [4])



### 2.4.1 Identification

Après avoir transformé la série étudiée de manière à la stationnariser ce qui est déjà vu; on arrive à l'étape de l'identification qui est une étape délicate qui conditionne la prévision de la chronique, elle consiste à déterminer les paramètres  $p$  et  $q$  du modèle  $ARMA$  à l'aide de la visualisation de la fonction d'autocorrélation simple et la fonction d'autocorrélation partiel. On a vu que la  $PACF$  d'un  $AR(p)$  est nulle à partir de l'ordre  $p + 1$  et l' $ACF$  d'un  $MA(q)$  est nulle à partir de l'ordre  $q + 1$ . D'autre part, l' $ACF$  a une décroissance exponentielle pour un  $AR(p)$ , la  $PACF$  a une décroissance exponentielle pour un  $MA(q)$ ; enfin ces deux fonctions ont une décroissance exponentielle pour un  $ARMA(p, q)$  dont les deux ordres sont non nuls. Ces comportements sont récapitulés dans le tableau suivant :

TABLE 2.2 – Les propriétés des fonctions d'autocorrélation simple et partielle des différence modèles

Fonction	$AR(p)$	$MA(q)$	$ARMA(p, q)$	bruit blanc
$ACF$	<i>Exp</i>	<i>Nulle(q)</i>	<i>Exp</i>	0
$PACF$	<i>Nulle(p)</i>	<i>Exp</i>	<i>Exp</i>	0

- Soit un processus stationnaire  $(X_t)_{t \in \mathbb{Z}}$  satisfaisant une représentation  $MA(q)$ , pour déterminer la valeur de  $q$ , on se base sur la fonction d'autocorrélation du processus  $MA$ . En fait,  $q$  correspondra au plus grand délai tel que l'autocorrélation n'est pas statistiquement égale à 0.
- Soit un processus stationnaire  $(X_t)_{t \in \mathbb{Z}}$  satisfaisant une représentation  $AR(p)$ . Globalement, l'identification d'un processus  $AR(p)$  s'effectue de la même façon que celle d'un processus  $MA(q)$ . La seule différence réside dans le fait que c'est l'autocorrélation partielle, plutôt que l'autocorrélation proprement dite, qui est utilisée.
- Il existe plusieurs méthodes pour déterminer les ordres  $p$  et  $q$  d'un processus  $ARMA$ . L'une d'elles est la méthode du coin élaborée par Beghin et al(1980). On peut également se baser sur les autocorrélations et les autocorrélations partielles. Pour obtenir l'ordre de la composante  $MA$ , il faut identifier l'autocorrélation significative dont l'ordre est le plus élevé; pour la composante  $AR$ , il faut identifier l'autocorrélation partielle significative dont l'ordre est le plus élevé.

## 2.4.2 Estimation

L'estimation des paramètres d'un modèle  $ARIMA(p, d, q)$  lorsque  $p, d, q$  sont supposés connus peut se réaliser par différentes méthodes dans le domaine temporel, et parmi ces méthodes on a :

1. **Maximum de vraisemblance** : Une méthode populaire pour estimer les paramètres d'un modèle est le maximum de vraisemblance. La fonction de vraisemblance associée à un échantillon  $X_1, \dots, X_T$  *i.i.d.* d'une loi dont la densité est  $f(x|\theta)$ , avec  $\theta = (\theta_1, \dots, \theta_k) \in \mathbb{R}^k$ ,

Est définie par :

$$L(\theta) = \prod_{t=1}^T f(X_t|\theta). \quad (2.31)$$

L'estimateur du maximum de vraisemblance (**EMV**) est la valeur  $\widehat{\theta}_{EMV}$  qui maximise  $L(\theta)$ . Parfois, il est possible de déduire cet estimateur en dérivant  $L(\theta)$  par rapport à chacun des paramètres  $(\theta_1, \dots, \theta_k)$  et de résoudre le système à  $k$  équations :

$$\frac{\partial L(\theta)}{\partial \theta_j} = 0 \text{ où } j = 1, \dots, k.$$

1. Dans le cas  $q = 0$ , c'est à dire dans le cas d'un  $AR(p)$ , on utilise les équations de **Yule Walker**.

### 2.4.2.1 Critères de choix des modèles

Souvent il n'est pas facile de déterminer un modèle unique. Le modèle qui est finalement choisi est celui qui minimise l'un des critères à partir de  $T$  observations.

#### 2.4.2.1.1 Critère standard

1. L'erreur absolue moyenne (Mean Absolute Error) :

$$MAE = \frac{1}{T} \sum_{t=1}^T |\varepsilon_t|.$$

2. L'erreur quadratique moyenne (Mean Squared Error) :

$$MSE = \frac{1}{T} \sum_{t=1}^T \varepsilon_t^2.$$

3. La racine carrée de l'erreur quadratique moyenne (Root Mean Square Error) :

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T \varepsilon_t^2}.$$

4. Ecart absolu moyen en pourcentage (Mean Absolute Percent Error) :

$$MAPE = \frac{100}{T} \sum_{t=1}^T \left| \frac{\varepsilon_t}{X_t} \right|.$$

Plus la valeur de ces critères est faible, plus le modèle estimé est proche des observations.

#### 2.4.2.1.2 Critère d'information

1. Akaike (1969) : (voir la référence [1])

$$AIC(p, q) = \log(\widehat{\sigma}_\varepsilon^2) + 2\frac{p+q}{T}. \quad (2.32)$$

2. Schwarz (1977) : (voir la référence [18])

$$BIC(p, q) = \log(\widehat{\sigma}_\varepsilon^2) + (p+q)\frac{\log T}{T}. \quad (2.33)$$

3. Hannan\_Quinn(1979) : (voir la référence [14])

$$\varphi(p, q) = \log(\widehat{\sigma}_\varepsilon^2) + (p+q)c\frac{\log(\log T)}{T}, \text{ avec } c > 2.$$

### 2.4.3 Adéquation et validation du modèle

Cette étape consiste à s'assurer que le modèle retenu décrit bien la série d'observations. Il s'agit de vérifier les conditions d'un processus bruit blanc à l'aide de tests appropriés et que les paramètres estimés sont statistiquement différents de 0.

#### Test de significativité des coefficients

Les coefficients du modèle doivent être significativement différent de zéro, pour ce faire on utilise le test classique de Student.

#### Test sur la moyenne des résidus

Le test sur la moyenne des résidus fait référence à la première condition d'un processus bruit blanc, à savoir une moyenne égale à 0. Ainsi, pour vérifier que les résidus d'un certain modèle sont de moyenne nulle, comme il se doit, il s'agit de confronter les hypothèses :

$$\begin{cases} H_0 : E(\varepsilon_t) = 0. \\ H_1 : E(\varepsilon_t) \neq 0. \end{cases}$$

Une statistique de test est définie par :

$$\xi = \frac{\bar{\varepsilon}}{S_\varepsilon/\sqrt{T}}.$$

Où

$$\bar{\varepsilon} = \frac{1}{T} \sum_{t=1}^T \varepsilon_t \quad \text{et} \quad S_\varepsilon = \frac{1}{T} \sum_{t=1}^T \varepsilon_t^2.$$

La statistique  $\xi$  est approximativement distribuée selon la loi normale centrée réduite. Ainsi, on rejette  $H_0$  si  $|\xi| > z_{\alpha/2}$  où  $z_{\alpha/2}$  est la fractile d'ordre  $\alpha/2$  de la distribution normale centrée réduite.

### Test sur la fonction d'autocorrélation

Le test de Box Pierce permet d'identifier les processus de bruit blanc. Ce test est :

$$\begin{cases} H_0 : \rho(1) = \rho(2) = \dots = \rho(h) = 0. \\ H_1 : \text{il existe au moins un } \rho(i) \text{ significativement différent de } 0. \end{cases}$$

Pour effectuer ce test, on recourt à la statistique  $Q$  qui est donnée par :

$$Q = T \sum_{k=1}^h \widehat{\rho}^2(k).$$

$h$  : nombre de retards,  $T$  : nombre d'observations,  $\widehat{\rho}(k)$  : autocorrélation empirique d'ordre  $k$ .

La statistique  $Q$  est distribuée de manière asymptotique comme un  $\chi^2$  à  $h$  degrés de liberté.

Nous rejetons donc l'hypothèse de bruit blanc, au seuil  $\alpha$ , si la statistique  $Q$  est supérieure au  $\chi^2$  lu dans la table au seuil  $(1 - \alpha)$  et  $h$  degrés de liberté.

### Test de normalité des résidus

Plusieurs des modèles de séries chronologiques supposent que les résidus sont indépendants et distribués selon la loi Normale. Un des tests permettant de vérifier la normalité des résidus est celui de Jarque & Bera. Les hypothèses à confronter sont :

$$\begin{aligned} H_0 &\sim \mathcal{N}(0, 1). \\ H_1 &\sim \mathcal{N}(0, 1). \end{aligned}$$

Avant de décrire le test pour les résidus d'un modèle de séries chronologiques, prenons le cas de  $n$  observations  $X_1, X_1, \dots, X_n$  indépendantes et de même loi. Dans ce cas, la statistique du test de Jarque & Bera, de loi asymptotique khi-carré à deux degrés de liberté, est définie par :

$$JB = \frac{n}{6} \beta_1^2 + \frac{n}{24} (\beta_2 - 3)^2.$$

$$\beta_1 = \frac{1}{n} \sum_{i=1}^n \left( \frac{X_i - \bar{X}}{S} \right)^3 \quad \text{et} \quad \beta_2 = \frac{1}{n} \sum_{i=1}^n \left( \frac{X_i - \bar{X}}{S} \right)^4.$$

Sont respectivement les coefficients d'asymétrie et d'aplatissement. Ici,  $X$  et  $S$  sont respectivement la moyenne et l'écart type empiriques. Sous l'hypothèse de normalité, on

peut montrer que :

$$\frac{\sqrt{n}\beta_1}{\sqrt{6}} \sim \mathcal{N}(0,1) \text{ et } \frac{\sqrt{n}(\beta_2 - 3)}{\sqrt{24}} \sim \mathcal{N}(0,1).$$

Et que :

$$\frac{n\beta_1^2}{6} \sim \chi_1^2 \text{ et } \frac{n(\beta_2 - 3)^2}{24} \sim \chi_1^2.$$

Où  $\chi_v^2$  représente la loi khi-carré à degrés de liberté  $v$ . Ainsi, la loi asymptotique de  $JB$  est la distribution khi-carré à deux degrés de liberté. On rejette donc l'hypothèse de normalité, si  $JB > \chi_{1-\alpha,2}^2$ .

#### 2.4.4 La prévision

C'est la dernière étape de la méthodologie de Box and Jenkins.

Etant donné une série stationnaire, observée entre 1 et  $T$ , on cherche à faire de la prévision à horizon  $h$ , et donc de prévoir  $X_{T+1}, \dots, X_{T+h}$ .

Il s'agit de calculer les prévisions optimales du modèle *ARIMA* estimé, à savoir  $X_T(h)$  la prévision de  $X_{T+h}$  sachant l'ensemble d'information disponible en  $T$ , noté :

$$\widehat{X}_{T+h} = E(X_{T+h}/X_T, X_{T-1}, \dots, X_1). \quad (2.34)$$

##### Principe de prévision

Soit  $\{X_t\}$  un processus aléatoire du second ordre défini sur  $(\Omega, P)$ . Le problème le plus couramment rencontré est la prévision des valeurs futures de ce processus. Prédire une valeur future  $X_{T+h}$  de  $\{X_t\}$  au vu d'une observation  $X = (X_1, X_2, \dots, X_T)$ , revient à chercher une fonction déterministe  $f : \mathbb{C}^h \rightarrow \mathbb{C}$  telle que  $\widehat{X}_{T+h} = f(X_1, X_2, \dots, X_T)$ .

La prédiction optimale au sens des moindres carrés consiste à choisir  $f$  de telle sorte que la moyenne du carré de l'erreur de prévision  $E \left[ |X_{T+h} - \widehat{X}_{T+h}|^2 \right]$  soit minimale.

Dans l'espace de Hilbert  $H_T = [X_1, X_2, \dots, X_T]$  engendré par :  $\{X_1, X_2, \dots, X_T\}$ , ce problème classique admet une solution unique.

$$\widehat{X}_{T+h} = E^{H_T}(X_{T+h}) = E(X_{T+h}/X_T, X_{T-1}, \dots, X_1) = f(X_1, X_2, \dots, X_T).$$

En d'autres termes,  $\widehat{X}_{T+h}$  est l'espérance conditionnelle de  $X_{T+h}$  lorsque l'on observe  $X_1, X_2, \dots, X_T$ .

