

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

**UNIVERSITÉ MOHAMED KHIDER, BISKRA**

FACULTÉ des SCIENCES EXACTES et des SCIENCES de la NATURE et de la VIE

**DÉPARTEMENT DE MATHÉMATIQUES**



Mémoire présenté en vue de l'obtention du Diplôme :

**MASTER en Mathématiques**

Option : **Statistique**

Par

**CHAOUCHKHOUANE Meriem**

Titre :

**Méthode de Box et Jenkins**

Membres du Comité d'Examen :

<b>Pr. YAHIA Djabrane</b>	<b>UMKB</b>	<b>Président</b>
<b>Dr. BENAMEUR Sana</b>	<b>UMKB</b>	<b>Encadreur</b>
<b>Dr. BENBRAIKA Ghazlene</b>	<b>UMKB</b>	<b>Examineur</b>

Juin 2021

## DÉDICACE

À la mémoire de mon père

À ma mère

À ma famille

## REMERCIEMENTS

Je remercie d'abord et avant tout **ALLAH** le tout puissant qui m'a donné le courage, la volonté et la patience pour continuer mes études et pour réaliser ce travail.

Tout d'abord, ce travail ne serait pas aussi riche et n'aurait pas pu avoir le jour sans l'aide et l'encadrement de **Dr. Benameur Sana**, je la remercie pour la qualité de son encadrement exceptionnel, pour sa patience, sa rigueur et sa disponibilité durant mon préparation de ce mémoire.

Je tiens à remercier particulièrement le **Pr. Yahia Djabrane** pour ses conseils bénéfiques et son apport précieux tout au long de la réalisation de ce mémoire. Il n'a ménagé aucun effort pour m'aider dans ma recherche.

Je tiens également à adresser mes sincères remerciements au **Pr. Necir Abdelhakim** pour l'aide qu'il m'a apportée et pour ses précieux conseils.

Un grand merci aux membres de jury : **Pr. Yahia Djabrane** et **Dr. Benbraïka Ghozlène** qui m'ont fait l'honneur d'examiner et d'évaluer mon travail.

Merci à tous et à toutes.

# Table des matières

<b>Remerciements</b>	ii
<b>Table des matières</b>	iii
<b>Table des figures</b>	vi
<b>Liste des tables</b>	vii
<b>Introduction</b>	1
<b>1 Modélisation des séries temporelles</b>	2
1.1 Concepts de base des séries temporelles . . . . .	2
1.1.1 Décomposition d'une série temporelle . . . . .	3
1.1.2 Schémas d'une série temporelle . . . . .	4
1.1.3 Opérateurs de différentiation et filtres . . . . .	5
1.1.4 Élimination et conservation de la tendance et la composante saisonnière . . . . .	6
1.1.5 Propriétés statistiques d'une série temporelle . . . . .	7
1.2 Modèles stationnaires . . . . .	8
1.3 Modèles linéaires . . . . .	10
1.4 Modèles SARIMA . . . . .	12
1.4.1 Caractéristiques d'un modèle AR . . . . .	12
1.4.2 Caractéristiques d'un modèle MA . . . . .	16
1.4.3 Caractéristiques d'un modèle ARMA . . . . .	18

1.4.4	Modèles ARIMA	20
1.4.5	Modèles SARIMA	20
<b>2</b>	<b>Prévision par la méthode de Box et Jenkins</b>	<b>22</b>
2.1	Identification	23
2.1.1	Estimation du paramètre d'intégration $d$	23
2.1.2	Estimation des ordres $p$ et $q$	25
2.2	Estimation des paramètres	28
2.3	Validation	30
2.3.1	Tests de significativité des paramètres	30
2.3.2	Test sur le Bruit Blanc	30
2.4	Choix du modèle	32
2.4.1	Critères de capacité prédictive	32
2.4.2	Critères d'information	33
2.5	Prévision	33
2.5.1	Prévision pour un modèle autorégressif	34
2.5.2	Prévision pour un modèle moyenne mobile	34
2.5.3	Prévision pour un modèle ARMA	36
2.5.4	Prévision pour un modèle ARIMA	37
<b>3</b>	<b>Application sous R</b>	<b>38</b>
3.1	Etape 1 : Identification	38
3.2	Etape 2 : Estimation des paramètres	40
3.3	Etape 3 : Validation	41
3.4	Etape 4 : Choix du modèle	42
3.5	Etape 5 : Prévision	43
3.6	Mise en œuvre sous R	44
	<b>Conclusion</b>	<b>46</b>

<a href="#">Bibliographie</a>	46
<a href="#">Annexe A : Abréviations et Notations</a>	49
<a href="#">Annexe B : Packages et Commandes</a>	51

# Table des figures

1.1	Décomposition d'une série temporelle ; le graphe de la série (en haut), la tendance et la saisonnalité (resp.) (au milieu) et le résidu (en bas).	3
1.2	Illustrations des schémas d'une série temporelle; schéma additif (à gauche), schéma multiplicatif (au milieu) et schéma mixte (à droite).	4
1.3	Représentation graphique d'une trajectoire (en haut), autocorrélogramme (au milieu) et autocorrélogramme partielle (en bas) d'un modèle $AR(1)$ (à gauche) et d'un modèle $AR(2)$ (à droite).	15
1.4	Représentation graphique d'une trajectoire (en haut) et l'autocorrélogramme (en bas) d'un modèle $MA(1)$ (à gauche) et d'un modèle $MA(2)$ (à droite).	17
2.1	Autocorrélogramme de la série de production laitière aux États-Unis.	23
3.1	Représentation graphique de la série stationnaire $\Delta(\Delta_{12})X_t$ .	39
3.2	ACF et PACF de la série stationnaire $\Delta(\Delta_{12})X_t$ .	39
3.3	Analyse des résidus standardisés pour le modèle $SARIMA(0, 1, 0)(3, 1, 0)_{12}$ .	41
3.4	QQ-plot des résidus du modèle : $SARIMA(0, 1, 0)(3, 1, 0)_{12}$ et la loi normale.	42
3.5	Représentation graphique des valeurs prédictives de la série "milk" sur une période de 4 ans à l'aide du modèle $SARIMA(0, 1, 0)(3, 1, 0)_{12}$ .	43

# Liste des tableaux

3.1	Résumé des résultats obtenus à partir de l'estimation des paramètres . . . . .	40
3.2	Résumé des résultats obtenus à partir du test de significativité des coefficients .	41
3.3	Résumé des résultats obtenus pour les valeurs des critères AIC et BIC . . . . .	42

# Introduction

Une série temporelle (ou série chronologique) est une séquence spécifique d'observations repérés dans le temps. Leur domaine d'application est très vaste et s'étend de l'astronomie à l'économie et la finance en passant par la biologie, la géophysique, et la médecine, ...

La théorie et la pratique de l'analyse des séries temporelles se sont développées rapidement depuis l'apparition des travaux fondateurs de Box et Jenkins en 1970 [3]. L'utilisation des séries chronologiques peut servir à diverses fins. L'objectif peut être la prédiction de l'avenir basée sur la connaissance du passé, il peut être le contrôle du processus produisant la série, ou tout simplement une brève description des principales caractéristiques de la série.

Ce mémoire porte sur la méthode de prévision de Box et Jenkins développée en 1976 [4]. Cette méthode constitue un outil systématique qui permet de déterminer un modèle de type *ARMA* restituant le mieux possible le comportement d'une série chronologique observée.

L'organisation de ce mémoire est la suivante :

**Chapitre 1 (Modélisation des séries temporelles) :** Dans ce chapitre, plusieurs notions de base en séries temporelles sont abordées et quelques modèles de séries chronologiques sont présentés telles que : *AR*, *MA*, *ARMA*, *ARIMA* et *SARIMA*.

**Chapitre 2 (Prévision par la méthode de Box et Jenkins) :** La procédure de la méthode de Box et Jenkins est décrite en détail, et les principales étapes de cette technique sont abordées dans ce chapitre.

**Chapitre 3 (Application sous R) :** Ce chapitre a pour but d'appliquer la méthode de Box et Jenkins sur la série de production laitière mensuelle aux États-Unis pour atteindre le meilleur modèle de prévision, en utilisant le logiciel R.

Enfin, on achève le travail par une brève conclusion.

# Chapitre 1

## Modélisation des séries temporelles

Le premier objectif de ce chapitre sera de fournir un aperçu sur les concepts fondamentaux des séries temporelles. Nous introduisons par la suite quelques notions de base et définitions sur la linéarité et la stationnarité. De plus, nous présentons certaines propriétés importantes sur les différents modèles des séries chronologiques, en particulier, les modèles : *AR*, *MA*, *ARMA*, *ARIMA* et *SARIMA*. Pour plus d'informations vous pouvez consulter les livres de Brockwell & Davis [6], [7] et Anderson [1].

### 1.1 Concepts de base des séries temporelles

#### Définition 1.1.1 (Série temporelle (ou chronologique))

*Une série temporelle est une suite d'observations indexées par le temps, notée  $X_t$ . Elle représente généralement l'évolution d'un phénomène au cours du temps, telle que :*

$$X_t := \{X_t, t = 1, \dots, T\}, T \in \mathbb{N}, \mathbb{R}, \mathbb{Z}, \dots$$

*L'indice  $t$  représente une unité de temps (qui peut être le mois, l'année, ...).*

**Remarque 1.1.1** *Une série chronologique peut être à temps continu ou à temps discret. Dans le cas continu, les observations sont mesurées à chaque instant de temps ( $T$  est un intervalle de  $\mathbb{R}$ ). Par contre, la série chronologique à temps discret contient des observations mesurées à des points discrets de temps (en général  $T \subset \mathbb{Z}$ ).*

On désignera dans la suite par série temporelle une série temporelle à temps discret.

### 1.1.1 Décomposition d'une série temporelle

Les séries temporelles sont modélisées selon les trois composantes suivantes :

**1- La tendance** : La tendance, notée  $Z_t$ , est une fonction représente le mouvement à long terme de la série étudiée. Il existe trois types de tendance.

**2- La saisonnalité** : La composante saisonnière, notée  $S_t$ , est une fonction périodique de période  $s$ , i.e.  $S_t = S_{t \pm s}$ . Elle est liée au rythme imposé par les saisons.  $S_t$  est entièrement déterminée par ses  $s$  composantes  $\gamma_1, \dots, \gamma_s$ , et par hypothèse  $\sum_{j=1}^s \gamma_j = 0$ , telle que :

$$S_t := \sum_{j=1}^s \gamma_j S^j, \text{ où } S^j = \begin{cases} 1 & \text{si } j = t, \\ 0 & \text{si } j \neq t. \end{cases}$$

**3- Le résidu** : On appelle résidu (ou erreur) ou partie irrégulière, noté  $\varepsilon_t$ , la variable aléatoire à effet de faible intensité et à courte durée.

La figure [1.1](#) illustre la décomposition d'une série temporelle réelle. Il s'agit de la production laitière mensuelle aux États-Unis (en millions de livres) de janvier 1994 à décembre 2005.

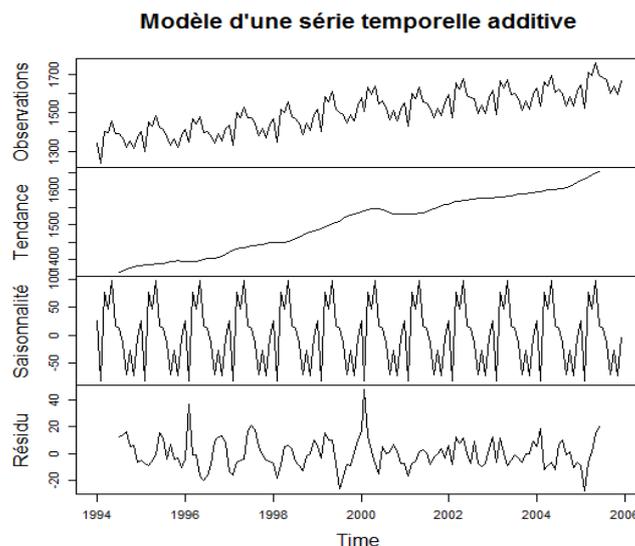


FIG. 1.1 – Décomposition d'une série temporelle ; le graphe de la série (en haut), la tendance et la saisonnalité (resp.) (au milieu) et le résidu (en bas).