En pratique la fonction  $f$  est difficile à déterminer, soit parce que la loi jointe de  $(X_1, X_2, \dots, X_T, X_{T+h})$  est mal connue, soit parce qu'elle ne se prête pas aux calculs explicites. Ceci nous amène à restreindre la classe des estimateurs considérés. Le principe reste toujours le même, *ie* on cherchera à minimiser  $E \left[ |X_{T+h} - \widehat{X}_{T+h}|^2 \right]$  où  $\widehat{X}_{T+h} = f(X_1, X_2, \dots, X_T)$ . Mais la fonction  $f$  sera déterminée parmi les fonctions linéaires. L'unique solution est évidemment donnée par  $\widehat{X}_{T+h} = \text{proj}(X_{T+h})$  où  $\text{proj}$  est la projection orthogonale sur le sous-espace vectoriel  $H_T$  engendré par  $X_1, X_2, \dots, X_T$  dans  $L^2(\Omega; P)$ .

Le principe de la prédiction linéaire revient à considérer que l'information fournie par l'observation  $X$  du processus ne dépend que de l'espace  $H_T$ , sous-espace vectoriel fermé de  $L^2(\Omega; P)$  engendré par les  $X$ . Le sous-espace  $H_T$  est souvent appelé l'enveloppe linéaire de  $X$ . Cela équivaut à ne considérer comme accessibles au calcul que les combinaisons linéaires des  $X_t$ ,  $t = 1, 2, \dots, T$  observés.

#### 2.4.4.1 Prédiction d'un modèle AR(p)

Le modèle s'écrit :

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t.$$

La prédiction optimale à la date  $T + 1$ , faite à la date  $T$  est

$$\widehat{X}_{T+1} = E(X_{T+1}/X_T, X_{T-1}, \dots, X_1).$$

Alors :

$$\widehat{X}_{T+1} = \phi_1 X_T + \dots + \phi_p X_{T+1-p}.$$

Pour prédire  $X_{T+2}$  à partir de  $X_1, \dots, X_T$  on a :

$$\widehat{X}_{T+2} = \phi_1 \widehat{X}_{T+1} + \dots + \phi_p X_{T+2-p}.$$

De façon générale

$$\widehat{X}_{T+h} = \phi_1 \widehat{X}_{T+h-1} + \dots + \phi_p X_{T+h-p}. \quad (2.35)$$

De façon récursive

$$\widehat{X}_{T+h} = \begin{cases} \phi_1 \widehat{X}_{T+h-1} + \dots + \phi_{h-1} \widehat{X}_{T+1} + \phi_h X_T + \dots + \phi_p X_{T+h-p} & \text{pour } h \leq p. \\ \phi_1 \widehat{X}_{T+h-1} + \dots + \phi_p \widehat{X}_{T+h-p} & \text{pour } h > p. \end{cases}$$

**Remarque 2.13** : Dans le cas des modèles d'ordre 1, on a  $\widehat{X}_{T+1} = \phi_1 X_T$  et  $\widehat{X}_{T+2} = \phi_1 \widehat{X}_{T+1} = \phi_1^2 X_T$ . On vérifie aisément par récurrence que :

$$\widehat{X}_{T+h} = \phi_1^h X_T.$$

Donc en particulier que  $\widehat{X}_{T+h} \rightarrow 0$  quand  $h$  tend vers l'infini.

#### 2.4.4.2 Prédiction d'un modèle MA(q)

Soit le modèle MA(q)

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} = \Theta(B)\varepsilon_t.$$

La prédiction optimale à la date  $T + 1$  est :

$$\widehat{X}_{T+1} = E(X_{T+1}/X_T, X_{T-1}, \dots, X_1).$$

Donc

$$\widehat{X}_{T+1} = \theta_1 \varepsilon_T + \dots + \theta_q \varepsilon_{T+1-q}.$$

De façons analogue

$$\widehat{X}_{T+h} = E(X_{T+h}/X_T, X_{T-1}, \dots, X_1).$$

$$\widehat{X}_{T+h} = \begin{cases} \theta_h \varepsilon_T + \dots + \theta_q \varepsilon_{T+h-q} & \text{pour } h \leq q. \\ 0 & \text{pour } h > q. \end{cases} \quad (2.36)$$

### 2.4.4.3 Prévision d'un modèle ARMA(p,q)

$X_t$  est un ARMA(p,q), alors :

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t - \sum_{j=1}^q \theta_j \varepsilon_{t-j}.$$

Donc

$$X_{T+h} = \sum_{i=1}^p \phi_i X_{T+h-i} + \varepsilon_{T+h} - \sum_{j=1}^q \theta_j \varepsilon_{T+h-j}.$$

On a alors :

$$\widehat{X}_{T+h} = \sum_{i=1}^p \phi_i \widehat{X}_{T+h-i} - \sum_{j=1}^q \theta_j \widehat{\varepsilon}_{T+h-j}. \quad (2.37)$$

Avec

$$\widehat{\varepsilon}_{T+h-j} = \begin{cases} 0 & \text{si } h > j. \\ \varepsilon_{T+h-j} & \text{si } h \leq j. \end{cases}$$

### 2.4.4.4 Prévisions dans le cas d'un processus ARIMA(p,d,q)

On considérons ici  $(X_t)$  satisfaisant une équation de la forme :

$$\Phi(B)(1-B)^d X_t = \Theta(B)\varepsilon_t.$$

Posons alors  $\Psi(B) = \Phi(B)(1-B)^d$  : La forme ARIMA(p,d,q) peut s'écrire :

$$X_t = \sum_{i=1}^{p+d} \psi_i X_{t-i} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j}.$$

Et donc :

$$X_{t+h} = \sum_{i=1}^{p+d} \psi_i X_{t+h-i} + \varepsilon_{t+h} + \sum_{j=1}^q \theta_j \varepsilon_{t+h-j}.$$

Notons  $\widehat{X}_{T+h}$  la prévision faite à la date  $T+h$

$$\widehat{X}_{T+h} = E(X_{T+h}/X_T, X_{T-1}, \dots, X_1).$$

Alors

$$\widehat{X}_{T+h} = \sum_{i=1}^{p+d} \psi_i \widehat{X}_{T+h-i} + \sum_{j=1}^q \theta_j \widehat{\varepsilon}_{T+h-j} \quad (2.38)$$

Où

$$\widehat{X}_{T+h-i} = X_{T+h-i} \text{ pour } i \geq h.$$

Et

$$\widehat{\varepsilon}_{T+h-j} = \begin{cases} 0 & \text{si } j < h. \\ \varepsilon_{T+h-j} & \text{si } j \geq h. \end{cases}$$

En particulier, pour  $h \geq q$ , on obtient une relation de récurrence de la forme :

$$\widehat{X}_{T+h} = \sum_{i=1}^{p+d} \psi_i \widehat{X}_{T+h-i}.$$

### 3.1 Introduction

*ARIMA* (autoregressive integrated moving average) est une technique couramment utilisée pour ajuster les données de séries chronologiques et la prévision. C'est une version généralisée du processus *ARMA* (autoregressive moving average), où le processus *ARMA* est appliqué à une version différenciée des données plutôt qu'à l'original.

Trois nombres  $p$ ,  $d$  et  $q$  spécifient le modèle *ARIMA* et l'*ARIMA* est dit d'ordre  $(p, d, q)$ . Ici  $p$ ,  $d$  et  $q$  sont l'ordre de la partie *AR*, de la *différence* et de la partie *MA* respectivement.

*AR* et *MA* sont deux techniques différentes pour la collecte de données chronologiques stationnaires. *ARMA* (et *ARIMA*) est une combinaison de ces deux méthodes pour un ajustement parfait du modèle.

Dans cet article, un aperçu des processus *AR* et *MA* sera donné les étapes de la construction d'un modèle *ARIMA* seront expliquées. Enfin, une démonstration utilisant *R* sera présentée.

### 3.2 Simulation

Pour les simulations on a utilisé le langage *R* et plus particulièrement les packages `timeSeries`, `tseries` et `forecast`.

#### 3.2.1 Simuler une série chronologique

Créer un objet de type série temporelle :

```
serie = ts (data,start,end,frequency)
```

— **data** contient le vecteur des données.

— **start** et **end** mentionne les dates de début et de fin de la série.

— **frequency** mentionne le nombre de données par unité de temps.

Pour représenter la série temporelle graphiquement on utilise la commande :  
`plot.ts(nom de la serie)`

### Exemple 3.1 Simuler un bruit blanc $BB(0,1)$

```
bruitblanc = rnorm(200,0,1)
plot(bruitblanc,type='l,col="red")
```

On obtient la trajectoire suivante :

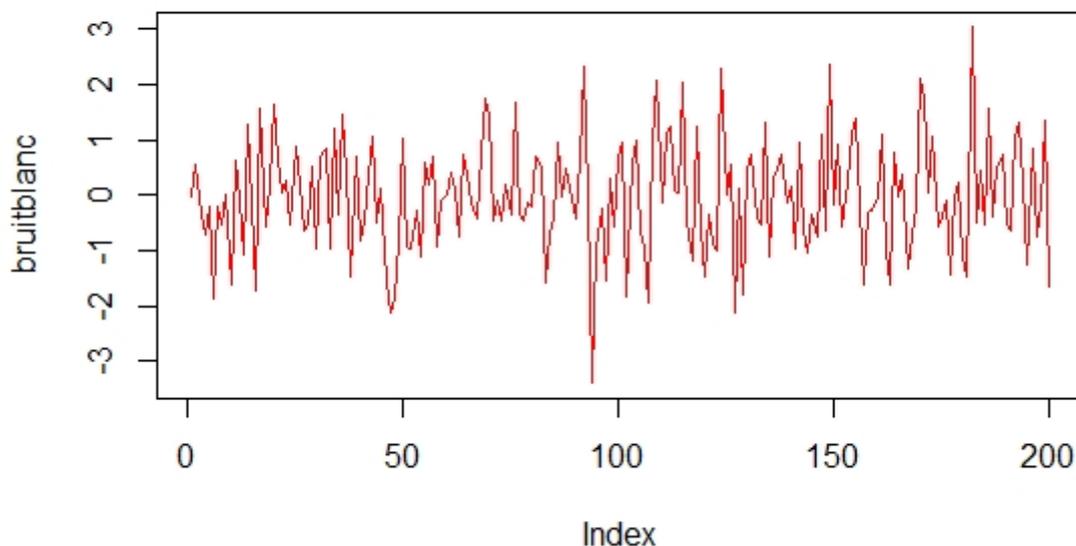


FIGURE 3.1 – Le graphe d'une série Bruit-blanc.

### 3.2.2 Transformation

La fonction **log** permet de stabiliser la variance.

La fonction **diff** permet d'appliquer l'opérateur de différenciation  $\Delta_T^k$  en utilisant la fonction :

$$\text{diff}(\text{serie}, \text{lag} = T, \text{difference} = k)$$

$\Delta_T^k$  élimine une tendance et une saisonnalité d'ordre  $k$ .

On peut aussi utiliser la fonction **decompose** pour extraire les composantes d'une série temporelle (via la méthode de la moyenne mobile) :

```
decomp<-decompose(serie,type=c('additive','mutliplicative'))
```

- **la composante saisonnière** : par `decomp$seasonal`, que l'on suppose additive ou multiplicative dans l'option `type`.
- **la tendance** : par `decomp$trend`.
- **la partie aléatoire stationnaire de la série** : par `decomp$random`.

Après avoir éliminer la tendance et la saisonnalité, l'étape suivante consiste à modéliser la série résiduelle obtenue. La première chose à faire est donc de tester s'il y a une dépendance entre les termes de cette série. Si cette série est un bruit blanc, il n'est donc pas utile d'aller plus loin dans la modélisation. Pour cela on utilise les fonctions suivantes :

(1) La fonction **Box.test(serie)** teste la bruité de la série.

Si p-value > à 0.05 la série est un bruit blanc.

(2) **qqplot**, **qqnorm** et **jarque.bera.test** ce sont des fonctions pour tester la normalité des résidus.

(3) La fonction **adf.test(serie)** teste la stationnarité de la série. Si p-value < à 0.05 la série est stationnaire. On peut

également utiliser la fonction **KPSS.test(série)**, Si p-value > à 0.05 la série est stationnaire.

On peut aussi utiliser la fonction **pp.test(série)**, Si p-value < à 0.05 la série est stationnaire et admet une racine unitaire.

### 3.2.3 Identification

La fonction **acf(x)** calcule (et trace si l'option plot est à **TRUE**) les lag.max premières auto-corrélations empiriques et auto-covariances.

**Exemple 3.2 (Identification du bruit) :** *On simule un bruit blanc gaussien par :*

$$x = rnorm(50)$$

*Et s'affiche le corrélogramme de x :*

$$r = acf(x)$$

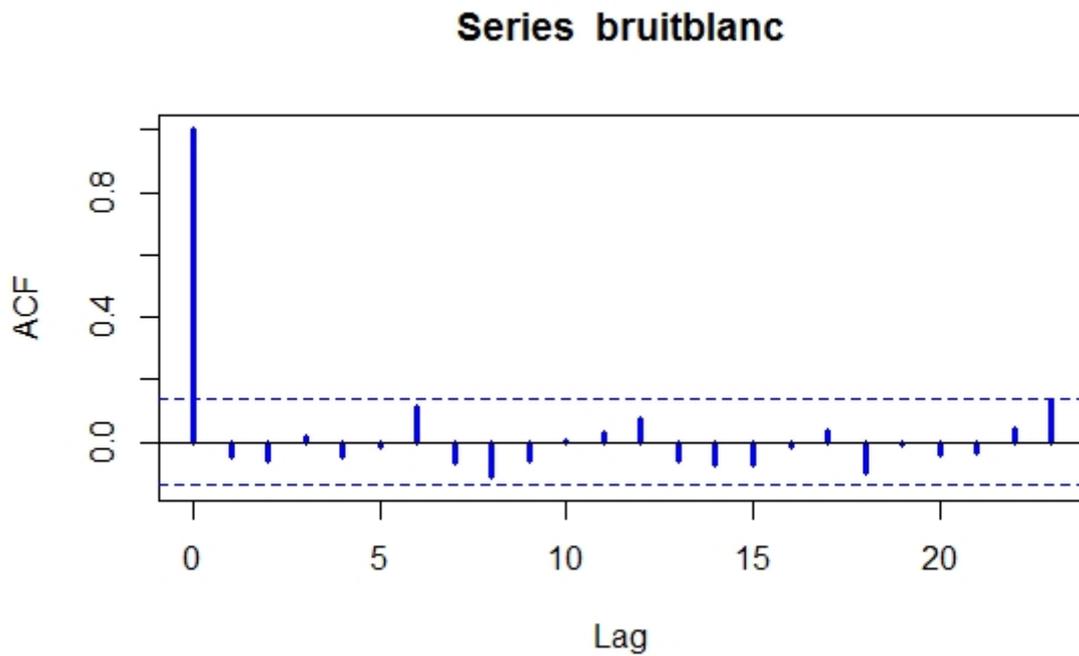


FIGURE 3.2 – Représentation graphique d'autocorrélation simple de la série Bruit blanc.

La fonction **pacf(x)** calcule et trace les autocorrélations partielles empiriques.

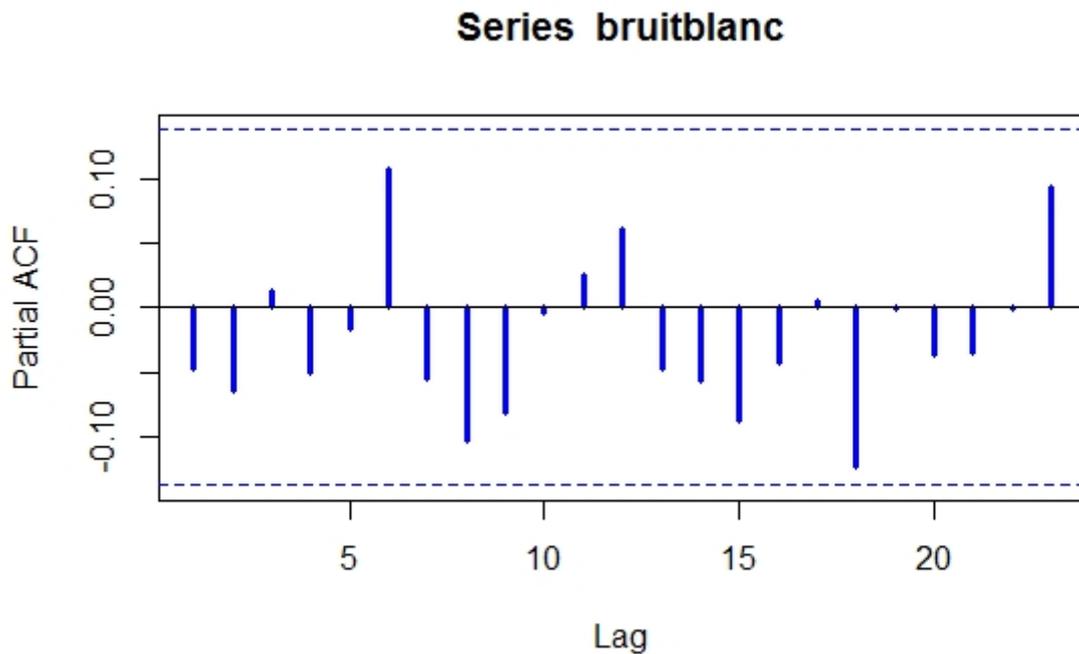


FIGURE 3.3 – Le corrélogramme partiel de la série Bruit blanc.

### 3.2.4 Estimation des paramètres

**Estimation des paramètres du modèle :** Dans le cas général la fonction `arima` permet d'estimer les paramètres :

- D'un ARMA(p,q) : `out=arima(serie,order=c(p,0,q))`.
- D'un ARIMA(p,d,q) : `out=arima(serie,order=c(p,d,q))`.
- D'un SARIMA(p,d,q,T) : `out=arima(serie,order=c(p,d,q),seasonal=list(order=c(P,D,Q),period=`

### 3.2.5 Tests d'adéquation et choix du modèle (validation)

Après avoir estimé le modèle, il faut s'assurer encore une fois que les résidus vérifient les propriétés d'un bruit blanc sans oublier de tester aussi la normalité, on utilise alors :

- `Box.test(out$resid)` # pour tester la bruité.
- `adf.test(out$resid)` # pour tester la stationnarité.
- `qqnorm(out$resid)` # pour tester la stationnarité graphiquement.
- `jarque.bera.test(out$resid)` # pour tester la normalité la normalité des résidus.

### 3.2.6 Prévision

Si à la fin on a plusieurs modèles valides on choisit parmi eux celui qui donne les meilleurs critères d'ajustement et les meilleures performances en prévision.

La fonction `predict` permet de faire les prévisions à l'horizon `h` : `prevision1=predict(out,n.ahead=h)`.

## 3.3 Application

### 3.3.1 Application sur les données simulées

Exemple 1 : Le modèle ARIMA(2,0,1)

A des fins de démonstration, nous travaillerons sur des données simulées. nous avons généré 1000 points de données à partir d'un certain processus ARIMA.

Soit le processus défini par ARIMA(2,0,1), avec comme paramètres

$$\begin{cases} \phi_1 = 0.5 \\ \phi_2 = -0.5 \\ \theta_1 = 0.1 \end{cases}$$

S'écrit comme :

$$(1 - 0.5B + 0.5B^2)X_t = (1 - 0.1B)\varepsilon_t$$

#### 3.3.1.1 Analyse graphique

```
set.seed(1234)
```

```
Serie_1=arima.sim(list(order=c(2,0,1),ar=c(0.5,-0.5),ma=0.1),n=1000)
```

`plot.ts(Serie_1,col='red')`

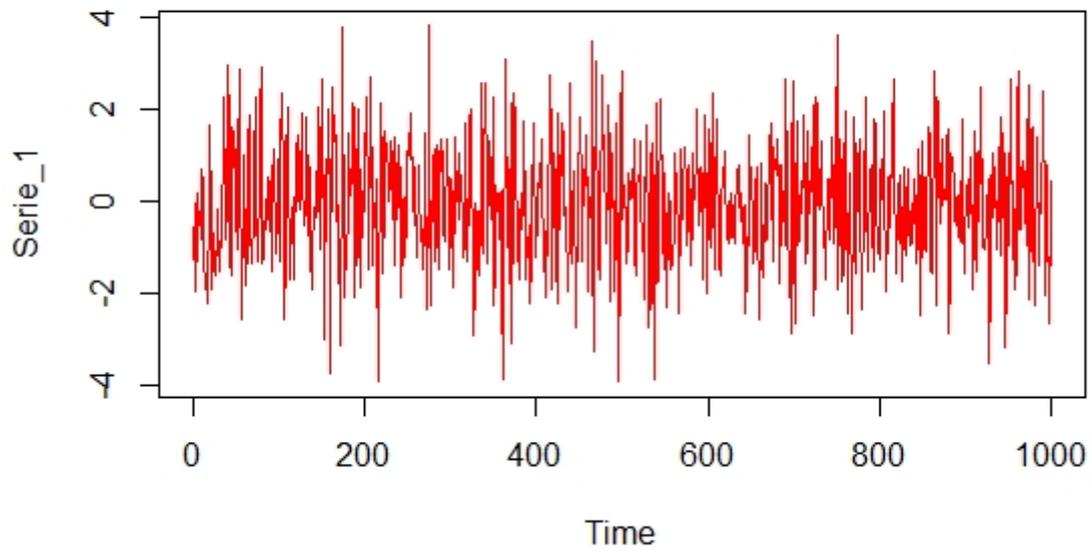


FIGURE 3.4 – Le graphe du modèle  $ARIMA(2,0,1)$ .

Son corrélogramme simple

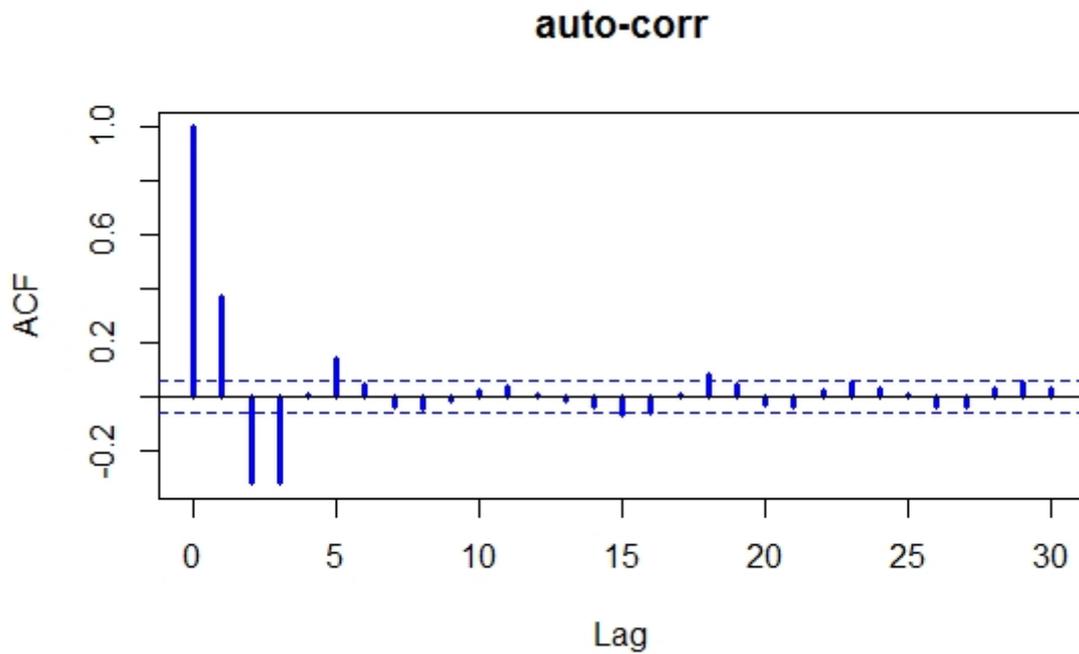


FIGURE 3.5 – Le corrélogramme simple du modèle  $ARIMA(2,0,1)$ .