### 1.1.2 Schémas d'une série temporelle

La décomposition d'une série chronologique repose sur un modèle qui porte le nom de schéma. Il existe essentiellement trois grands types des schémas :

**1. Schéma additif :** La série  $X_t$  s'écrit comme une somme de ses trois composantes :

$$X_t := Z_t + S_t + \varepsilon_t, \quad t = 1, \dots, T.$$

**2. Schéma multiplicatif :** Le modèle de la série s'écrit uniquement par un produit des composantes de  $X_t$  :

$$X_t := Z_t \times S_t \times \varepsilon_t, \quad t = 1, \dots, T.$$

**Remarque 1.1.2** Si le modèle étudié est multiplicatif, un passage au log permet de se ramener à un modèle additif :  $\log(X_t) = \log(Z_t \times S_t \times \varepsilon_t) = \log(Z_t) + \log(S_t) + \log(\varepsilon_t)$ .

**3. Schéma mixte :** Il s'écrit comme des différentes combinaisons de ces trois composantes. On donne comme exemples :

$$X_t = Z_t \times (S_t + \varepsilon_t), \quad t = 1, \dots, T.$$

$$X_t = Z_t \times S_t + \varepsilon_t, \quad t = 1, \dots, T.$$

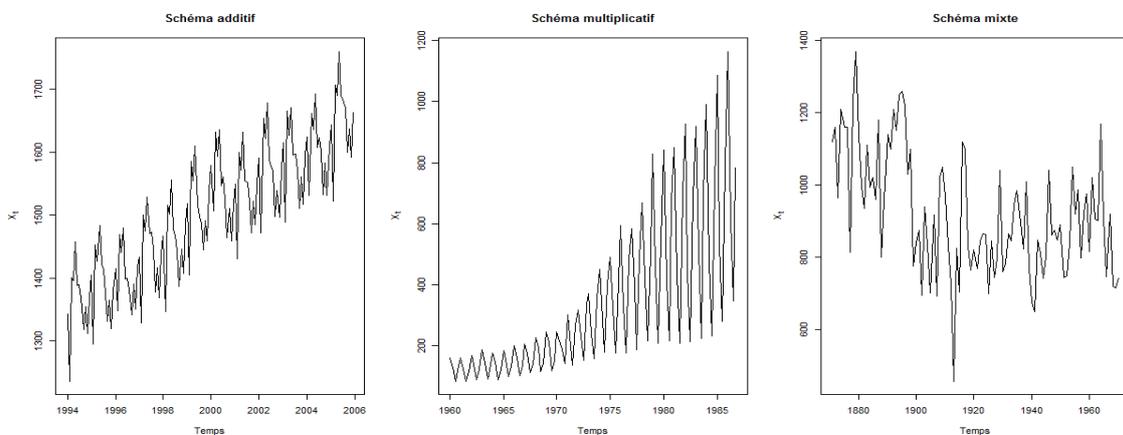


FIG. 1.2 – Illustrations des schémas d'une série temporelle; schéma additif (à gauche), schéma multiplicatif (au milieu) et schéma mixte (à droite).

L'examen visuel du graphique de gauche de la figure [1.2](#), indique que la trajectoire de la série augmente de manière régulière (ou l'amplitude de la série est plus constante) au cours du temps. Au milieu, la courbe de la série croît de façon exponentielle et l'amplitude de la série n'est plus constante au cours du temps. La dernière figure à droite, représente une série avec une fluctuation irrégulière.

### 1.1.3 Opérateurs de différentiation et filtres

#### Définition 1.1.2 (Opérateurs de retard et d'avance)

L'opérateur de retard, noté  $L$  (*Lag*) (ou *B Backward*), sur une série temporelle est défini par :

$$LX_t = X_{t-1}, \quad \forall t \in \mathbb{N}.$$

Cet opérateur a pour fonction de transformation d'une variable aléatoire  $X_t$  en sa valeur passée.

On note de manière générale :

$$L^k X_t = X_{t-k}, \quad \forall t \in \mathbb{N} \text{ et } \forall k \in \mathbb{N}^*.$$

L'opérateur d'avance, noté  $F$  (*Forward*), sur une série temporelle est défini par :

$$FX_t = L^{-1}X_t = X_{t+1}, \quad \forall t \in \mathbb{N}.$$

De façon plus générale :

$$F^k X_t = L^{-k} X_t = X_{t+k}, \quad \forall t \in \mathbb{N} \text{ et } k \in \mathbb{N}^*.$$

#### Définition 1.1.3 (Opérateurs de différentiation)

L'opérateur de différentiation linéaire d'une série  $X_t$ , noté  $\Delta$ , est défini par :

$$\Delta X_t = (1 - L)X_t = X_t - X_{t-1}, \quad \forall t \in \mathbb{N}.$$

L'opérateur de différentiation d'ordre  $k$  correspondant est :

$$\Delta^k X_t = (1 - L)^k X_t, \quad \forall t \in \mathbb{N} \text{ et } k \in \mathbb{N}^*.$$

**Définition 1.1.4 (Filtres)**

Les filtres (ou les moyennes mobiles), notées  $M(L)$ , sont des opérateurs qui permettent d'éliminer et/ou de conserver une tendance  $Z_t$  et/ou une composante saisonnière  $S_t$ , d'une série temporelle  $X_t$ . Ils s'écrivent :

$$M(L) = \sum_{j=-m_1}^{m_2} \theta_j L^j, \quad \theta_j \in \mathbb{R} \text{ et } (m_1, m_2) \in \mathbb{N}^* \times \mathbb{N}^*.$$

1. Le degré du polynôme caractéristique de  $M(L)$  est :  $m = m_1 + m_2$ .
2. L'ordre de la moyenne mobile  $M(L)$  est :  $k = m + 1$ .
3. La moyenne mobile est dite :
  - **normalisée** : si  $\sum_{j=-m_1}^{m_2} \theta_j = 1$ .
  - **centrée** : si  $m_1 = m_2$ .
  - **symétrique** : si  $\theta_j = \theta_{-j}$ , pour  $j = 1, \dots, m$ .
  - **arithmétique** : si les coefficients  $\theta_j$  sont égaux  $\forall j$ .

**1.1.4 Élimination et conservation de la tendance et la composante saisonnière**

1) L'opérateurs  $\Delta^d$  permet d'éliminer les tendances polynômiales de degré  $d$

$$Z_t = \sum_{j=0}^d a_j t^j = a_0 + a_1 t + a_2 t^2 + \dots + a_d t^d,$$

d'un modèle additif.

2) L'opérateurs  $\Delta_s = (1 - L^s)$  permet d'éliminer une composante saisonnière de période  $s$  d'un modèle additif.

3) Les moyennes mobiles arithmétiques d'ordre  $2d + 1$ , telles que

$$M_{2d+1}(L) = \frac{1}{2d+1}(1 + L + \dots + L^{2d}),$$

laissent invariante les tendances polynômiales de degré  $d$ .

4) Pour enlever une composante  $S_t$  d'une série temporelle, on utilise des filtres d'ordre égale à la période  $s$  de  $S_t$  :

- Si  $s = 2k + 1$ . On utilise le filtre suivant :

$$M_1(L) = \frac{1}{2k+1}(L^k + \dots + L^1 + I + L^{-1} + \dots + L^{-k}).$$

- De même, si  $s = 2k$ . Le filtre utilisé est :

$$M_2(L) = \frac{1}{2k}(\frac{1}{2}L^k + \dots + L^1 + I + L^{-1} + \dots + \frac{1}{2}L^{-k}).$$

6) L'opérateur  $\Delta_s^{d+1} = (1 - L^s)^{d+1}$  permet d'éliminer une tendance polynômiale de degré  $d$  et une composante saisonnière de période  $s$  d'un modèle mixte de la forme :

$$X_t = (a_0 + a_1t + \dots + a_d t^d) \times S_t + \varepsilon_t.$$

### 1.1.5 Propriétés statistiques d'une série temporelle

#### Définition 1.1.5 (Espérance mathématique)

L'espérance d'une série chronologique  $X_t$  est une fonction donnée par :

$$\mathbb{E}(X_t) := \mu_t, \text{ avec } t = 1, 2, \dots, T.$$

#### Définition 1.1.6 (Variance)

La variance d'une série chronologique  $X_t$  est définie par :

$$\text{var}(X_t) := \mathbb{E}(X_t - \mathbb{E}(X_t))^2 = \mathbb{E}(X_t^2) - \mathbb{E}^2(X_t), \text{ avec } t = 1, 2, \dots, T.$$

**Définition 1.1.7 (Fonction d'autocovariance)**

Si  $X_t$  est une série temporelle, telle que  $\text{var}(X_t) < +\infty, \forall t$ , alors la fonction d'autocovariance de  $X_t$ , notée  $\gamma_X(\cdot, \cdot)$ , est définie par :

$$\gamma_X(t, s) := \text{cov}(X_t, X_s) = \mathbb{E}[(X_t - E(X_t))(X_s - E(X_s))].$$

Elle vérifiée :

$$\gamma_X(t, s) = \gamma_X(|t - s|), \text{ avec } t, s = 1, 2, \dots, T.$$

**Définition 1.1.8 (Fonction d'autocorrélation)**

La fonction d'autocorrélation, notée  $\rho_X(\cdot, \cdot)$ , prend ses valeurs dans  $[-1; 1]$ , est donnée par :

$$\rho_X(t, s) := \text{corr}(X_t, X_s) = \frac{\gamma_X(t, s)}{(\text{var}(X_t))^{\frac{1}{2}}(\text{var}(X_s))^{\frac{1}{2}}}, \text{ avec } t, s = 1, 2, \dots, T.$$

## 1.2 Modèles stationnaires

**Définition 1.2.1 (Bruit Blanc)**

Un bruit blanc est une suite de variables aléatoires indépendantes et identiquement distribuées (iid), notée  $\varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2)$ , vérifiée les propriétés suivantes :

$$\mathbb{E}(\varepsilon_t) = 0, \text{ Var}(\varepsilon_t) = \sigma_\varepsilon^2 \text{ et } \text{Cov}(\varepsilon_t, \varepsilon_s) = 0, \text{ si } t \neq s.$$

**Définition 1.2.2 (Stationnarité strict)**

On dit qu'une série chronologique  $X_t$  est strictement (ou fortement) stationnaire si les distributions jointes de  $(X_{t_1}, \dots, X_{t_k})$  et  $(X_{t_1+h}, \dots, X_{t_k+h})$  sont les mêmes, pour tous entier positif  $k$  et  $\forall (t_1, \dots, t_k) \in T^k, h \in \mathbb{Z}$ .

**Définition 1.2.3 (Stationnarité faible)**

Une série chronologique carré intégrable, i.e.  $\mathbb{E}(|X_t|^2) < +\infty$ , est dite stationnaire (ou faiblement stationnaire ou bien stationnarité du second ordre), si son espérance  $\mathbb{E}(X_t)$  et ses autocovariances sont invariante par translation dans le temps.