Et son corrélogramme partiel est

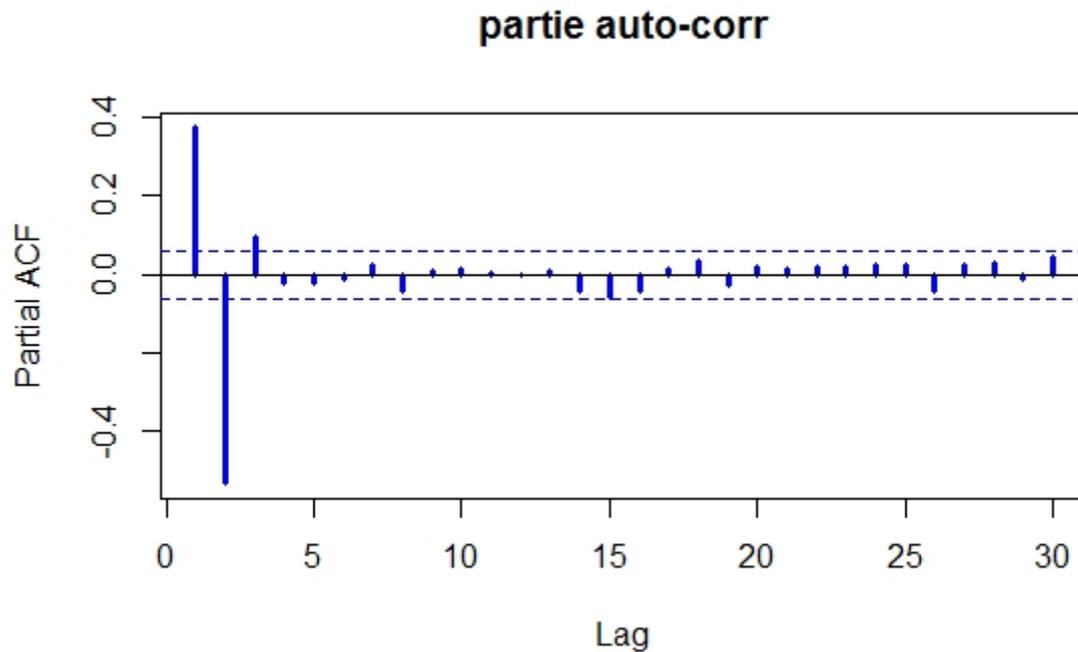


FIGURE 3.6 – Représentation graphique d'autocorrélation partielle du modèle  $ARIMA(2,0,1)$ .

### 3.3.1.2 Tests de stationnarité

- **ADF** : avec le packages `library(tseries)`, et la commande : `adf.test(Serie)` on obtient :

```
adf.test(Serie_1)
```

**Augmented Dickey-Fuller Test**

```
data : Serie_1
```

```
Dickey-Fuller = -10.377, Lag order = 9, p-value = 0.01
```

$p - value = 0.01 < 0.05$  donc la séries de modèle  $ARIMA(2,0,1)$  est stationnaire.

- **PP** : avec le packages : `library(tseries)`, et la commande : `pp.test(Serie)` on obtient :

```
pp.test(Serie_1)
```

**Phillips-Perron Unit Root Test**

```
data : Serie_1
```

```
Dickey-Fuller Z(alpha) = -420.26, Truncation lag parameter = 7, p-value = 0.01
```

$p - value = 0.01 < 0.05$  donc la séries de modèle  $ARIMA(2,0,1)$  est stationnaire.

- **KPSS** : avec le packages : `library(tseries)`, et la commande : `kpss.test(Serie)` on obtient :

```
kpss.test(Serie_1)
```

**KPSS Test for Level Stationarity**

```
data : Serie_1
```

```
KPSS Level = 0.10107, Truncation lag parameter = 7, p-value = 0.1
```

$p - value = 0.1 > 0.05$  donc la séries de modèle  $ARIMA(2,0,1)$  est stationnaire.

Donc d'après les trois test de stationnarité on resume que la série qui suit le modèle  $ARIMA(2,0,1)$  est stationnaire.

### 3.3.1.3 Prédiction

```
library(forecast)
m_aa=auto.arima(Serie_1)
f_aa=forecast(m_aa,h=50)
plot(f_aa,col='red')
```

On obtient

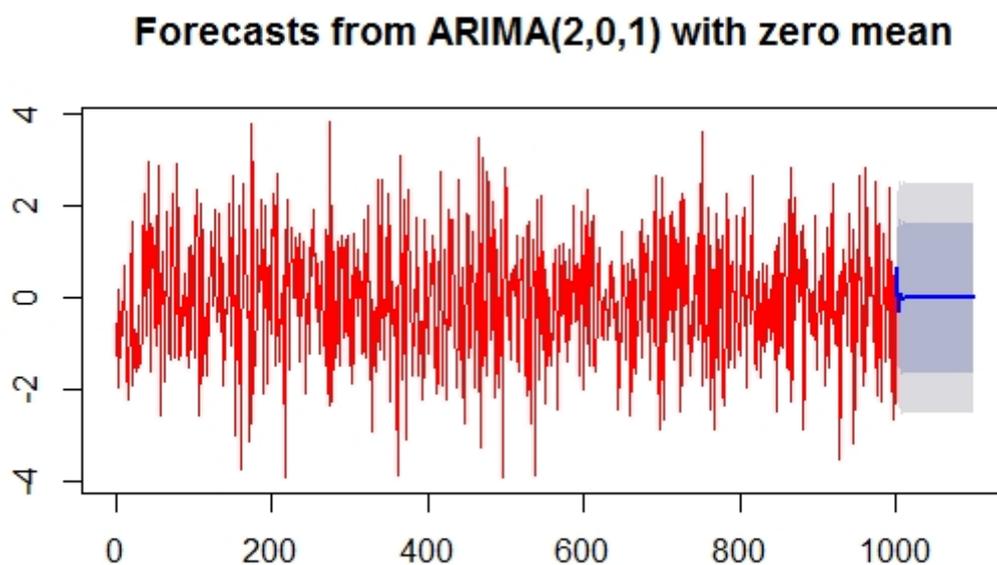


FIGURE 3.7 – Représentation graphique des valeurs prédictives.

### Exemple 2 : Le modèle $SARIMA(2,1,1)(1,1,1)[12]$

On doit généré 500 observations à partir du modèle  $SARIMA(2,1,1)(1,1,1)[12]$  par la commande suivant :

```
set.seed(123)
library(forecast)
modèle <- arima(ts(rnorm(100), frequency = 12), order = c(2, 1, 1),
seasonal = c(1, 1, 1), fixed = c(phi1 = 0.5, phi2 = -0.5, theta = -0.4, Phi = 0.3, Theta =
-0.2))
modèle <- simulate(modèle, n = 500)
plot.ts(modèle)
```

On trouve la représentation graphique suivante :

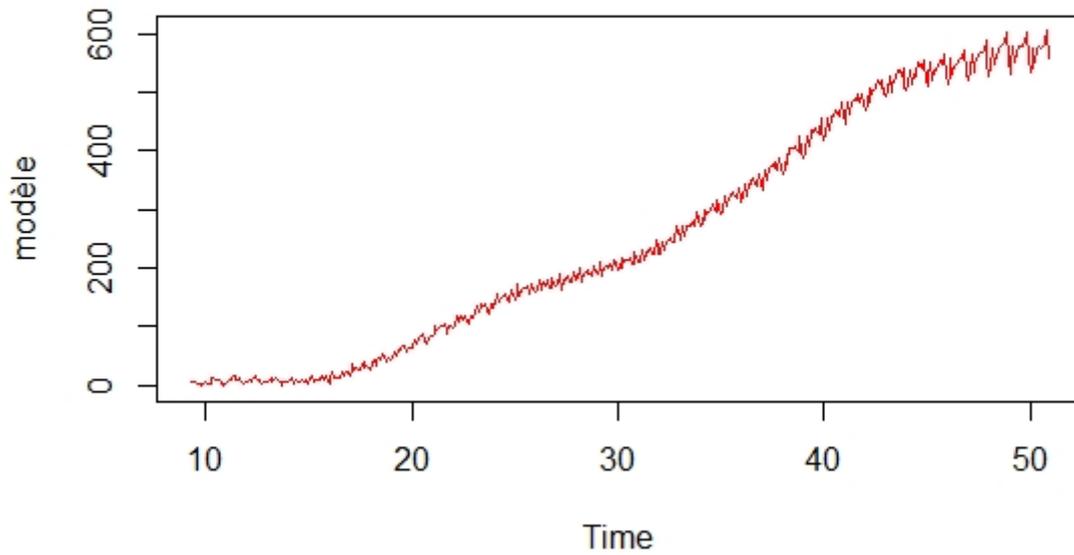


FIGURE 3.8 – Représentation graphique du modèle SARIMA.

Son corrélogramme simple est :

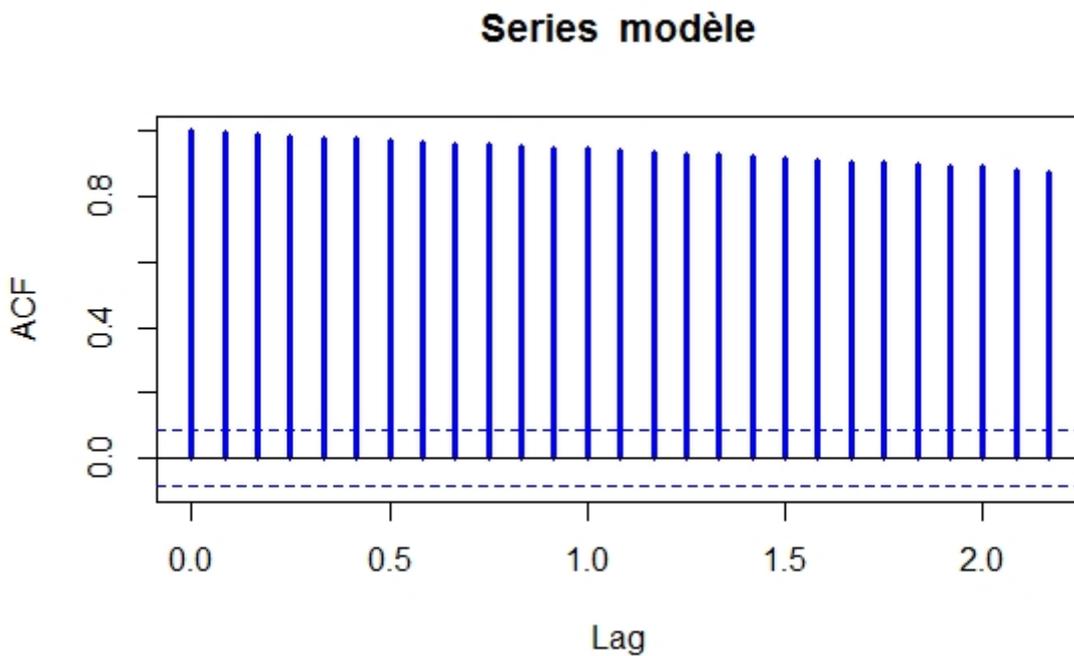


FIGURE 3.9 – Le corrélogramme simple.

Et son corrélogramme partiel est :

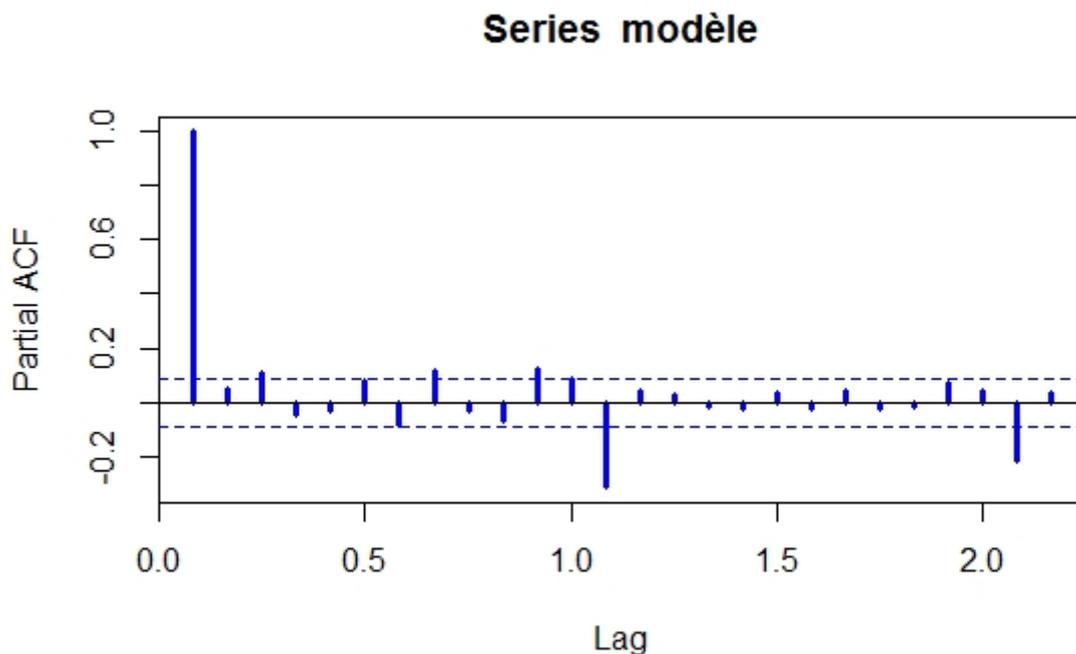


FIGURE 3.10 – Représentation graphique d'autocorrélation partielle.