- i)  $\mathbb{E}(X_t) = \mu, \forall t \in T,$   
 ii)  $cov(X_t, X_s) = cov(X_{t+h}, X_{s+h}), \forall (t, s) \in T^2, \forall h \in \mathbb{Z}.$

Dans la suite, on notera la fonction d'autocovariance et d'autocorrélation (resp.), d'une série stationnaire par :

$$\gamma_X(h) = cov(X_t, X_{t+h}) \text{ et } \rho_X(h) = corr(X_t, X_{t+h}) = \frac{\gamma_X(h)}{\gamma_X(0)}.$$

**Exemple 1.2.1** *Un bruit blanc est stationnaire.*

**Propriétés 1.2.1**

1. Si la série chronologique  $X_t$  peut s'écrire sous la forme  $\Phi(L)X_t = \Theta(L)\varepsilon_t + c$ , où :  $\Phi$  et  $\Theta$  sont des polynômes en  $L$ . Alors, on dit que  $X_t$  est stationnaire si toutes les racines de  $\Phi(z) = 0$  sont de module supérieur strictement à 1.
2. Il est clair que la stationnarité forte implique la stationnarité faible, mais la réciproque est fautive.
3. Soit  $Y_t$  un modèle stationnaire centrée à variance finie. Soit une suite  $(\phi_j)_{j \in \mathbb{Z}}$  telle que

$\sum_{j \in \mathbb{Z}} |\phi_j| < +\infty$ . La série  $X_t$  où  $X_t = \sum_{j \in \mathbb{Z}} \phi_j Y_{t-j} = \Phi(L)Y_t$ , est stationnaire centrée de fonction d'autocovariance  $\gamma_X(h) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \phi_j \phi_k \gamma_Y(h+k-j), \forall h \in \mathbb{Z}.$

**Proposition 1.2.1 (Propriétés élémentaires)**

- i) *La fonction d'autocovariance d'une série temporelle stationnaire  $X_t$ , vérifie les propriétés suivantes :*

$$1) \gamma_X(0) \geq 0, \quad 2) |\gamma_X(h)| \leq \gamma_X(0), \quad 3) \gamma_X(h) = \gamma_X(-h), \forall h \in \mathbb{N}.$$

- ii) *La fonction d'autocorrélation d'une série temporelle stationnaire  $X_t$ , vérifie des propriétés similaires :*

$$1) \rho_X(0) = 1, \quad 2) |\rho_X(h)| \leq 1, \quad 3) \rho_X(h) = \rho_X(-h), \forall h \in \mathbb{N}.$$

**Propriétés 1.2.2 (Estimation des propriétés statistiques d'une série temporelle)**

Soit  $X_t$  une série chronologique stationnaire :

1) L'estimateur de l'espérance est la moyenne empirique :

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

2) La fonction d'autocovariance est estimée par :

$$\hat{\gamma}_X(h) = \frac{1}{T-h} \sum_{t=h+1}^T (X_t - \bar{X})(X_{t-h} - \bar{X}).$$

3) On en déduit l'estimateur de la fonction d'autocorrélation :

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

## 1.3 Modèles linéaires

**Définition 1.3.1 (Modèle linéaire)**

On dit qu'un modèle  $X_t$  est linéaire s'il s'écrit comme suit :

$$X_t = c + \sum_{j \in \mathbb{Z}} \psi_j \varepsilon_{t-j}, \quad \forall t \in T, \tag{1.1}$$

où :  $c \in \mathbb{R}$ ,  $(\psi_j)_{j \in \mathbb{Z}}$  est une suite de constantes, telle que  $\sum_{j=-\infty}^{+\infty} |\psi_j| < +\infty$ , et  $\varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2)$ .

**Définition 1.3.2 (Modèle causal)**

On dit qu'un modèle  $X_t$  est causal s'il s'écrit en fonction de  $(\varepsilon_k, k \leq t)$  seulement :

$$X_t = \sum_{j=0}^{+\infty} \varphi_j \varepsilon_{t-j}, \quad \varphi_j \in \mathbb{R} \text{ et } \varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2).$$

**Définition 1.3.3 (Modèle inversible)**

On dit qu'un modèle  $X_t$  est inversible si  $\varepsilon_t$  s'écrit en fonction de  $(X_k, k \leq t)$  seulement :

$$\varepsilon_t = \sum_{j=0}^{+\infty} \phi_j X_{t-j}, \quad \phi_j \in \mathbb{R}.$$

**Propriétés 1.3.1**

1- Si la série  $X_t$  est causale, alors elle est linéaire. La réciproque n'est pas toujours vraie.

2- Si  $X_t$  peut s'écrire sous la forme :

$$\Phi(L)X_t = \Theta(L)\varepsilon_t.$$

où  $\Phi$  et  $\Theta$  sont des polynômes en  $L$ . Alors,  $X_t$  est dite inversible si toute les racines de  $\Theta(z) = 0$  sont de module supérieur strictement à 1.

3- La fonction d'autocovariance d'un modèle linéaire  $X_t$  est donnée par :

$$\gamma_X(h) = \sigma_\varepsilon^2 \sum_{j=0}^{+\infty} \psi_j \psi_{j+h}.$$

En particulier, en fixant  $h = 0$ , on constate que sa variance est :

$$\gamma_X(0) = \sigma_\varepsilon^2 \sum_{j=0}^{+\infty} \psi_j^2.$$

**Exemple 1.3.1 (Marche aléatoire)**

Une marche aléatoire  $X_t$  est définie par :

$$X_t = X_{t-1} + \varepsilon_t, \text{ avec } X_0 = 0, \text{ et } \varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2).$$

On remarque que  $X_t$  est inversible. De plus, elle est causale et linéaire. En effet :

$$X_t = X_{t-1} + \varepsilon_t = \sum_{j=0}^{t-1} \varepsilon_{t-j}.$$

## 1.4 Modèles SARIMA

Dans cette section, on considère les modèles linéaires proposés par Box et Jenkins et qui sont les plus fréquemment utilisés pour modéliser une série chronologique  $X_t$ , tels que le modèle autorégressif, le modèle à moyenne mobile et le modèle ARMA qui combine ces deux modèles.

### 1.4.1 Caractéristiques d'un modèle AR

#### Définition 1.4.1 (Modèle Autorégressif)

On dit que  $X_t$  est un modèle autorégressif ("Auto-Regressive") d'ordre  $p$  ( $p \geq 1$ ), abrégé  $AR(p)$ , si :

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t = \sum_{j=1}^p \phi_j X_{t-j} + \varepsilon_t, \quad (1.2)$$

où :  $\phi_1, \phi_2, \dots, \phi_p$  sont des constantes ( $\phi_p \neq 0$ ), et  $\varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2)$ . On utilise généralement la notation suivante :

$$\Phi(L)X_t = \varepsilon_t,$$

avec  $\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$ .

On dit dans ce cas, qu'un modèle  $AR(p)$  est stationnaire si toutes les racines du polynôme caractéristique  $\Phi(z)$  sont de module supérieur strictement à 1.

#### Propriétés 1.4.1 (Système de Yule-Walker)

Sous hypothèse de stationnarité, un modèle  $AR(p)$  vérifie les propriétés suivantes :

1) L'espérance mathématique :  $\mathbb{E}(X_t) = 0$ ,

2) La fonction d'autocovariance :

$$\gamma_X(h) = \sum_{j=1}^p \phi_j \gamma_X(h-j), \quad (1.3)$$

3) La fonction d'autocorrélation :

$$\rho_X(h) = \sum_{j=1}^p \phi_j \rho_X(h-j). \quad (1.4)$$

4) *La variance :*

$$\gamma_X(0) = \frac{\sigma_\varepsilon^2}{1 - \sum_{j=1}^p \phi_j \rho_X(j)}. \quad (1.5)$$

Les trois équations (1.3), (1.4) et (1.5) s'appellent le système de Yule-Walker.

**Remarque 1.4.1** *Si le modèle n'est pas centré, i.e.  $\Phi(L)X_t = c + \varepsilon_t$ , on obtient les mêmes résultats mais sur la translation  $Y_t = X_t - \mu$ . Dans ce cas :*

$$\mathbb{E}(X_t) = \mu = \frac{c}{1 - \phi_1 - \phi_2 - \dots - \phi_p} = \frac{1}{\Phi(1)} \neq 0, \quad \mathbb{E}(Y_t) = 0, \quad \gamma_X(h) = \gamma_Y(h), \quad \text{et } \rho_X(h) = \rho_Y(h).$$

**Propriété 1.4.1 (Système de Yule-Walker sous forme matricielle)**

*Les équations de Yule-Walker peuvent être écrites sous forme matricielle, telle que :*

$$\rho = \mathcal{R}_p \phi,$$

où :  $\rho = (\rho_X(1), \rho_X(2), \dots, \rho_X(p))^t$  et  $\phi = (\phi_1, \phi_2, \dots, \phi_p)^t$ , et :

$$\mathcal{R}_p = \begin{pmatrix} 1 & \rho_X(1) & \dots & \rho_X(p-1) \\ \rho_X(1) & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho_X(1) \\ \rho_X(p-1) & \dots & \rho_X(1) & 1 \end{pmatrix}. \quad (1.6)$$

*Cette dernière s'appelle matrice des autocorrélations, elle est de rang  $p$  et de forme Toeplitz.*

**Remarques 1.4.1**

1) *L'inversion de la matrice des autocorrélations  $\mathcal{R}_p$  nous permet de calculer les  $\phi_j$  en fonction de  $\rho_X(j)$  seulement :  $\phi = \mathcal{R}_p^{-1} \rho$ .*

2) *Le vecteur des paramètres  $\phi$  est estimé par :*

$$\hat{\phi} = \hat{\mathcal{R}}_p^{-1} \hat{\rho},$$

Les fonctions  $\gamma_X(h)$  et  $\rho_X(h)$  n'interviennent que les variables  $X_t$  et  $X_{t\pm h}$ , on introduit ici une notion faisant intervenir de plus les variables intermédiaires  $(X_{t\pm 1}, X_{t\pm 2}, X_{t\pm 3}, \dots, X_{t\pm(h-1)})$ .

**Définition 1.4.2 (Fonction d'autocorrélation partielle)**

La fonction d'autocorrélation partielle de  $X_t$ , notée  $\phi_h$ , est définie par :

$$\phi(h) = \text{corr} \left( X_t - \mathbb{E} \left( X_t \mid X_{t-1}, \dots, X_{t-h+1} \right); X_{t-h} - \mathbb{E} \left( X_{t-h} \mid X_{t-1}, \dots, X_{t-h+1} \right) \right), \forall h \in \mathbb{N}^*.$$

**Algorithme de Durbin**

Pour calculer les autocorrélations partielles d'un modèle  $AR(p)$  on utilise l'algorithme de Durbin suivant :

$$\begin{cases} \phi(0) = \rho_X(0) = 1, \\ \phi(1) = \rho_X(1), \\ \phi(h) = \frac{\det \mathcal{R}_h^*}{\det \mathcal{R}_h}, h \geq 2. \end{cases}$$

où  $\mathcal{R}_h^*$  est la matrice  $\mathcal{R}_h$ , définie précédemment par (1.6), juste en remplaçant la dernière colonne par :  $(\rho_X(1), \rho_X(2), \dots, \rho_X(h))^t$ .

**Propriété 1.4.2**

Pour un modèle  $AR(p)$ , les autocorrélations partielles sont nulle au-delà de rang  $p$ ,  $\phi(h) = 0$  si  $h > p$ . Cette propriété permet d'estimer l'ordre ou le degré  $p$  du modèle  $AR(p)$ .

**Exemples 1.4.1**

1. Soit  $X_t \sim AR(1)$ , tel que  $X_t = \phi X_{t-1} + \varepsilon_t$  avec  $\varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2)$ . Si  $|\phi| < 1$ , le modèle est stationnaire avec

$$\gamma_X(h) = \phi \gamma_X(h-1) = \phi^2 \gamma_X(h-2) = \dots = \phi^h \gamma_X(0), \quad h \geq 1.$$

$$\rho_X(h) = \phi \rho_X(h-1) = \phi^2 \rho_X(h-2) = \dots = \phi^h, \quad h \geq 1.$$

$$\gamma_X(0) = \frac{\sigma_\varepsilon^2}{1 - \phi \rho_X(1)} = \frac{\sigma_\varepsilon^2}{1 - \phi^2}.$$

Et par conséquent  $\phi(h) = \begin{cases} 1 & \text{si } h = 0, \\ \phi & \text{si } h = 1, \\ 0 & \text{si } h > 1. \end{cases}$

2. On considère un modèle  $AR(2)$  stationnaire, tel que  $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t$ . D'après le système de Yule-Walker, on peut trouver facilement :

$$\rho_X(1) = \frac{\phi_1}{1 - \phi_2} \text{ et } \rho_X(2) = \frac{\phi_1^2 - \phi_2^2 + \phi_2}{1 - \phi_2}$$

$$\gamma_X(0) = \frac{(1 - \phi_2) \sigma_\varepsilon^2}{1 - \phi_2 - \phi_1^2 - \phi_2 \phi_1^2 - \phi_2^2 + \phi_2^3}$$

Il résulte, d'après l'algorithme de Durbin :  $\phi(h) = \begin{cases} 1 & \text{si } h = 0, \\ \frac{\phi_1}{1 - \phi_2} & \text{si } h = 1, \\ \phi_2 & \text{si } h = 2, \\ 0 & \text{si } h > 2. \end{cases}$

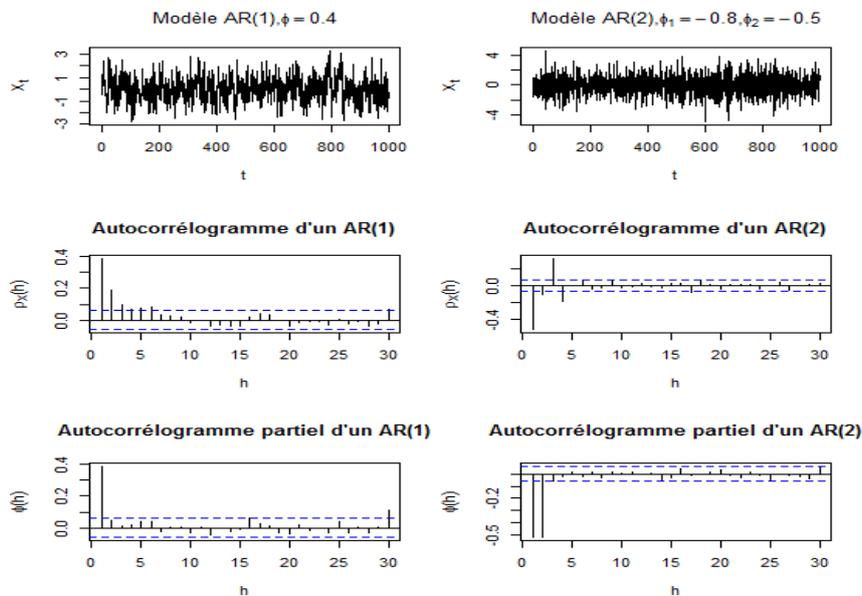


FIG. 1.3 – Représentation graphique d'une trajectoire (en haut), autocorrélogramme (au milieu) et autocorrélogramme partiel (en bas) d'un modèle  $AR(1)$  (à gauche) et d'un modèle  $AR(2)$  (à droite).

La figure (1.3) représente (en haut) les trajectoires de deux modèles stationnaires,  $AR(1)$ ;  $X_t = 0.4X_{t-1} + \varepsilon_t$  (à gauche) et  $AR(2)$ ;  $X_t = -0.8X_{t-1} - 0.5X_{t-2} + \varepsilon_t$  (à droite). Le graphique de la fonction d'autocorrélation estimée (ou autocorrélogramme) pour le modèle  $AR(1)$  (au milieu) décroît de façon exponentielle, alors qu'il alterne en signe pour le modèle  $AR(2)$ . Le graphique de la fonction d'autocorrélation partielle estimée (ou autocorrélogramme partiel) (en bas) présente exactement un pic au premier retard ( $h = 1$ ) pour le modèle  $AR(1)$ , et deux pics aux retards ( $h = 1, 2$ ) pour le modèle  $AR(2)$ .

## 1.4.2 Caractéristiques d'un modèle MA

### Définition 1.4.3 (Modèle moyenne mobile)

On appelle modèle moyenne mobile ("Moving Average") d'ordre  $q$ , abrégé  $MA(q)$ , un modèle stationnaire  $X_t$  qui vérifié :

$$X_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} = c + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i},$$

où :  $c \in \mathbb{R}$ ,  $\theta_1, \dots, \theta_q \in \mathbb{R}$  ( $\theta_q \neq 0$ ) et  $\varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2)$ . Cette relation est équivalente à l'écriture :

$$X_t = \Theta(L)\varepsilon_t,$$

où :  $\Theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q$ .

### Remarque 1.4.2

Les modèles  $MA(q)$  sont toujours stationnaires de moyenne  $c$ , il est possible aussi d'inverser les polynômes  $\Theta(L)$  dans le cas où ses racines sont de module différent de 1. Mais, le modèle est inversible seulement si ces racines sont de module supérieur strictement à 1.

### Propriété 1.4.3 (Fonction d'autocovariance et d'autocorrélation)

Un modèle  $MA(q)$  vérifie les propriétés suivantes :

1) La variance :  $\gamma_X(0) = (1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2) \sigma_\varepsilon^2$ .

2) La fonction d'autocovariance :

$$\gamma_X(h) = \begin{cases} (\theta_h + \theta_{h+1}\theta_1 + \cdots + \theta_q\theta_{q-h}) \sigma_\varepsilon^2 & \text{si } 1 \leq h \leq q, \\ 0 & \text{si } h > q. \end{cases}$$

3) La fonction d'autocorrélation :

$$\rho_X(h) = \begin{cases} \frac{\theta_h + \theta_{h+1}\theta_1 + \cdots + \theta_q\theta_{q-h}}{1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2} & \text{si } 1 \leq h \leq q, \\ 0 & \text{si } h > q. \end{cases}$$

**Propriété 1.4.4**

La fonction d'autocorrélation d'un modèle  $MA(q)$  s'annule au-delà du rang  $q$  ( $h > q$ ). Cette propriété sera relativement pratique pour faire l'estimation de l'ordre du modèle  $MA(q)$ .

**Exemple 1.4.1 (Modèle  $MA(1)$ )**

Soit  $X_t = \varepsilon_t + \theta\varepsilon_{t-1}$  avec  $\varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2)$ . Il vérifie les propriétés suivantes :

$$\mathbb{E}(X_t) = 0, \quad \gamma_X(0) = (1 + \theta^2)\sigma_\varepsilon^2, \quad \gamma_X(h) = \begin{cases} \theta\sigma_\varepsilon^2 & \text{si } h = 1, \\ 0 & \text{si } h > 1. \end{cases}$$

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \begin{cases} 1 & h = 0, \\ \frac{\theta}{(1 + \theta^2)} & h = 1, \\ 0 & h > 1. \end{cases}$$

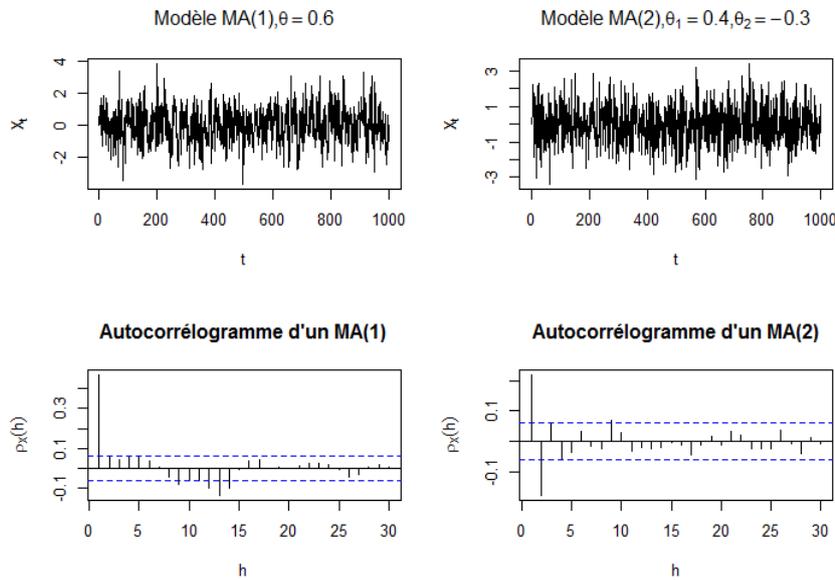


FIG. 1.4 – Représentation graphique d’une trajectoire (en haut) et l’autocorrélogramme (en bas) d’un modèle  $MA(1)$  (à gauche) et d’un modèle  $MA(2)$  (à droite)

Cette figure illustre les trajectoires (en haut) des deux modèles stationnaires,  $MA(1)$ ;  $X_t = 0.6\varepsilon_{t-1} + \varepsilon_t$  (à gauche) et  $MA(2)$ ;  $X_t = 0.4\varepsilon_{t-2} - 0.3\varepsilon_{t-1} + \varepsilon_t$  (à droite). Les autocorrélogrammes (en bas) présentent un seul pic au premier retard ( $h = 1$ ) pour le modèle  $MA(1)$ , et deux pics aux retards ( $h = 1, 2$ ) pour le modèle  $MA(2)$ .

**Définition 1.4.4 (Moyenne mobile d'ordre infini  $MA(\infty)$ )**

On dit qu'un modèle  $X_t$  est  $MA(\infty)$ , s'il peut s'écrire sous la forme :

$$X_t = \sum_{j=-\infty}^{+\infty} \varphi_j \varepsilon_{t-j}, \quad \varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2).$$

De plus, on peut dire qu'un modèle  $MA(\infty)$  est un modèle linéaire, encore si  $\varphi_j = 0, \forall j < 0$  :

$$X_t = \sum_{j=0}^{+\infty} \varphi_j \varepsilon_{t-j}, \quad \text{avec } t \in \mathbb{N},$$

on dit dans ce cas que le modèle  $X_t$  est causal.

**Théorème 1.4.1**

Tout modèle  $X_t$  centré et stationnaire (ou seconde ordre) peut être représenté sous une forme proche de la forme  $MA(\infty)$

$$X_t = \sum_{j=-\infty}^{+\infty} \varphi_j \varepsilon_{t-j}, \quad \forall t \in \mathbb{N},$$

où :  $\varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2)$ .

**1.4.3 Caractéristiques d'un modèle ARMA**

**Définition 1.4.5 (Modèle ARMA)**

Un modèle  $X_t$  est dit autorégressif moyenne mobile ("Auto-Regressive Moving Average"), abrégé  $ARMA(p, q)$ , s'il vérifie, pour tout  $t$  :

$$\begin{aligned} X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} &= c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \\ X_t - \sum_{i=1}^p \phi_i X_{t-i} &= c + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j}, \end{aligned} \tag{1.7}$$

où :  $\varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2)$ , et  $\phi_i, \theta_j$  sont des réels. L'équation (1.7) peut être écrite symboliquement sous la forme :

$$\Phi(L)X_t = c + \Theta(L)\varepsilon_t,$$

avec  $\Phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$  et  $\Theta(L) = 1 + \theta_1 L + \dots + \theta_q L^q$ .

**Remarques 1.4.2**

- On suppose que les polynômes  $\Phi$  et  $\Theta$  ont tous les racines de module strictement supérieure à 1, et n'ont pas des racines communes.
- On suppose aussi, que les degrés de  $\Phi$  et  $\Theta$  sont respectivement  $p$  et  $q$ , au sens où  $\phi_p \neq 0$  et  $\theta_q \neq 0$ . On dit dans ce cas que cette écriture est minimale.
- Un modèle  $AR(p)$  est un  $ARMA(p, 0)$ , et un modèle  $MA(q)$  est un  $ARMA(0, q)$ .
- On dit que  $X_t$  est un  $ARMA(p, q)$  de moyenne  $\mu$  si  $Y_t = X_t - \mu$  est un  $ARMA(p, q)$  centré, avec  $\mu = \frac{c}{\Phi(1)}$ . Alors, on obtient :  $\Phi(L)Y_t = \Theta(L)\varepsilon_t$ . Ce changement de variable nous permet de simplifier les calculs, tel que :

$$\gamma_X(h) = \mathbb{E}[(X_t - \mu)(X_{t-h} - \mu)] = \mathbb{E}(Y_t Y_{t-h}) = \gamma_Y(h), \quad \forall h \geq 0.$$

**Propriété 1.4.5 (Fonction génératrice d'autocovariance)**

La fonction génératrice d'autocovariance, notée  $G(z)$  pour un modèle stationnaire  $X_t$  permet de calculer les  $\gamma_X(h)$  en spécifiant les coefficients de  $z^h$ ,

$$G(z) = \sum_{h=-\infty}^{+\infty} \gamma_X(h) z^h = \sigma_\varepsilon^2 \psi(z) \psi(z^{-1}), \quad \text{où } \psi(z) = \frac{\Theta(z)}{\Phi(z)}.$$

**Exemple 1.4.2**

Soit  $X_t \sim ARMA(0, 2)$ , tel que :

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} = \Theta(L) \varepsilon_t,$$

où :  $\Theta(L) = 1 + \theta_1 L + \theta_2 L^2$ .

$$\begin{aligned} G(z) &= \sigma_\varepsilon^2 \psi(z) \psi(z^{-1}) \\ &= \sigma_\varepsilon^2 \Theta(z) \Theta(z^{-1}) \\ &= \sigma_\varepsilon^2 [(1 + \theta_1^2 + \theta_2^2) + (\theta_1 + \theta_1 \theta_2)(z + z^{-1}) + \theta_2(z^2 + z^{-2})]. \end{aligned}$$

Par conséquent

$$\begin{aligned}\gamma_X(0) &= (1 + \theta_1^2 + \theta_2^2) \sigma_\varepsilon^2, \\ \gamma_X(1) &= (\theta_1 + \theta_1\theta_2) \sigma_\varepsilon^2, \\ \gamma_X(2) &= \theta_2\sigma_\varepsilon^2, \\ \gamma_X(h) &= 0, \quad \forall h > 2.\end{aligned}$$

### 1.4.4 Modèles ARIMA

#### Définition 1.4.6 (Modèle ARMA intégré)

On dit que  $X_t$  est un modèle ARMA intégré ("Auto-Regressive Integrated Moving Average"), abrégé ARIMA( $p, d, q$ ), si  $\Delta^d X_t$  est un ARMA( $p, q$ ).