### 3.3.2 Vérification de la stationnarité

```

— library(tseries)
— adf.test(modle)
— AugmentedDickey-FullerTest
— data :modle
— Dickey-Fuller = (-2.685, Lagorder=7, p-value=0.2883)
— kpss.test(modle)
— KPSSTestforLevelStationarity
— data :modle
— KPSSLevel=8.3592, Truncationlagparameter=5, p-value=0.01
— pp.test(modle)
— Phillips-PerronUnitRootTest
— data :modle
— Dickey-FullerZ(alpha)= -18.409, p-value=0.09546

```

D'après les tests de stationnarité on remarque que le modèle  $SARIMA(2, 1, 1)(1, 1, 1)[12]$  n'est pas stationnaire.

### 3.3.3 Stationnarisation de la série

```
ndiffs(modle)[1]1
```

Donc l'ordre de différenciation est  $d = 1$ .

Et la représentation graphique de la série diff est :

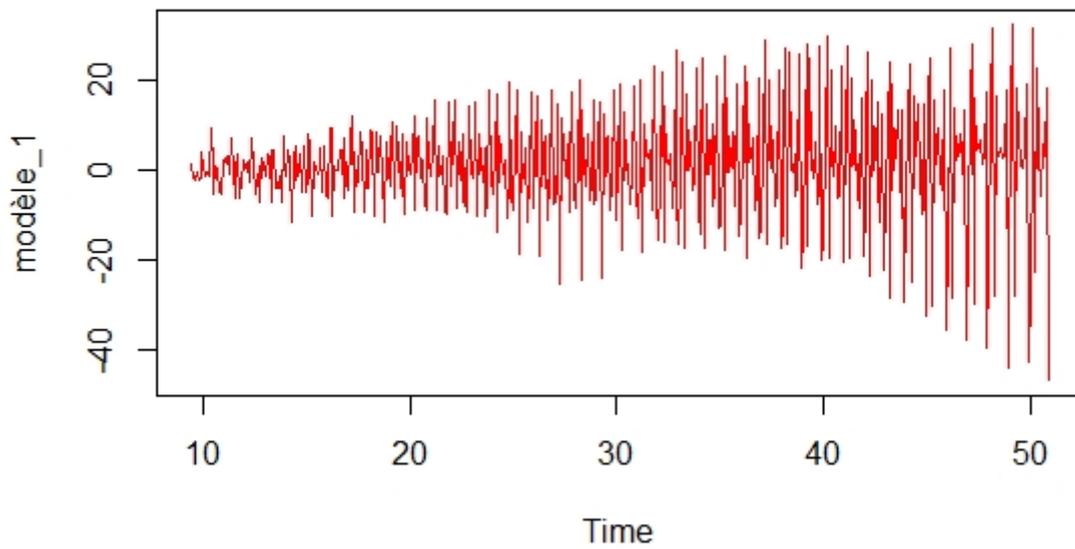


FIGURE 3.11 – Représentation graphique du modèle diff\_SARIMA.

Son corrélogramme simple est

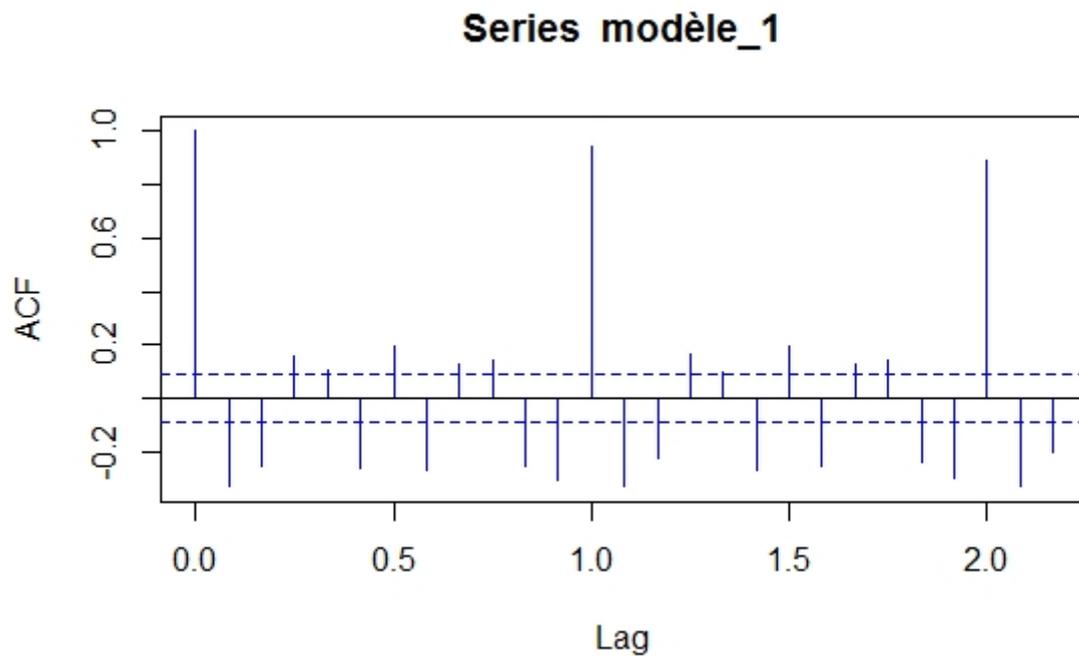


FIGURE 3.12 – Le corrélogramme simple du modèle fiff\_SARIMA.

Et son corrélogramme partiel est :

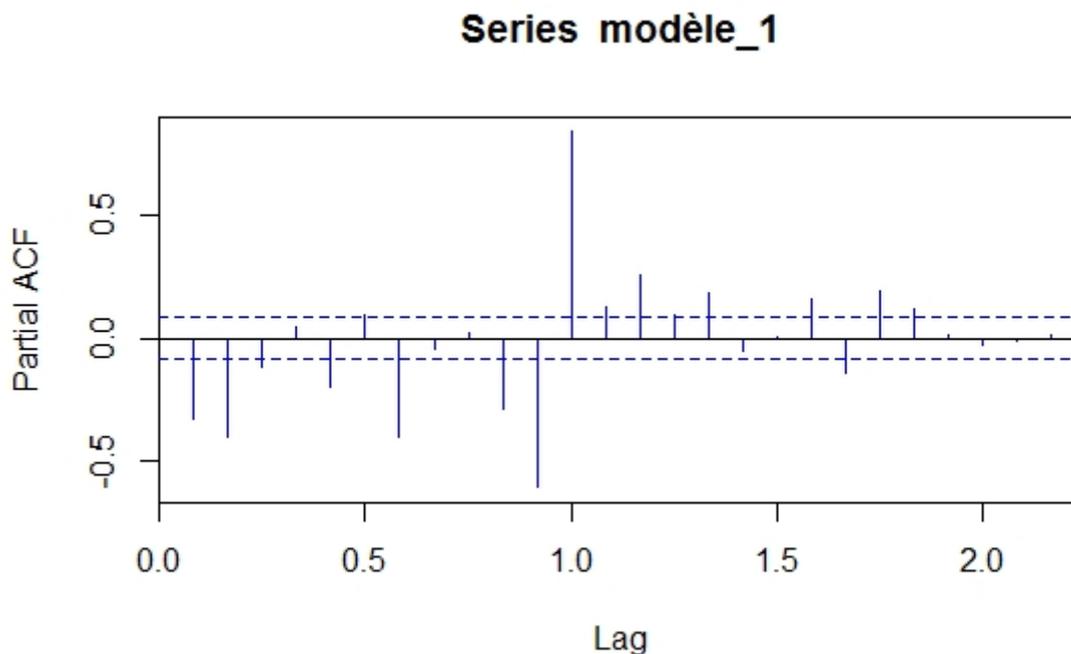


FIGURE 3.13 – Représentation graphique de PACF.

D'après les corrélogramme simple est partiel on remarque que la série diff est stationnaire.

On doit confirmer la stationnarité avec les tests de stationnarité on trouve :

```
— adf.test(modèle_1)
```

*AugmentedDickey – FullerTest*

```
data : modèle_1
```

```
Dickey – Fuller = -13.268, Lagorder = 7, p – value = 0.01
```

```
— kpss.test(modèle_1)
```

*KPSSTestforLevelStationarity*

```
data : modèle_1
```

```
KPSSLevel = 0.15697, Truncationlagparameter = 5, p – value = 0.1
```

```
— pp.test(modèle_1)
```

*Phillips – PerronUnitRootTest*

```
data : modèle_1
```

```
Dickey – FullerZ(alpha) = -530.57, p – value = 0.01
```

Les résultats des tests vérifient la stationnarité.

On doit utiliser la commande *auto.arima* pour identifier le meilleur d'ajustement

```
Modèle = auto.arima(modèle, trace = TRUE, test = "kpss", ic = "aic")
```

```
summary(Modèle)
```

Ce qui donne que le meilleur modèle est *ARIMA(2,1,1)(0,1,0)[12]*

TABLE 3.1 – Coefficients estimé du modèle.

	ar1	ar2	ma1
coéfs	0.5323	-0.5156	-0.4788
s.e.	0.0663	0.0397	0.0759

avec  $AIC = 2165.45$ .

On doit faire des tests de résidus, on trouve :

— `adf.test(Modèle$residuals)`

*AugmentedDickey – FullerTest*

*data : Modèle\$residuals*

*Dickey – Fuller = -7.9993, Lagorder = 7, p – value = 0.01*

— `Box.test(Modèle$residuals)`

*Box – Piercetest*

*data : Modèle\$residuals*

*X – squared = 0.051227, df = 1, p – value = 0.8209*

— `qqnorm(Modèle$residuals)`

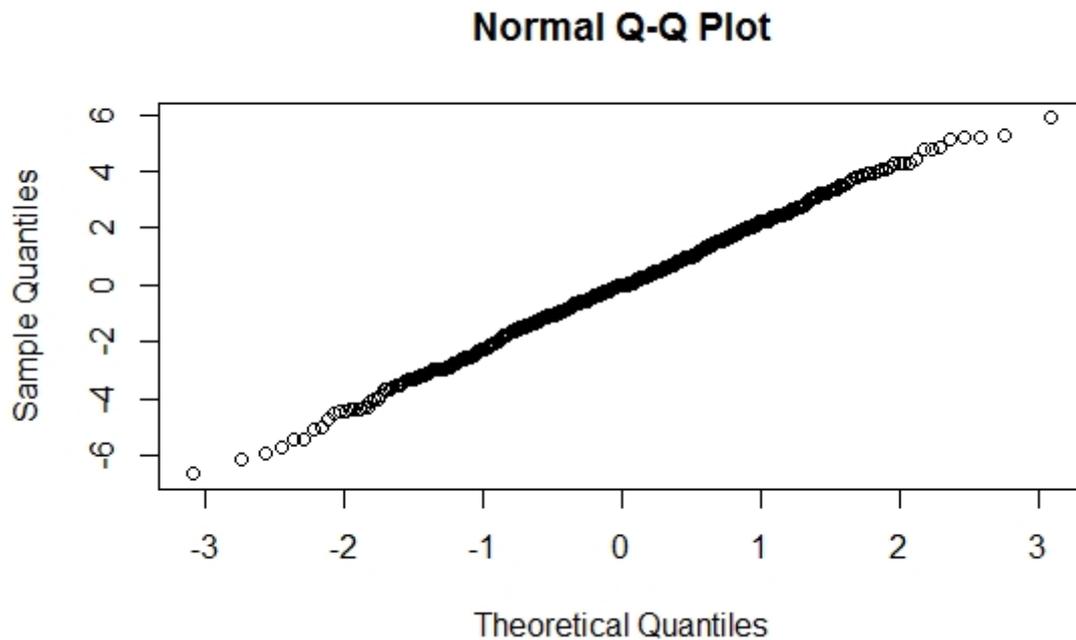


FIGURE 3.14 – Le graphe qqnorm.