En d'autre terme,  $X_t \sim \text{ARIMA}(p, d, q)$  s'il vérifie une équation du type :

$$\Phi(L)(1 - L)^d X_t = \Theta(L)\varepsilon_t, \quad \forall t, \quad \text{et } \varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2), \quad (1.8)$$

$$\text{où : } \begin{cases} \Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p, \quad \text{où } \phi_p \neq 0 \\ \Theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q, \quad \text{où } \theta_q \neq 0 \end{cases}$$

sont des polynômes dont les racines sont de module supérieur strictement à 1.

#### Propriétés 1.4.2

- 1) Un modèle de type ARIMA( $p, d, q$ ) est non stationnaire ; il contient des tendances de degré  $d$ .
- 2) Le paramètre  $d$  indique le nombre de différences prises pour obtenir la stationnarité.

### 1.4.5 Modèles SARIMA

#### Définition 1.4.7 (Modèle ARMA intégré avec une composante saisonnière)

Un modèle ARMA intégré avec une composante saisonnière ("Seasonal Auto-Regressive Integrated Moving Average") peut être obtenu sous forme de généralisation d'un modèle ARMA avec des termes d'ordre non saisonniers ( $p, d, q$ ) et des termes d'ordre saisonniers ( $P, D, Q$ ),

abrégé  $SARIMA(p, d, q)(P, D, Q)_s$ . Il peut s'écrire sous forme :

$$\Phi(L)\Phi_P(L^s)(1-L)^d(1-L^s)^D X_t = c + \Theta(L)\Theta_Q(L^s)\varepsilon_t, \quad \varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2), \quad c \in \mathbb{R}. \quad (1.9)$$

où :  $\Phi_P, \Theta_Q$  désignent des polynômes en  $L^s$  d'ordre  $P, Q$  (resp.), tels que :

$$\begin{cases} \Phi_P(L) = 1 - \beta_1 L^s - \beta_2 L^{2s} - \dots - \beta_P L^{Ps}, \text{ où } \beta_P \neq 0. \\ \Theta_Q(L) = 1 + \gamma_1 L^s + \gamma_2 L^{2s} + \dots + \gamma_Q L^{Qs}, \text{ où } \gamma_Q \neq 0. \end{cases}$$

### Propriété 1.4.6

Lors de l'ajustement des modèles  $SARIMA$ , on doit d'abord choisir des valeurs pour les deux ordres de différenciation, saisonniers ( $D$ ) et non saisonniers ( $d$ ), de manière à rendre la série stationnaire.

# Chapitre 2

## Prévision par la méthode de Box et Jenkins

Après avoir décrit divers modèles de séries temporelles, la question suivante qui nous intéresse est de savoir comment sélectionner un modèle approprié qui peut produire des prévisions précises basées sur une description des données d'une série chronologique, et comment déterminer les ordres de modèle optimaux ?

Les statisticiens George Box et Gwilym Jenkins en 1976 [4] ont proposés une démarche de prévision pour les séries temporelles univariées, fondée sur l'utilisation des modèles *ARIMA*. Leur concept est d'une importance fondamentale dans le domaine de l'analyse et de la prévision des séries chronologiques. Les références de base utilisées dans ce chapitre est le livre de Gouriéroux [17] et le livre de Charpentier [9].

La méthodologie de Box-Jenkins utilise une approche itérative en cinq étapes :

- ✓ Identification à priori.
- ✓ Estimation des modèles retenus.
- ✓ Validation.
- ✓ Choix du modèle.
- ✓ Prévision.

## 2.1 Identification

Avant d'appliquer les méthodes d'estimation et de sélection de modèle, il convient de représenter graphiquement la série chronologique observée et de faire une première analyse de ses éventuelles composantes : tendances, saisonnalités ou autres particularités. Si la série étudiée n'est pas stationnaire, il convient de la transformer pour obtenir un modèle stationnaire. L'étape suivante consiste à analyser l'autocorrélogramme (ACF) et l'autocorrélogramme partielle (PACF) afin de déterminer les ordres du modèle.

L'identification est basée généralement sur l'estimation des ordres  $(p, d, q)$  pour un modèle de type  $ARIMA(p, d, q)$ , comme suit :

### 2.1.1 Estimation du paramètre d'intégration $d$

#### 1) Approche empirique par l'autocorrélogramme

Supposons que le modèle  $X_t \sim ARIMA(p, d, q)$  avec  $d > 0$ . En pratique, si  $\hat{\rho}_X(h)$  est proche de 1 pour un grand nombre de retards  $h$ , alors c'est une indication que la série doit être différenciée pour la rendre stationnaire, comme le montre la figure 2.1. Autrement dit, on considère que la série a besoin d'une différence si les premiers  $\hat{\rho}(h)$  sont proches les uns des autres, même si  $\hat{\rho}(1)$  est assez différent de 1. Lorsque la décision de différencier une série a été prise, on applique le même critère à la série différenciée, afin de comprendre si une seconde différenciation est nécessaire, et ainsi de suite.

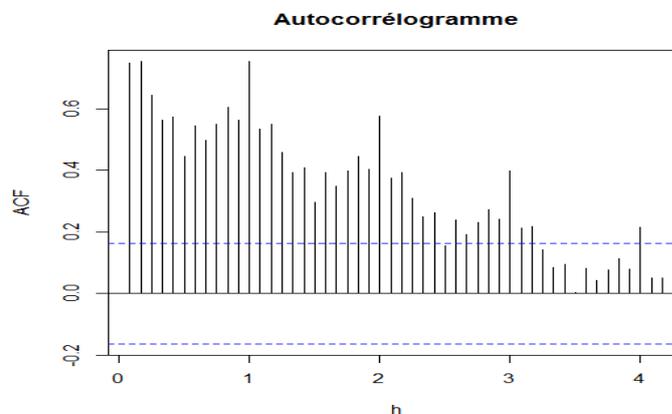


FIG. 2.1 – Autocorrélogramme de la série de production laitière aux États-Unis.

## 2) Test de racine unitaire

i) **Test de Dickey-Fuller simple (DF, 1979)** : Ce test permet de tester la non stationnarité d'un modèle  $AR(1)$  :

$$\begin{cases} H_0 : & \text{Le modèle suit une marche aléatoire,} \\ H_1 : & \text{Le modèle est un } AR(1). \end{cases}$$

Ces tests peuvent être regroupés en quatre cas possibles :

- $X_t = \phi X_{t-1} + \varepsilon_t$ . Sous l'hypothèse  $H_0 : \phi = 1$ .
- $X_t = \alpha + \phi X_{t-1} + \varepsilon_t$ . On teste  $H_0 : \alpha = 0$  et  $\phi = 1$ .
- $X_t = \alpha + \phi X_{t-1} + \varepsilon_t$ . On teste  $H_0 : \alpha \neq 0$  et  $\phi = 1$ .
- $X_t = \alpha + \beta t + \phi X_{t-1} + \varepsilon_t$ . On teste  $H_0 : \alpha = 0$  et  $\beta = 0$  et  $\phi = 1$ .

Le test de Dickey-Fuller, dans le premier cas, se construit comme un test de Student.

ii) **Test de Dickey-Fuller augmenté (ADF)** : Ce test permet de tester l'hypothèse :

$$\begin{cases} H_0 : & \text{Le modèle est intégré d'ordre au moins 1,} \\ H_1 : & \text{Le modèle suit un modèle } AR(1). \end{cases}$$

Ces tests peuvent être regroupés en trois cas possibles :

- $\Delta X_t = \phi X_{t-1} + \sum_{i=1}^p a_i \Delta X_{t-i} + \varepsilon_t$ .
- $\Delta X_t = \alpha + \phi X_{t-1} + \sum_{i=1}^p a_i \Delta X_{t-i} + \varepsilon_t$ .
- $\Delta X_t = \alpha + \beta t + \phi X_{t-1} + \sum_{i=1}^p a_i \Delta X_{t-i} + \varepsilon_t$ .

L'hypothèse  $H_0$  est ainsi testée à l'aide de la statistique de Student suivante :  $t_{\hat{\phi}} = \frac{\hat{\phi}}{\sqrt{\text{var}(\hat{\phi})}}$ .

Les valeurs critiques usuelles ne sont pas valides dans ce type de test. Les valeurs critiques à utiliser dans chacun des trois cas possibles ont été tabulées par Dickey et Fuller (on accepte  $H_0$  si  $t_{\hat{\phi}}$  est supérieur à  $t$  tabulé).

iii) **Test de Phillips et Perron** : Ces tests non paramétriques ont été introduits en 1988.

La distribution théorique à la base des tests DF repose sur l'hypothèse d'hétéroscédasticité du

bruit. La généralisation des tests DF consiste à ne plus supposer que  $\varepsilon_t$  est un Bruit Blanc, et à autoriser qu'il soit autocorrélé.

### 3) Tests de racines unitaires saisonnières

Dans le cas d'une modélisation *SARIMA*, avec une saisonnalité d'ordre  $s$ , il peut être intéressant de tester l'ordre  $s$ . Les tests les plus connues sont :

#### i) Test de Hasza et Fuller et de Osborn, Chui, Smith et Birchenhall (OCSB) :

Hasza et Fuller (1982) ont considérés le modèle :

$$X_t = \phi_1 X_{t-1} + \phi_s X_{t-s} + \phi_{s+1} X_{t-s-1} + \varepsilon_t, \text{ où } \varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2).$$

L'hypothèse à tester ici est  $H_0 : \phi_1 = \phi_s = -\phi_{s+1} = 1$ .

Osborn, chui, Smith et Birchenhall (1988) ont étendu cette approche sous la forme :

$$\Theta(L)(1-L)(1-L^s)X_t = \sum_{i=1}^s \delta_s D_{s,t} + \lambda(1-L^s)X_{t-1} + \mu(1-L)X_{t-s} + \varepsilon_t.$$

Si l'on accepte l'hypothèse  $\mu = 0$ , la différence à l'ordre  $s$  est appropriée, et si  $\lambda = \mu = 0$ , alors le filtre  $(1-L)(1-L^s)$  est nécessaire.

ii) **Test de Hylleberg, Engle, Granger et Yoo (HEGY 1990) :** Ce test utilise la décomposition des polynômes  $(1-L^4)(1-L^s)$ .

iii) **Test de Franses (1990) :** Ce test a mis en place pour tester une saisonnalité à l'ordre 12.

## 2.1.2 Estimation des ordres $p$ et $q$

Pour estimer les ordres  $p$  et  $q$ , nous utilisons les propriétés que nous avons vues dans le premier chapitre sur les formes des autocorrélogrammes et des autocorrélogrammes partielles.

### 1) Cas d'un modèle $MA(q)$

Les modèles  $MA(q)$  présentent exactement  $q$  pics aux  $q$  premières valeurs de l'ACF (voir la figure [1.4](#)).

On peut noter que la variance des autocorrélations empiriques (estimées) est donnée par la relation

$$\text{var}(\hat{\rho}_X(h)) = \sigma^2(\hat{\rho}_X(h)) = \frac{1}{T} \left( 1 + 2 \sum_{i=1}^q \rho_X^2(i) \right), \quad \forall h > q,$$

où :  $T$  est le nombre d'observations dans la série après différenciation  $d$  fois. Ainsi, l'estimateur de l'écart-type est

$$\hat{\sigma}(\hat{\rho}_X(h)) = \frac{1}{\sqrt{T}} \sqrt{1 + 2 \sum_{i=1}^q \hat{\rho}_X^2(i)}.$$

En fait, sous l'hypothèse nulle que le processus est un  $MA(q)$ , on a :

$$\forall h > q : \sqrt{T} \hat{\rho}_X(h) \xrightarrow{\mathcal{D}} \mathcal{N} \left( 0, 1 + 2 \sum_{i=1}^q \rho_X^2(i) \right).$$

Pratiquement, on identifie l'ordre  $q$  comme la première valeur à partir de laquelle les  $\hat{\rho}_X(h)$  sont dans l'intervalle à 95% dont les extrémités sont délimités par (lignes en tirets dans la figure

**1.4**) :

$$\pm \frac{1.96}{\sqrt{T}} \left( 1 + 2 \sum_{i=1}^q \hat{\rho}_X^2(i) \right)^{\frac{1}{2}}.$$

## 2) Cas d'un modèle $AR(p)$

Les modèles  $AR(p)$  présentent une ACF dont les valeurs décroissent exponentiellement avec des alternances possibles de valeurs positives et négatives. De plus, leur PACF présente exactement  $p$  pics aux  $p$  premières valeurs (voir la figure **1.3**).

De la même manière, en utilisant un résultat de Quenouille (1949), on peut prendre  $\frac{1}{\sqrt{T}}$  comme une erreur standard approximative de l'autocorrélation partielle empirique  $\hat{\phi}(h)$  pour  $h > p$  si le modèle est un  $AR(p)$ . Afin d'identifier le degré  $p$  nous pouvons représenter la séquence des autocorrélations partielles estimées  $\hat{\phi}(h)$  et examiner à partir de quelle valeur il reste dans l'intervalle à 95% (lignes en tirets dans la figure **1.3**) :

$$\left[ -\frac{1.96}{\sqrt{T}}, \frac{1.96}{\sqrt{T}} \right].$$

Une telle étape consistant à tester séparément l'hypothèse nulle d'autocorrélation partielle nulle  $\phi(h)$  pour chaque retard  $h$  au seuil de signification de 5% n'est pas strictement rigoureuse d'un

point de vue statistique, mais fournit une ligne directrice utile en pratique.

### 3) Cas d'un modèle $ARMA(p, q)$

On peut identifier les ordres  $p$  et  $q$  à partir de l'autocorrélation partielle qui s'annule au-delà de l'ordre  $p$ , et l'autocorrélation qui s'annule au-delà de l'ordre  $q$ . On peut aussi appliquer une méthode dite "la méthode du coin" qui permet d'estimer conjointement  $p$  et  $q$  lorsque les deux sont non-nuls. Elle est basée sur les propriétés suivantes :

- i)  $\Delta(i, j) = 0, \forall i \geq p \text{ et } j \geq q,$
- ii)  $\Delta(i, q - 1) \neq 0, \forall i \geq p - 1,$
- iii)  $\Delta(p - 1, j) \neq 0, \forall j \geq q - 1,$

où :

$$\Delta(i, j) = \det \begin{pmatrix} \rho_X(j+1) & \rho_X(j+2) & \cdots & \rho_X(j+i+1) \\ \rho_X(j) & \rho_X(j+1) & \cdots & \rho_X(j+i) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_X(j+1-i) & \rho_X(j+2-i) & \cdots & \rho_X(j+1) \end{pmatrix}.$$

Les valeurs de  $\Delta(i, j)$  sont inconnues, mais elles peuvent être estimées par les  $\hat{\Delta}(i, j)$  obtenus en remplaçant les autocorrélations  $\rho_X(h)$  par leurs estimateurs  $\hat{\rho}_X(h)$  dans le déterminant. Puisque les  $\hat{\Delta}(i, j)$  sont des estimateurs cohérents des  $\Delta(i, j)$ , il est probable que nous remarquons une "rupture" entre les lignes  $i = p - 1$  et  $i = p$ , et entre les colonnes  $j = q - 1$  et  $j = q$  dans la matrice  $\hat{\Delta}(i, j)$ . La variance asymptotique de  $\hat{\Delta}(i, j)$  est une fonction différentiable du vecteur des autocorrélations  $\hat{\rho}_X(h)$ , avec une loi normale asymptotique. Un test de nullité ( $\hat{\Delta}(i, j) = 0$ ) est basé sur l'utilisation de la statistique de Student

$$\frac{|\hat{\Delta}(i, j)|}{\sqrt{\widehat{var}(\hat{\Delta}(i, j))}},$$

qui doit être comparée à 1.96 pour un seuil de signification de 5%.

## 2.2 Estimation des param etres

A cette  tape, les ordres  $p, d$  et  $q$  ont  t  fix es. Il reste donc d'estimer les param etres  $\phi_i, \theta_j$  et aussi  $\sigma_\varepsilon^2$ . Nous serons int eress es par les trois m ethodes suivantes :

### Estimation par la m ethode des moindres carr es

Soit  $X_t$  un mod ele  $AR(p)$ , s' crit sous la forme :

$$X_t = c + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t = Z^t \beta + \varepsilon_t,$$

o  :  $\varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2)$ ,  $Z^t = (1, X_{t-1}, \dots, X_{t-p})$  et  $\beta^t = (c, \phi_1, \dots, \phi_p)$ .

L'estimation des param etres de ce mod ele par la m ethode des moindres carr es est donn ee par :

$$\hat{\beta} = (ZZ^t)^{-1} Z X_t \quad \text{et} \quad \hat{\sigma}_\varepsilon^2 = \frac{1}{T - (p + 1)} \sum_{t=1}^T (X_t - Z^t \hat{\beta})^2.$$

#### Proposition 2.2.1

*Si les racines du polyn ome caract eristique  $\Phi(z)$  sont sup erieures strictement   1. Alors*

$$\hat{\beta} \xrightarrow{\mathcal{P}} \beta \quad \text{et} \quad \hat{\sigma}_\varepsilon^2 \xrightarrow{\mathcal{P}} \sigma_\varepsilon^2, \quad \text{quand } T \longrightarrow +\infty$$

*et de plus :*

$$\sqrt{T}(\hat{\beta} - \beta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma_\varepsilon^2 V), \quad \text{quand } T \longrightarrow +\infty \quad \text{avec } V = p \lim_{T \rightarrow +\infty} \frac{1}{T} Z Z^t.$$

### Estimation par la m ethode de maximum de vraisemblance

Soit  $X_t \sim ARMA(p, q)$ . Pour appliquer cette m ethode, il est n ecessaire de supposer que  $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$  ( $\varepsilon_t$  est par hypoth ese un Bruit Blanc).

$$\mathcal{L}(X, \phi, \theta, \sigma_\varepsilon^2) = \frac{1}{(2\pi\sigma_\varepsilon^2)^{T/2}} \frac{1}{\sqrt{\det \Omega}} \exp\left(-\frac{1}{2\sigma_\varepsilon^2} X^t \Omega^{-1} X\right).$$

où :  $\Omega$  est la matrice  $(T \times T)$  de variance-covariance du vecteur  $X = (X_1, \dots, X_T)^t$ , et  $\mathcal{L}(X, \phi, \theta, \sigma_\varepsilon^2)$  est la fonction de vraisemblance.

La maximisation et le calcul de cette quantité étaient relativement difficile à cause du calcul de l'inverse  $\Omega^{-1}$  et du déterminant de  $\Omega$ , surtout quand  $T$  est assez grand. Newbold a proposé une autre expression plus facile à calculer. On note par  $H$  la matrice triangulaire inférieure, à éléments positifs sur la diagonale, telle que  $HH^t = \Omega$  (décomposition de Cholesky), et soit  $e$  le vecteur où  $e = H^{-1}X$ . La log-vraisemblance du modèle s'écrit donc :

$$\begin{aligned} \log \mathcal{L} &= -\frac{T}{2} \log 2\pi - \frac{1}{2} \log \sigma_\varepsilon^2 - \frac{1}{2} |\det \Omega| - \frac{1}{2\sigma_\varepsilon^2} X^t \Omega^{-1} X \\ &= -\frac{T}{2} \log 2\pi - \frac{T}{2} \log(e^t e) - \log |\det H| \\ &= -\frac{T}{2} \log \left( |\det H|^{\frac{1}{T}} e^t e |\det H|^{\frac{1}{T}} \right). \end{aligned}$$

La méthode de maximum de vraisemblance revient finalement à chercher le minimum de :

$$l = \left( |\det H|^{\frac{1}{T}} e^t e |\det H|^{\frac{1}{T}} \right).$$

## Estimation par les équations de Yule-Walker

Dans le cas d'un modèle  $AR(p)$ , on peut utiliser le système de Yule-Walker, pour estimer les paramètres  $\phi = (\phi_1, \phi_2, \dots, \phi_p)$ , comme suit :

$$\hat{\phi} = \hat{\mathcal{R}}_p^{-1} \hat{\rho},$$

où :  $\hat{\mathcal{R}}_p$  est la matrice des autocorrélations estimées et  $\hat{\rho} = (\hat{\rho}_X(1), \hat{\rho}_X(2), \dots, \hat{\rho}_X(p))^t$  avec  $\hat{\rho}_X(h) = \frac{\hat{\gamma}_X(h)}{\hat{\gamma}_X(0)}$ . De plus, l'estimateur de la variance  $\sigma_\varepsilon^2$  est donné par :

$$\hat{\sigma}_\varepsilon^2 = \hat{\gamma}_X(0) \left( 1 - \sum_{j=1}^p \hat{\phi}_j \hat{\rho}_X(j) \right).$$

## 2.3 Validation

Les tests utilis es pour v erifier la validation des mod eles estim es sont de deux types : les tests concernant les param etres  $\phi_j$  et  $\theta_j$  du mod ele et ceux concernant l'hypoth ese sur le Bruit Blanc.

### 2.3.1 Tests de significativit e des param etres

Pour comparer une repr esentation  $ARMA(p, q)$  avec  $ARMA(p_1, q_1)$ . Tout d'abord, on suppose qu'on a le mod ele  $ARMA(p, q)$  et on examine les tests correspondants pour diff erentes valeurs de  $p_1$  et  $q_1$ .

$$\begin{cases} H_0 : \text{ Le mod ele est } ARMA(p, q), \\ H_1 : \text{ Le mod ele est } ARMA(p_1, q_1). \end{cases}$$

**i)**  $p_1 = p - 1, q_1 = q$  (ou  $p_1 = p, q_1 = q - 1$ ), pour tester la signification du coefficient  $\phi_p$  ( $\phi_p = 0$ ), on utilise le test de Student. On accepte  $H_1$  au seuil 5% si

$$\frac{|\hat{\phi}_p|}{\sqrt{\widehat{var}(\hat{\phi}_p)}} > 1.96.$$

**ii)**  $p_1 = p + 1, q_1 = q$ , cela peut ˆetre fait en estimant le mod ele  $ARMA(p_1, q_1) = ARMA(p + 1, q)$  et de tester l'hypoth ese  $\phi_{p+1} = 0$ . C'est la mˆeme que la proc edure d ecrite pr ec edemment.

**iii)**  $p_1 = p + 1, q_1 = q + 1$ , un mod ele  $X_t$  avec une repr esentation  $ARMA(p, q)$  admet aussi une repr esentation  $ARMA(p + 1, q + 1)$  en multipliant  $\Phi(L)$  et  $\Theta(L)$  par le mˆeme polynˆome  $(1 + \lambda L)$ . Donc, avec le mod ele  $ARMA(p + 1, q + 1)$ , l'hypoth ese nulle d'un  $ARMA(p, q)$  n'est pas identifiable. Les tests usuels tels que les tests de rapport de vraisemblance ne sont pas utilisables (en g en eral, nous choisissons le mod ele qui a la valeur  elev ee du rapport de vraisemblance).