On résume que le modèle ajusté qui est  $ARIMA(2, 1, 1)(0, 1, 0)[12]$  est valide.

On passe à la prévision :

`library(forecast)`

`m_aa = auto.arima(modèle)`

`f_aa = forecast(m_aa, h = 120)`

`plot(f_aa, col = red)`

### Forecasts from ARIMA(2,1,1)(0,1,0)[12]

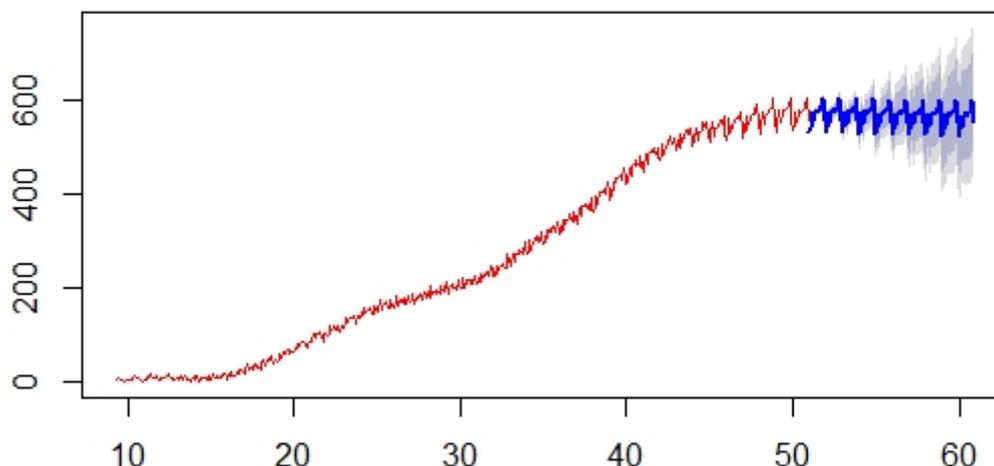


FIGURE 3.15 – Représentation graphique des valeurs prédictives.

### 3.3.4 Application sur des données réelles

Le tableau suivant présente le prix du baril de pétrole en dollar (usa \$) délivré par ARAMCO de l'Arabie Saoudite, les données considérées sont mensuelles et la période retenue pour l'étude est de 1<sup>er</sup> mois de 2010 à 3<sup>ème</sup> mois 2019.

TABLE 3.2 – Tableau représente le prix du baril de pétrole en dollar à l'Arabie Saoudite.

DATE	Jan	fév	Mars	Avr	Mai	Juin	Juil	Août	Sept	Oct	Nov	Dé
2010	53.52	54.3	58.43	63.4	60.65	61.31	58.55	59.49	59.52	59.68	62.39	69.
2011	72.09	76.18	81.76	85.23	79.78	79.07	81.72	76.75	80.65	79.87	81.37	81.
2012	86.08	90.54	94.64	91.53	86.37	76.32	84.03	91.41	88.1	86.3	85.58	83.
2013	85.04	87.23	84.27	78.98	79.36	78.19	82.34	83.37	83.59	80.28	80.17	80.
2014	78.91	79.68	77.7	78.04	79.85	82.3	79	76.54	75.52	68.87	62.9	50.
2015	41.46	51.06	51.5	55.11	57.87	55.6	50.83	42.2	42.09	42.86	41.32	34.
2016	28.37	29.94	35.21	37.26	41.64	43.18	40.73	41.16	41.19	45.13	42.93	51.
2017	51.68	52.14	48.64	49.44	46.04	41.76	42.27	43.51	46.3	49.01	53.37	54.
2018	56.58	53	53.87	58.35	64.85	64.39	63.69	63.33	67.63	70.07	57.34	49.
2019	51.91	56.49	58.64									

#### 3.3.4.1 Analyse graphique de la série "Prix du baril de pétrole"

A partir des données de la série, et avec l'ensemble des commandes suivantes : `plot(f_aa,col='red')`  
`S<-read.csv2("prix du baril de pétrole.CSV",header=TRUE)` `S<-t(S[, -1])` `S<-as.vector(S)`  
`data<-ts(S,start = c(2010, 1),end=c(2019, 3),frequency = 12)` `data`

Nous obtenons le graphique suivant :

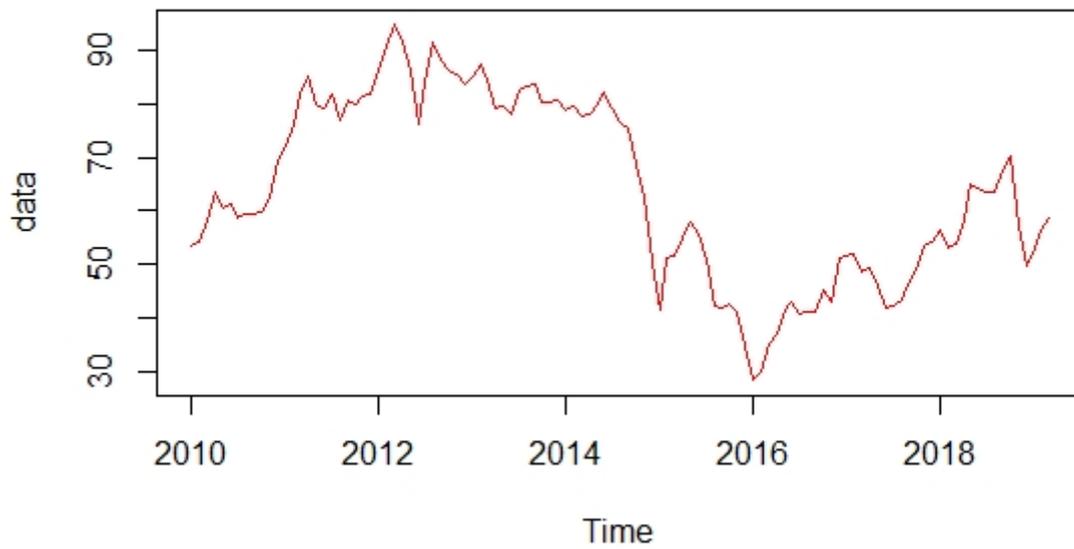


FIGURE 3.16 – Représentation graphique de la série des Prix du pétrole.

Le corrélogramme simple :

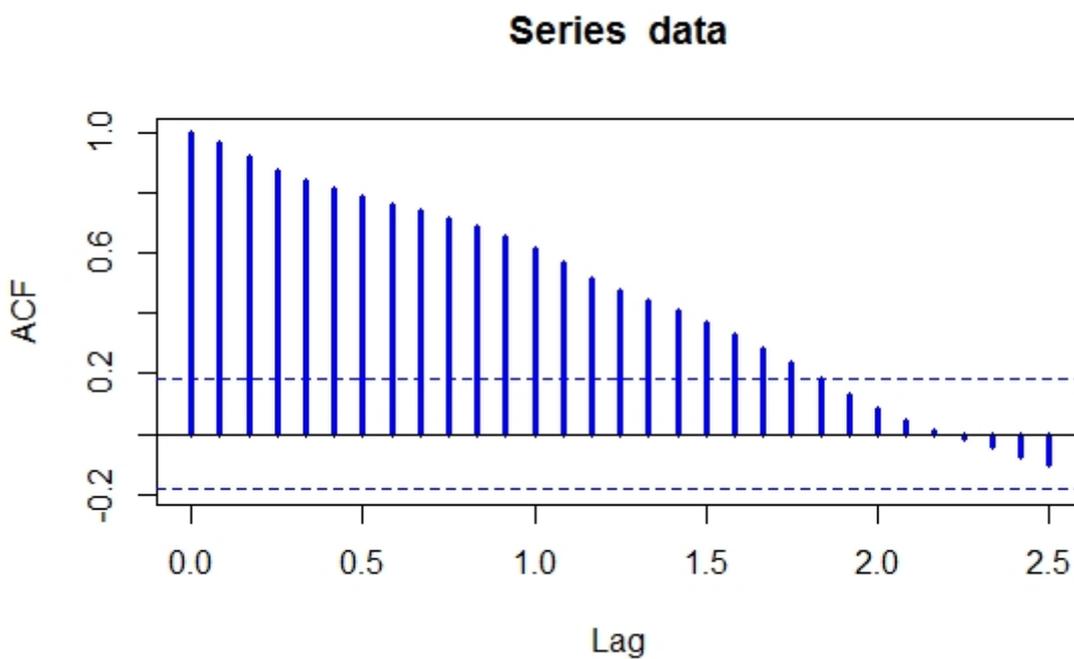


FIGURE 3.17 – Représentation graphique d'autocorrélation simple de la série Prix du baril de pétrole.

Et sa corrélogramme partiel est

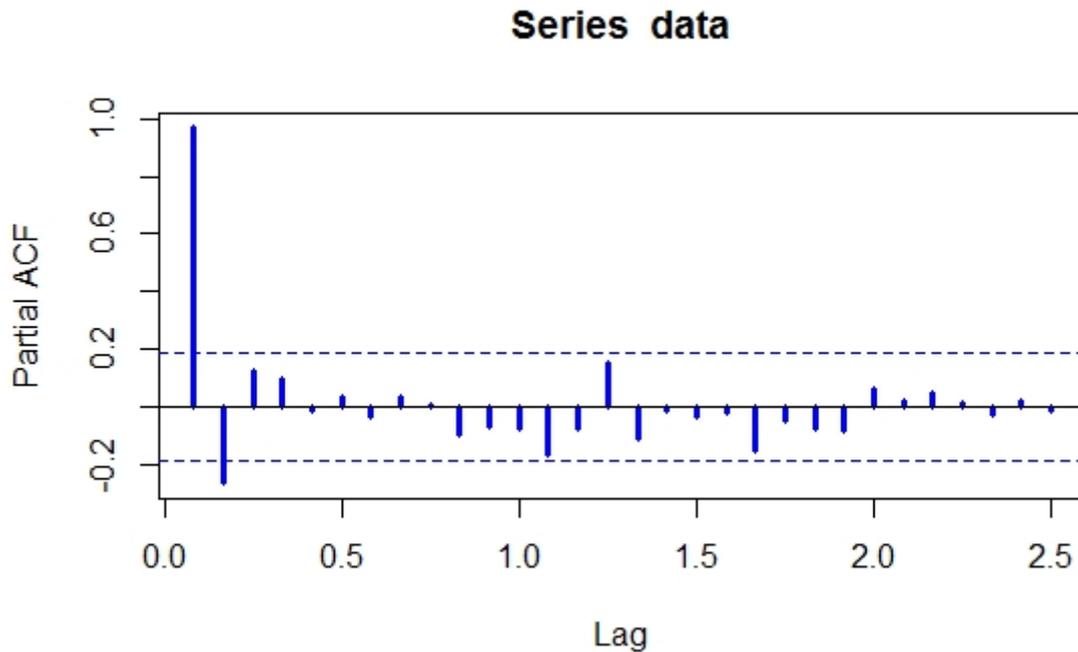


FIGURE 3.18 – Le corrélogramme partiel de la série Prix du baril de pétrole.

La visualisation des corrélogramme simple et partiel montre une non stationnarité propre, l'ACF ne coupe pas.

On utilise aussi le test de racine unitaire (test ADF) pour confirmer la non stationnarité :

```
library(tseries)
```

```
adf.test(data)
```

**Augmented Dickey-Fuller Test**

```
data : data
```

```
Dickey-Fuller = -2.0149, Lag order = 4, p-value = 0.5704
```

$p\text{-value} = 0.5704 > 0.05$  donc la série n'est pas stationnaire (admet une racine unitaire).