### 2.3.2 Test sur le Bruit Blanc

L'hypoth ese que  $X_t$  est un  $ARIMA(p, d, q)$  peut s' ecrire de mani ere  equivalente  a l' equation (1.8) :

$$\varepsilon_t = \Theta(L)^{-1}(1 - L)^d \Phi(L) X_t.$$

L'estimation des ordres  $p, d, q$  et des param etres  $\phi_i$  et  $\theta_j$ , permettant d'obtenir les r esidus

estimés :

$$\hat{\varepsilon}_t = \hat{\Theta}(L)^{-1}(1 - L)^d \hat{\Phi}(L)X_t.$$

Alors, tester l'hypothèse ( $H_0$  : Le modèle est  $ARIMA(p, d, q)$ ), est équivalent à tester ( $H_0$  :  $\hat{\varepsilon}_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2)$ ). Le but est donc de vérifier si les résidus estimés  $\hat{\varepsilon}_t$ , est bien un Bruit Blanc.

**i) Test de Blancheur des résidus (Test de Portmanteau) :** Ce test a été proposé par Box et Pierce en 1970, pour tester l'hypothèse :

$$\begin{cases} H_0 : \rho_{\hat{\varepsilon}}(1) = \rho_{\hat{\varepsilon}}(2) = \dots = \rho_{\hat{\varepsilon}}(h) = 0, \\ H_1 : \text{il existe } i \text{ tel que } \rho_{\hat{\varepsilon}}(i) \neq 0. \end{cases}$$

Il est basé sur la statistique

$$Q = T \sum_{h=1}^K \hat{\rho}_{\hat{\varepsilon}}^2(h).$$

où :  $h$  est le nombre de retards,  $T$  est le nombre d'observations, et  $\hat{\rho}_{\hat{\varepsilon}}(k)$  est l'autocorrélation empirique de  $\hat{\varepsilon}$ . Sous l'hypothèse nulle, on peut montrer que  $Q$  est asymptotiquement distribué selon une loi de Khi-deux à  $K - p - q$  degrés de liberté, notée  $\chi_{K-p-q}^2$ . Donc, on rejette  $H_0$  au niveau de signification  $\alpha$  si :

$$Q > \chi_{1-\alpha}^2(K - p - q),$$

où :  $\chi_{1-\alpha}^2(K - p - q)$  est le quantile de la loi de Khi-deux à  $K - p - q$  degrés de liberté.

Les propriétés du petit échantillon de  $Q$  sont assez différentes des propriétés du grand échantillon, même pour un  $T$  assez grand. Pour cette raison, une statistique modifiée (améliorée) a été proposée par Ljung et Box en 1978, donnée par :

$$Q_{LB} = T(T + 2) \sum_{h=1}^K \frac{1}{T - h} \hat{\rho}_{\hat{\varepsilon}}^2(h),$$

le nombre  $K$  doit être choisi suffisamment grand, généralement entre 15 et 30 en pratique.

**ii) Test de normalité des résidus :** Leur but est de vérifier la normalité des résidus. Un test possible est celui de Bera et Jarque (1984). Notons par  $\mu_k = E([X - E(X)]^k)$  le moment d'ordre  $k$  de la distribution. On appelle : coefficient d'asymétrie (skewness)  $s = \frac{\mu_3}{\mu_2^{3/2}}$ , et

coefficient d'aplatissement (kurtosis)  $k = \frac{\mu_4}{\mu_2^2}$ . Sous l'hypothèse de normalité, on a :

$$s \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sqrt{6/T}\right) \text{ et } k \xrightarrow{\mathcal{D}} \mathcal{N}\left(3, \sqrt{24/T}\right) \text{ quand } T \rightarrow +\infty.$$

Alors le test de Bera et Jarque est défini par la statistique :

$$BJ = \frac{T}{6}s^2 + \frac{T}{24}[k - 3]^2,$$

qui suit une loi de  $\chi_2^2$ . Aussi, si  $BJ \geq \chi_{1-\alpha}^2(2)$  on rejette l'hypothèse  $H_0$  de normalité des résidus au seuil  $\alpha$ .

**Remarque 2.3.1** *En général, les logiciels fournissent la valeur de la statistique de test, ainsi que la  $p$ -valeur. On rejette l'hypothèse  $H_0$  au niveau  $\alpha$  si :  $p$ -valeur  $< \alpha$ .*

## 2.4 Choix du modèle

### 2.4.1 Critères de capacité prédictive

Dans le cas du modèle  $ARIMA(p, d, q)$ , l'erreur de prévision dépend de la variance  $\sigma_\varepsilon^2$ , Alors le but est de choisir le modèle qui réduit cette erreur, grâce à l'une des critères suivants :

- i) L'estimateur de la variance du résidu  $\hat{\sigma}_\varepsilon^2$ , ou la somme des carrés des résidus ( $SCR$ ).
- ii) Le coefficients de détermination  $R^2$ , qui correspondant à une normalisation de la variance.

$$R^2 = 1 - \frac{\hat{\sigma}_\varepsilon^2}{V_X}.$$

où :  $V_X$  est la variance empirique de la série différenciée  $d$  fois. Ce deuxième critère est simplement une version standardisée de la précédente.

- iii) Le coefficient de détermination modifié :

$$\bar{R}^2 = 1 - \frac{\hat{\sigma}_\varepsilon^2/(T - p - q)}{V_X/(T - 1)}.$$

où :  $T$  est la taille de la série temporelle.

iv) La statistique de Fisher (cas du modèle linéaire) :

$$F = \frac{(V_X - \hat{\sigma}_\varepsilon^2)/(p + q)}{\hat{\sigma}_\varepsilon^2/(T - p - q)}.$$

Alors, le critère *i*) doit être minimisé, tandis que *ii*), *iii*) et *iv*) doivent être maximisés.

## 2.4.2 Critères d'information

Ces critères sont établis en calculant l'écart entre le modèle proposé et la vraie loi, appelée distance de Kullback, dont le but est de minimiser le nombre de paramètre ( $p + q$ ).

i) Le critère *AIC* (Aikaïke Information Criterion 1969) :

$$AIC(p, q) = \ln \hat{\sigma}_\varepsilon^2 + \frac{2(p + q)}{T}.$$

ii) Le critère *BIC* (Schwartz Bayesian Information Criterion 1977) :

$$BIC(p, q) = \ln \hat{\sigma}_\varepsilon^2 + \frac{(p + q) \ln(T)}{T}.$$

iii) Le critère de Hannan-Quinn (1979) :

$$\phi(p, q) = \ln \hat{\sigma}_\varepsilon^2 + \frac{(p + q)c \ln(\ln(T))}{T}, \quad c > 2.$$

Pratiquement, un modèle choisi est celui associé aux plus petites valeurs de l'un des précédents critères.

## 2.5 Prévision

Il s'agit de la dernière étape de la méthode de Box et Jenkins. Soit  $\{X_t, t = 1, \dots, T\}$  une série temporelle stationnaire, on cherche à faire de la prévision à horizon  $h$ , et alors à prévoir  $X_{T+1}, X_{T+2}, \dots, X_{T+h}$ . Tous les processus *AR*, *MA* et *ARMA* seront supposés mis sous forme canonique, i.e. toutes les racines des polynômes autorégressifs  $\Phi$  et des polynômes moyennes mobiles  $\Theta$  auront leurs racines supérieures strictement à 1 et n'avoir aucune racine commune.

Ainsi, pour tous les mod eles  $X_t$  tels que :  $\Phi(L)X_t = \Theta(L)\varepsilon_t$ ,  $\varepsilon_t$  soit un Bruit Blanc centr ee et de variance finie.

### 2.5.1 Pr evision pour un mod ele autor egressif

Consid erons un mod ele  $AR(p)$

$$X_t = c + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2).$$

On appelle pr evision optimale   la date  $T + h$ , la valeur  $\hat{X}_{T+h}$  telle que

$$\begin{aligned} \hat{X}_{T+h} &= \hat{X}_T(h) = E(X_{T+h} | X_T, X_{T-1}, \dots, X_1) \\ &= E(X_{T+h} | \sigma(X)), \end{aligned}$$

o  :  $\sigma(X)$  est l'ensemble des informations disponible  $f(X_1, X_2, \dots, X_T)$ . De fa on r ecursive, et comme  $(\varepsilon_{T+h}, h \geq 1)$  est ind ependant de  $\sigma(X)$  alors  $E(\varepsilon_{T+h} | \sigma(X)) = 0$ , on trouve

$$\begin{aligned} \hat{X}_{T+1} &= c + \phi_1 X_T + \phi_2 X_{T-1} + \dots + \phi_p X_{T+1-p} \\ \hat{X}_{T+2} &= c + \phi_1 \hat{X}_{T+1} + \phi_2 X_T + \dots + \phi_p X_{T+2-p} \\ &\vdots \\ \hat{X}_{T+h} &= c + \phi_1 \hat{X}_{T+h-1} + \phi_2 \hat{X}_{T+h-2} + \dots + \phi_p \hat{X}_{T+h-p}, \quad h > p. \end{aligned}$$

On peut donc g en eraliser la formule de la pr evision pour un mod ele  $AR(p)$  par :

$$\hat{X}_{T+h} = \begin{cases} c + \phi_1 \hat{X}_{T+h-1} + \dots + \phi_{h-1} \hat{X}_{T+1} + \phi_h X_T + \dots + \phi_p X_{T+h-p} & \text{si } h \leq p, \\ c + \phi_1 \hat{X}_{T+h-1} + \phi_2 \hat{X}_{T+h-2} + \dots + \phi_p \hat{X}_{T+h-p} & \text{si } h > p. \end{cases}$$

### 2.5.2 Pr evision pour un mod ele moyenne mobile

On suppose que l'on a un mod ele  $MA(q)$

$$X_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}, \quad \varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2).$$

La prévision optimale à la date  $T + h$  est donnée par :

$$\hat{X}_T(h) = \hat{X}_{T+h} = E(X_{T+h} | X_T, X_{T-1}, \dots, X_1) = E(X_{T+h} | \varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_1).$$

On obtient donc

$$\begin{aligned} \hat{X}_{T+1} &= c + 0 + \theta_1 \varepsilon_T + \theta_2 \varepsilon_{T-1} + \dots + \theta_q \varepsilon_{T+1-q} \\ \hat{X}_{T+2} &= c + 0 + 0 + \theta_2 \varepsilon_T + \dots + \theta_q \varepsilon_{T+2-q} \\ &\vdots \\ \hat{X}_{T+h} &= c + \hat{\varepsilon}_{T+h} + \theta_1 \hat{\varepsilon}_{T+h-1} + \theta_2 \hat{\varepsilon}_{T+h-2} + \dots + \theta_q \hat{\varepsilon}_{T+h-q}, \end{aligned}$$

où :

$$\hat{\varepsilon}_{T+h-q} = E(\varepsilon_{T+h-q} | \sigma(X)) = \begin{cases} \varepsilon_{T+h-q} & \text{si } h \leq q, \\ 0 & \text{si } h \geq q + 1. \end{cases}$$

Finalement, la prévision optimale à la date  $T + h$ , pour un modèle  $MA(q)$  est :

$$\hat{X}_{T+h} = \begin{cases} c + \theta_h \varepsilon_T + \dots + \theta_q \varepsilon_{T+h-q} & \text{si } h \leq q, \\ c & \text{si } h > q. \end{cases}$$

### Remarque 2.5.1

*Cette méthode présente le désavantage d'estimer  $X_{T+h}$  à partir des résidus passés, car le passé du Bruit Blanc ne peut pas être connu.*

### Utilisation de l'écriture $AR(\infty)$ du modèle $MA(q)$

L'équation  $X_t = c + \Theta(L)\varepsilon_t$  peut être réécrite  $\Theta^{-1}(L)(X_t - c) = \varepsilon_t$ , ce qui est équivalent à

$$\Theta^{-1}(L)X_t = c + \varepsilon_t,$$

où :  $\Theta^{-1}(L) = \sum_{k=0}^{+\infty} \theta_k L^k$ ,  $\theta_0 = 1$ , de plus  $\sum_{k=0}^{+\infty} |\theta_k| < \infty$ .