On peut aussi utiliser la fonction **decompose** pour extraire les composantes d'une série temporelle

```
dec.data=decompose(data)
```

```
plot(dec.data)
```

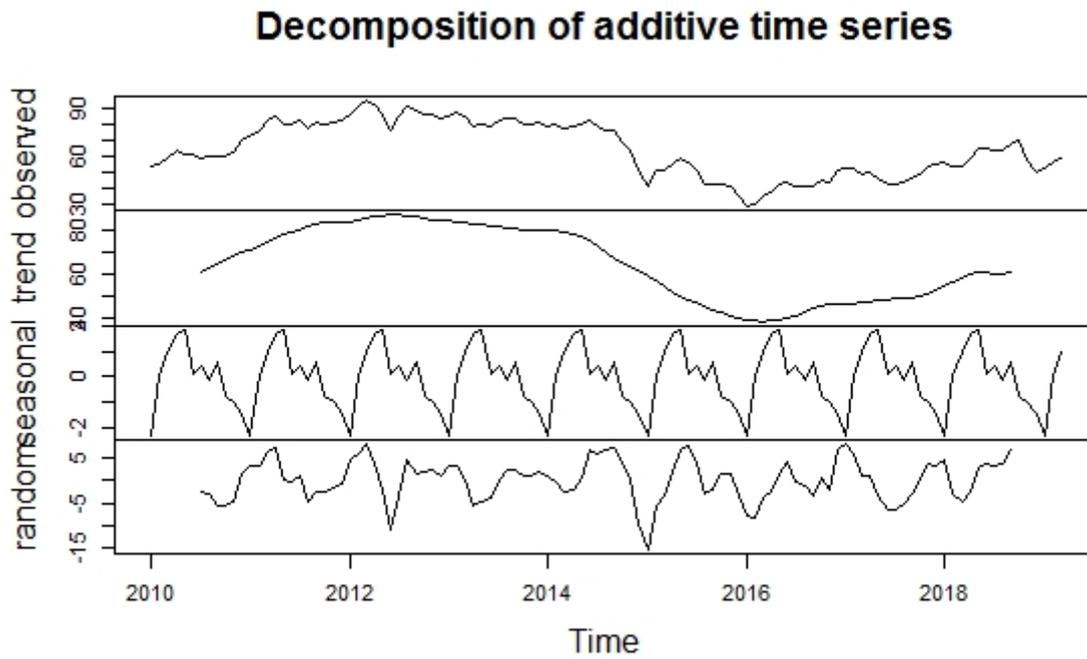


FIGURE 3.19 – Les composantes de la série Prix du baril de pétrole.

### 3.3.4.2 Transformation de la séries

On utilise la commande `ndiffs` pour savoir l'ordre de différenciation qui éliminé la tandance et la saisonnalité.

```
library(forecast)
```

```
ndiffs(data)
```

```
[1] 1
```

Donc  $d = 1$  .

```
data_2=diff(data,df=1)
```

```
data_2
```

On obtient l'ensemble des données suivantes :

TABLE 3.3 – Tableau représente les données de la fonction diff du prix du baril de pétrole.

DATE	Jan	fév	Mars	Avr	Mai	Juin	Juil	Août	Sept	Oct	Nov	Déc
2010		0.78	4.13	4.97	-2.75	0.66	-2.76	0.94	0.03	0.16	2.71	7.06
2011	2.64	4.09	5.58	3.47	-5.45	-0.71	2.65	-4.97	3.90	-0.78	1.50	0.53
2012	4.18	4.46	4.10	-3.11	-5.16	-10.05	7.71	7.38	-3.31	-1.80	-0.72	-1.91
2013	1.37	2.19	-2.96	-5.29	0.38	-1.17	4.15	1.03	0.22	-3.31	-0.11	0.62
2014	-1.88	0.77	-1.98	0.34	1.81	2.45	-3.30	-2.46	-1.02	-6.65	-5.97	-12.35
2015	-9.09	9.60	0.44	3.61	2.76	-2.27	-4.77	-8.63	-0.11	0.77	-1.54	-6.64
2016	-6.31	1.57	5.27	2.05	4.38	1.54	-2.45	0.43	0.03	3.94	-2.20	8.36
2017	0.39	0.46	-3.50	0.8	-3.40	-4.28	0.51	1.24	2.79	2.71	4.36	0.88
2018	2.33	-3.58	0.87	4.48	6.50	-0.46	-0.70	-0.36	4.30	2.44	-12.73	-7.72
2019	2.29	4.58	2.15									

La représentation graphique est

```
plot.ts(data_2,col='red')
```

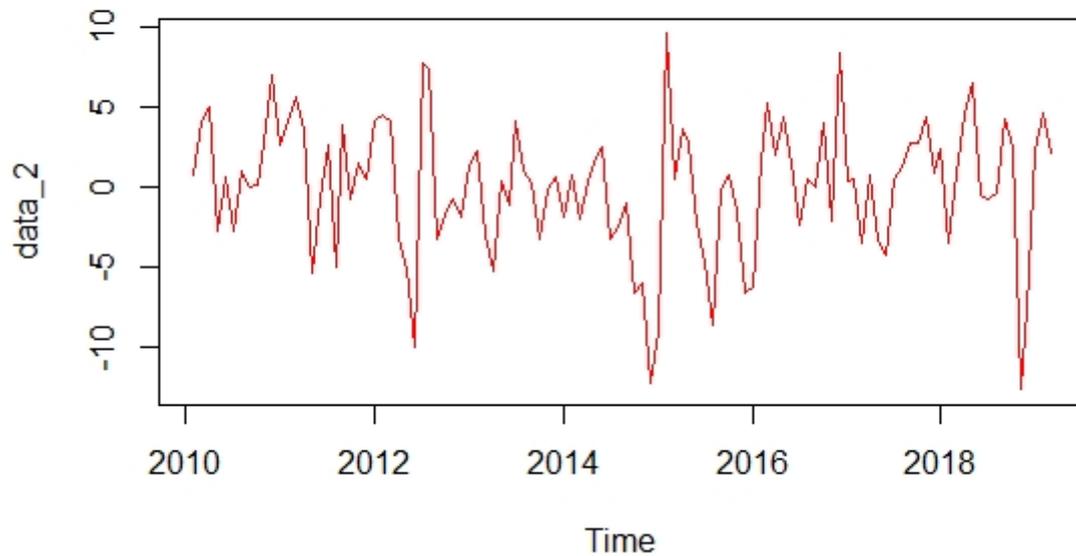


FIGURE 3.20 – La représentation graphique de la série différencier.

On applique l'ensemble des test de la stationnarité on obtient :

```
— library(tseries)
```

```
  kps.test(data_2)
```

**KPSS Test for Level Stationarity**

**data : data\_2**

**KPSS Level = 0.17156, Truncation lag parameter = 4, p-value = 0.1**

```
— library(tseries)
```

```
  pp.test(data_2)
```

**Phillips-Perron Unit Root Test**

**data : data\_2**

**Dickey-Fuller Z(alpha) = -70.249, Truncation lag parameter = 4, p-value = 0.01**

```
— library(tseries)
```

```
  adf.test(data_2)
```

**Augmented Dickey-Fuller Test**

**data : data\_2**

**Dickey-Fuller = -4.7578, Lag order = 4, p-value = 0.01**

D'après les résultats des trois test on résume que la série différencier est stationnaire.

### 3.3.4.3 Identification

Les graphes de autocorrélation simple et partielle de la série **diff** sont :

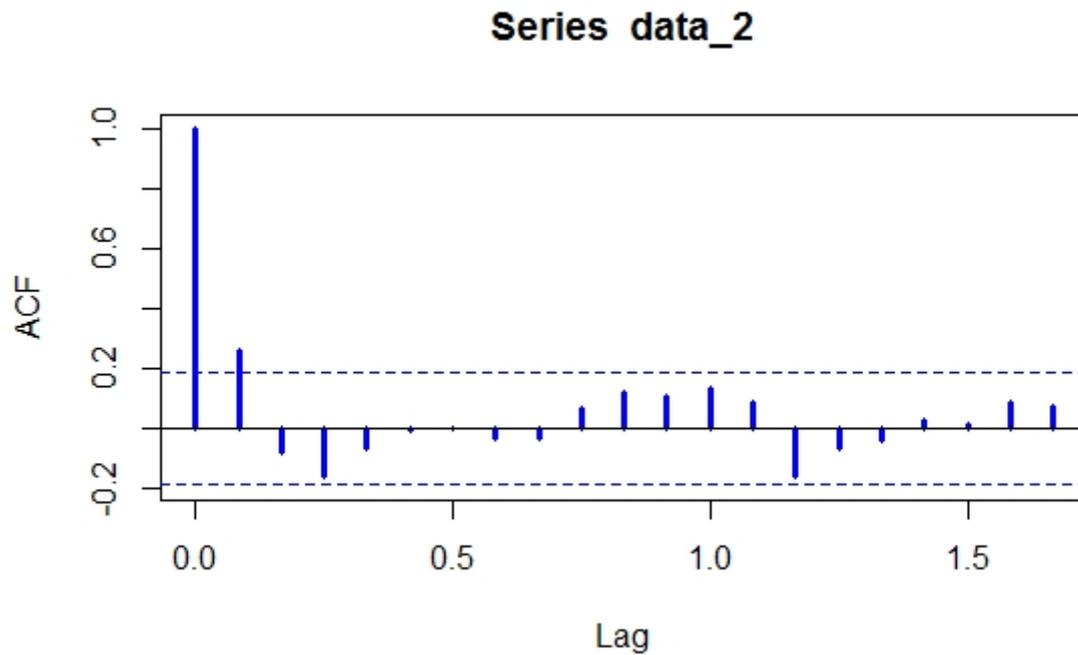


FIGURE 3.21 – Le corrélogramme simple de la série diff.

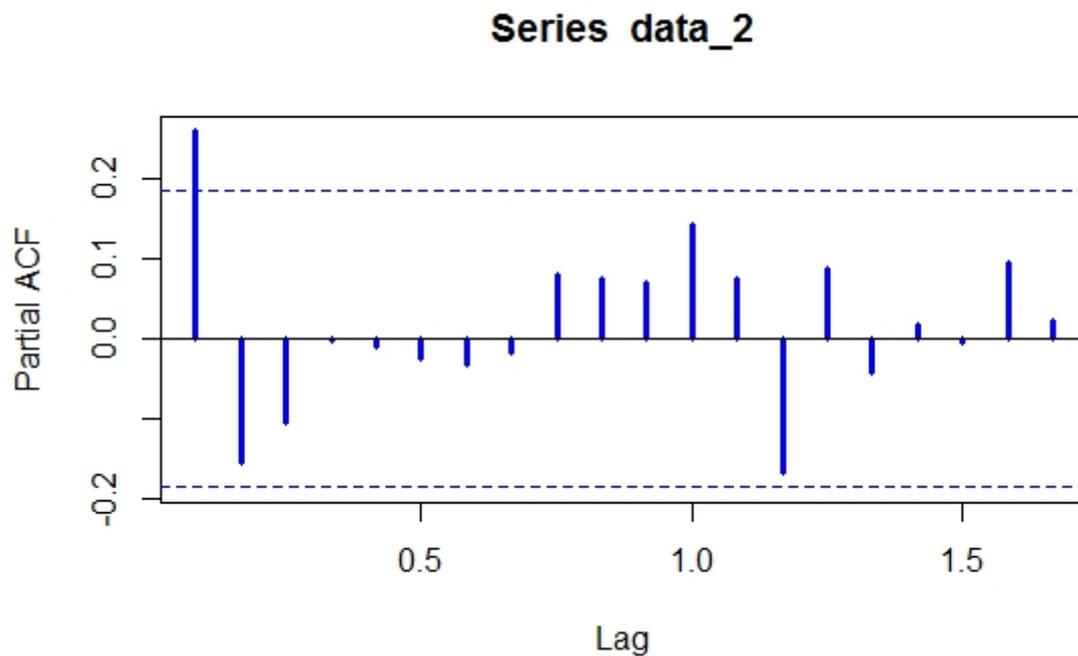


FIGURE 3.22 – Le corrélogramme partiel de la série diff.

Nous voyons que l'échantillon ACF est coupé après le décalage 1 et que l'échantillon

PACF est coupé après le décalage 0. Nous proposons donc deux modèles *ARMA* pour les données différenciées : *ARMA*(0,0) et *ARMA*(0,1)

C'est à dire que, pour la série chronologique originale, nous proposons deux modèles *ARIMA*, *ARIMA*(0,1,0) et *ARIMA*(0,1,1).

### 3.3.4.4 Estimation des paramètres

On utilise la commande *auto.arima* pour construire le meilleur modèle qui représente la série.

**library(forecast)**

**Modèle=auto.arima(data,trace=TRUE,test="kpss",ic="aic")**

on trouve

TABLE 3.4 – Tableau represente les différents modèles ajustés.

Modèles	La valeur de AIC
ARIMA(2,1,2)(1,0,1)[12] with drift	628.6978
ARIMA(0,1,0) with drift	629.8849
ARIMA(1,1,0)(1,0,0)[12] with drift	624.8739
ARIMA(0,1,1)(0,0,1)[12] with drift	623.4131
ARIMA(0,1,0)	627.8986
ARIMA(0,1,1) with drift	622.7928
ARIMA(0,1,1)(1,0,0)[12] with drift	623.7082
ARIMA(0,1,1)(1,0,1)[12] with drift	625.2132
ARIMA(1,1,1) with drift	624.7928
ARIMA(0,1,2) with drift	624.7928
ARIMA(1,1,0) with drift	624.1918
ARIMA(1,1,2) with drift	625.8999
ARIMA(0,1,1)	620.803
ARIMA(0,1,1)(1,0,0)[12]	621.7265
ARIMA(0,1,1)(0,0,1)[12]	621.4312
ARIMA(0,1,1)(1,0,1)[12]	623.2288
ARIMA(1,1,1)	622.803
ARIMA(0,1,2)	622.803
ARIMA(1,1,0)	622.2034
ARIMA(1,1,2)	623.908

Qui donne que le meilleur modèle est *ARIMA*(0,1,1) car il est minimise le critère AIC

Aussi on a la commande **summary(Modèle)** qui donne

**Series : data**

**ARIMA(0,1,1)**

	ma1
coef	0.2926
s.e	0.0880
AIC	620.8

### 3.3.4.5 Validation

Nous avons diagnostiqué notre modèle à partir des tests suivants pour montrer que le modèle choisis est valide.

#### 3.3.4.5.1 Test ADF `adf.test(Modèle$residuals)`

Augmented Dickey-Fuller Test

data : `Modèle$residuals`

Dickey-Fuller = -5.0071, Lag order = 4, p-value = 0.01

#### 3.3.4.5.2 Test de Box-pierce `Box.test(Modèle$residuals)`

Box-Pierce test

data : `Modèle$residuals`

X-squared = 0.00071317, df = 1, p-value = 0.9787

#### 3.3.4.5.3 Tester de la stationnarité graphiquement La commande :`qqnorm(modèle_1$residuals)`

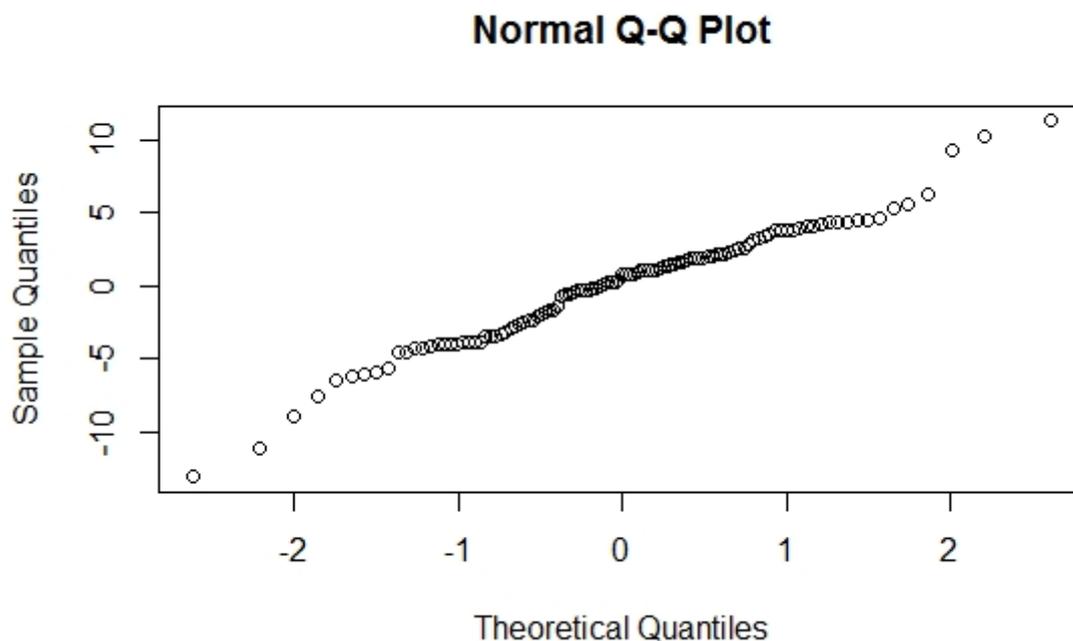


FIGURE 3.23 – Le graphe qqnorm du modèle choisis.

### 3.3.4.6 Prévision

D'après le modèle choisis précédent on peut prédire des résultats pour les deux prochaines années.

On utilise l'ensemble des commandes suivantes :

`library(forecast)`

```
m_aa=auto.arima(data)
f_aa=forecast(m_aa,h=24)
f_aa
On aura
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Apr 2019	58.96289	53.82346	64.10233	51.102805	66.82298
May 2019	58.96289	50.56379	67.36200	46.117574	71.80821
Jun 2019	58.96289	48.25420	69.67159	42.585363	75.34042
Jul 2019	58.96289	46.36102	71.56477	39.689992	78.23580
Aug 2019	58.96289	44.71725	73.20854	37.176062	80.74973
Sep 2019	58.96289	43.24445	74.68134	34.923608	83.00218
Oct 2019	58.96289	41.89829	76.02750	32.864838	85.06095
Nov 2019	58.96289	40.65083	77.27496	30.957005	86.96878
Dec 2019	58.96289	39.48309	78.44270	29.171099	88.75469
Jan 2020	58.96289	38.38149	79.54429	27.486358	90.43943
Feb 2020	58.96289	37.33594	80.58985	25.887320	92.03847
Mar 2020	58.96289	36.33865	81.58714	24.362101	93.56369
Apr 2020	58.96289	35.38351	82.54228	22.901334	95.02445
May 2020	58.96289	34.46558	83.46021	21.497478	96.42831
Jun 2020	58.96289	33.58082	84.34497	20.144359	97.78143
Jul 2020	58.96289	32.72588	85.19991	18.836844	99.08894
Aug 2020	58.96289	31.89793	86.02785	17.570610	100.35518
Sep 2020	58.96289	31.09458	86.83121	16.341979	101.58381
Oct 2020	58.96289	30.31374	87.61205	15.147786	102.77800
Nov 2020	58.96289	29.55362	88.37217	13.985289	103.94050
Dec 2020	58.96289	28.81266	89.11313	12.852091	105.07370
Jan 2021	58.96289	28.08948	89.83631	11.746081	106.17971
Feb 2021	58.96289	27.38285	90.54293	10.665392	107.26040
Mar 2021	58.96289	26.69170	91.23409	9.608361	108.31743

FIGURE 3.24 – Interval de confiance pour le prix du baril avec un seuil de 80% et 95%

La représentation graphique des valeurs prédire est : `plot(f_aa,col='red')`

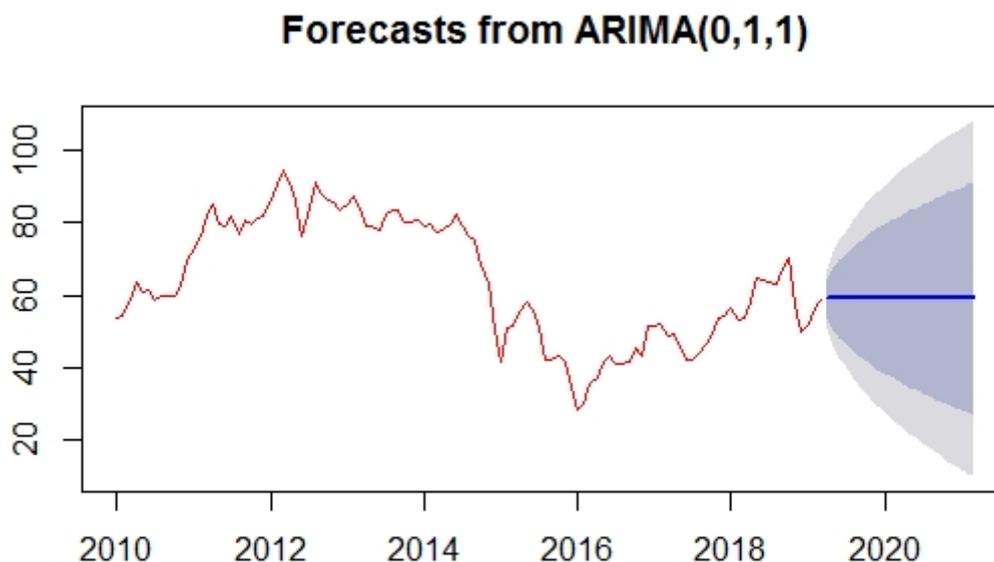


FIGURE 3.25 – Graphe représente les valeurs prédictives de prix du baril de pétrole.

**Conclusion 3.3.1** *D'après nos études on peut conclure que les valeurs prédictives de prix du baril de pétrole en dollar pour les deux années prochaines sera augmente*

## CONCLUSION GÉNÉRALE

L'objectif de ce travail a été d'introduire l'essentiel de la modélisation de séries chronologiques. Dans la première partie de ce travail, on a effectué une revue des principales notions de base concernant les séries chronologiques à une variable, cela a permis d'introduire, entre autres, la famille des processus ARIMA qui sont des modèles mathématiques efficaces pour modéliser des phénomènes temporelles.

Ainsi que la procédure de sélection de modèles de Box & Jenkins qui nous permet de modéliser et prédire le comportement d'un caractère statistique dans horizon. Le logiciel *R* facilite de manipuler et traiter le modèle temporel pour choisir le meilleur modèle d'ajustement, et donne les valeurs des prévision numériquement et visuellement.

On peut conclure à des valeurs prévisionnelles et des résultats satisfaisants et homogènes qui nous conforte dans notre conviction que la méthode utilisée est la plus adéquate pour effectuer notre étude.

## BIBLIOGRAPHIE

- [1] Akaike, H. (1969). Fitting autoregressive models for prediction. *Annals of the institute of Statistical Mathematics*, 21(1), 243-247.
- [2] Bosq, D. et Lectoutre, J.P. (2000). *Analyse et prsion des ses chronolo-gique*. Masson.
- [3] Bowerman, B. L., O'Connell, R. T. (1979). *Time series and forecasting*. North Scituate, MA : Duxbury Press.
- [4] Box, G. E., Jenkins, G. M., Reinsel, G. C., Ljung, G. M. (2015). *Time series analysis : forecasting and control*. John Wiley Sons.
- [5] Corinne, P. (2005). *Ses chronologiques : Quelques ments du cours*. Paris.
- [6] Charpentier, A. (2004). *Cours des ses temporelles. Thie et Applications*. Universit Paris
- [7] Cheung, Y. W., Lai, K. S. (1995). Lag order and critical values of the augmented DickeyFuller test. *Journal of Business Economic Statistics*, 13(3), 277-280.
- [8] Degerine, S. (2005). *Cours de série chronologique*. Université Joseph Fourier.
- [9] Eshel, G. (2003). The yule walker equations for the AR coefficients. *Internet resource*, 2, 68-73
- [10] Florenece, N. (2006). *Généralités sur les séries chronologiques*.
- [11] Girard, Y. (2011). *Série chronologiques e et plusieurs variables : synthèse des méthodes classiques et modèles à base de copules*. Universit Québec.
- [12] Gimet, C. (2007). L'impact des chocs externes dans les économies du Mercosur : un modèle VAR Structurel. *Économie internationale*, (2), 107-136.
- [13] Lévy, C.L. (2013). *Introduction étude des séries temporelles*.
- [14] Lütkepohl, H. (1985). Comparison of criteria for estimating the order of a vector autoregressive process. *Journal of time series analysis*, 6(1), 35-52.
- [15] Mechgoug, R. (2013). *La Prédiction des Séries Temporelles utilisant les Paradigmes de Soft Computing*. Algérie.
- [16] Marine, E. (2015). *Centre de Formation aux Métier de la Montagne*.

- [17] Melard, G. (1979). Modèles ARIMA pour des séries chronologiques non homogènes. *Statistique et analyse des données*, 4(2), 41-50.
- [18] MCLEOD, A. Ian. Diagnostic checking of periodic autoregression models with application. *Journal of Time Series Analysis*, 1994, vol. 15, no 2, p. 221-233.
- [19] Phillips, P. C., Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2), 335-346.
- [20] Woodward, W. A. et Gray, H. L. (1996). On the relationship between the S-array , and the box-jenkins method of ARMA model identification, *Journal of American Statistical Association*, vol 76, n375
- [21] Yves, T. (2004). Résumé du cours de séries temporelles.
- [22] Yves, A. (2011). *Séries temporelles avec R Methodes et cas*, Springer-Verlang, France.