Donc :

$$X_t = c - \sum_{k=1}^{+\infty} \theta_k X_{t-k} + \varepsilon_t, \text{ et } X_{T+h} = c - \sum_{k=1}^{+\infty} \theta_k X_{T+h-k} + \varepsilon_{T+h}, \quad h \geq 0.$$

On peut écrire la formule de la prévision optimale  $\hat{X}_{T+h}$  de façon itérative comme suit :

$$\begin{aligned} \hat{X}_{T+h} &= c - \sum_{k=1}^{+\infty} \theta_k \hat{X}_{T+h-k} = c - \sum_{k=1}^{h-1} \theta_k \hat{X}_{T+h-k} - \sum_{k=h}^{+\infty} \theta_k X_{T+h-k} \\ \hat{X}_{T+h} &= c - \sum_{k=1}^{+\infty} \theta_k \hat{X}_{T+h-k} = c - \sum_{k=1}^{h-1} \theta_k \hat{X}_{T+h-k} - \sum_{k=h}^{T+h-1} \theta_k X_{T+h-k} - \sum_{k=T+h}^{+\infty} \theta_k X_{T+h-k}. \end{aligned}$$

A partir du troisième terme, le problème qui se pose est que les  $X_t$  ne sont pas observées pour tout  $t \leq 0$ . (donc on le néglige). On approxime  $\hat{X}_{T+h}$  et on obtient

$$\hat{X}_{T+h} = c - \sum_{k=1}^{h-1} \theta_k \hat{X}_{T+h-k} - \sum_{k=h}^{T+h-1} \theta_k X_{T+h-k}.$$

### 2.5.3 Prévision pour un modèle ARMA

Soit  $X_t$  un modèle de type  $ARMA(p, q)$  qui s'écrit sous la forme  $\Phi(L)X_t = \Theta(L)\varepsilon_t$ . Alors :

$$X_t = \sum_{j=1}^p \phi_j X_{t-j} + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i}, \text{ et } X_{t+h} = \sum_{j=1}^p \phi_j X_{t+h-j} + \varepsilon_{t+h} + \sum_{i=1}^q \theta_i \varepsilon_{t+h-i}.$$

La prévision optimale d'un modèle  $ARMA(p, q)$  est donnée par :

$$\hat{X}_{T+h} = \sum_{j=1}^p \phi_j \hat{X}_{T+h-j} + \sum_{i=1}^q \theta_i \hat{\varepsilon}_{T+h-i}.$$

#### Utilisation de la forme $AR(\infty)$

La formule  $\Phi(L)X_t = \Theta(L)\varepsilon_t$  peut s'écrire :

$$\Theta^{-1}(L)\Phi(L)X_t = \varepsilon_t,$$

$$X_t = -\sum_{k=1}^{\infty} \phi_k X_{t-k} + \varepsilon_t, \text{ et } X_{t+h} = -\sum_{k=1}^{\infty} \phi_k X_{t+h-k} + \varepsilon_{t+h}, \quad h \geq 0.$$

D'où la prévision optimale est :

$$\hat{X}_{T+h} = - \sum_{k=1}^{h-1} \phi_k \hat{X}_{T+h-k} - \sum_{k=h}^{T+h-1} \phi_k X_{T+h-k}.$$

### Utilisation de la forme $MA(\infty)$

De la même manière, si on a  $\Phi(L)X_t = \Theta(L)\varepsilon_t$  on peut l'écrire sous forme :

$$X_t = \Phi^{-1}(L)\Theta(L)\varepsilon_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \dots + \theta_q\varepsilon_{t-q} + \dots = \varepsilon_t + \sum_{k=1}^{\infty} \theta_k\varepsilon_{t-k}.$$

Alors, la prévision optimale est :

$$\hat{X}_{T+h} = \sum_{k=h}^{\infty} \theta_k \hat{\varepsilon}_{T+h-k}.$$

## 2.5.4 Prévision pour un modèle ARIMA

Soit  $X_t$  un modèle de type  $ARIMA(p, d, q)$  s'écrit sous la forme  $\Phi(L)(1-L)^d X_t = \Theta(L)\varepsilon_t$  avec les conditions initiales  $Z = (X_{-1}, \dots, X_{-p-q}, \varepsilon_{-t}, \dots, \varepsilon_{-q})^t$ . On note par  $\Psi(L) = \Phi(L)(1-L)^d$ , donc :

$$X_t = \sum_{i=1}^{p+d} \psi_i X_{t-i} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j}, \text{ et de plus } X_{t+h} = \sum_{i=1}^{p+d} \psi_i X_{t+h-i} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t+h-j}.$$

Alors la prévision  $\hat{X}_{T+h}$  faite à la date  $T$  est :

$$\hat{X}_{T+h} = E(X_{T+h} | X_T, X_{T-1}, \dots, X_0, Z) = \sum_{i=1}^{p+d} \psi_i \hat{X}_{T+h-i} + \sum_{j=1}^q \theta_j \hat{\varepsilon}_{T+h-j},$$

$$\text{où : } \hat{X}_{T+h-i} = X_{T+h-i}, \text{ si } i \geq h \text{ et } \hat{\varepsilon}_{T+h-j} = \begin{cases} \varepsilon_{T+h-j}, & \text{si } j \geq h, \\ 0, & \text{si } j < h. \end{cases}$$

**Remarque 2.5.2** Si  $h \geq q$  on obtient une relation de récurrence de la forme :

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

# Chapitre 3

## Application sous R

Dans ce chapitre, nous menons une étude pratique prévisionnelle à l'aide des données réelles extraites à partir de la base des données du logiciel R (version 4.1.0), selon la méthode de Box et Jenkins.

On considère la série "milk", prise à partir du package (**TSA**), qui représente la production laitière mensuelle aux États-Unis (en millions de livres), de Janvier 1994 à Décembre 2005. Nous invitons le lecteur à consulter ces bonnes références [19] et [20] pour plus d'informations.

### 3.1 Etape 1 : Identification

D'après le schéma additif de la série (figure 1.2) et le graphe de décomposition (figure 1.1), on constate que la série est non stationnaire (a une tendance linéaire et une composante saisonnière annuelle  $s = 12$ ). Pour confirmer ce résultat on applique le test d'ADF en utilisant la commande `adf.test` (du package `tseries`), ce qui nous donne un  $p$  – valeur = 0.363 > 0.05, on accepte alors l'hypothèse nulle de non stationnarité. Donc, la série sera modélisée à travers un modèle de type  $SARIMA(p, d, q)(P, D, Q)_s$ .

L'étape suivante consiste à identifier la valeur préliminaire des ordres de différentiations  $d$  et  $D$ . La représentation graphique de l'ACF (figure 2.1), obtenue à l'aide de la commande `acf`, montre une lente décroissance de  $\hat{\rho}(h)$ , cela indique qu'une différentiation peut être nécessaire pour éliminer la tendance ( $d = 1$ ) et une autre différentiation pour supprimer la saisonnalité ( $D = 1$ ). Il convient donc de transformer la série initiale, en utilisant une double différentiation

pour obtenir une série stationnaire  $\Delta(\Delta_{12})X_t = (1 - L)(1 - L^{12})X_t$  (figure 3.1).

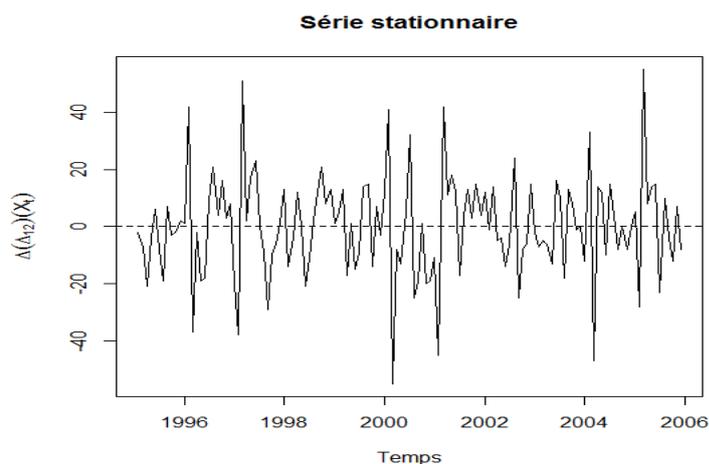


FIG. 3.1 – Représentation graphique de la série stationnaire  $\Delta(\Delta_{12})X_t$ .

Dans la figure 3.2, nous illustrons l'ACF et le PACF de la série stationnaire  $\Delta(\Delta_{12})X_t$ . L'analyse de ces deux graphiques nous permet de déterminer les termes d'ordre non saisonniers  $(p, q)$  et les termes d'ordre saisonniers  $(P, Q)$  du modèle.

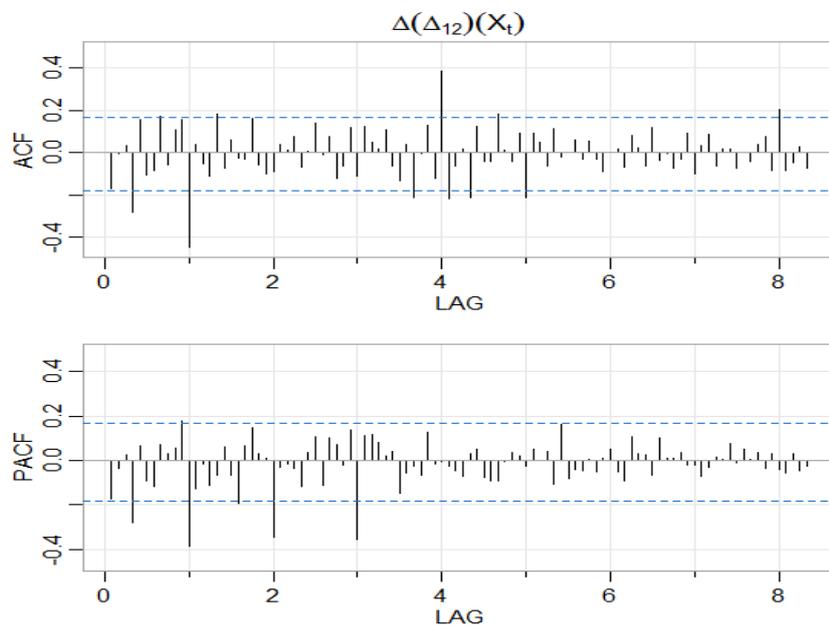


FIG. 3.2 – ACF et PACF de la série stationnaire  $\Delta(\Delta_{12})X_t$

**Les ordres saisonniers :**

- Il semble qu’au fil des saisons (à l’échelle annuelle), l’ACF a un pic prononcé au décalage saisonnier  $1s = 12$  (i.e. une forte autocorrélation au 1<sup>er</sup> retard) et à  $4s = 48$  (mais aucun à  $2s$  et  $3s$ ).
- D’autre part, le PACF affiche des pics prononcé aux décalages saisonniers  $1s = 12, 2s = 24$  et  $3s = 36$ .

Les deux dernières observations sont cohérentes avec les choix suivants :  $(P, Q) = (3, 0), (P, Q) = (3, 1)$  ou  $(P, Q) = (2, 1)$ .

**Les ordres non saisonniers :**

- Concentrant juste sur une saison, il semble que l’ACF et le PACF n’ont pas vraiment de décrochage sur une année. On prend donc,  $p = 0$  et  $q = 0$ .

Par conséquent, il existe trois modèles qui émergent comme des possibilités fortes :

- Modèle1 :  $SARIMA(0, 1, 0)(3, 1, 0)_{12}$ ,
- Modèle2 :  $SARIMA(0, 1, 0)(3, 1, 1)_{12}$ ,
- Modèle3 :  $SARIMA(0, 1, 0)(2, 1, 1)_{12}$ .

### 3.2 Etape 2 : Estimation des paramètres

La procédure `Arima` (du package `forecast`) permet d’estimer les coefficients et la variance résiduelle  $\sigma_\varepsilon^2$  des trois modèles précédents à travers la méthode du maximum de vraisemblance.

Les résultats sont résumés dans le tableau [3.1](#) :

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\gamma}_1$	$\hat{\sigma}_\varepsilon^2$
Modèle1	-0.9133	-0.8146	-0.6002		124.4
Modèle2	-0.9839	-0.8615	-0.6317	0.1166	124.7
Modèle3	-0.3054	-0.3117		-0.6475	155.6

TAB. 3.1 – Résumé des résultats obtenus à partir de l’estimation des paramètres

### 3.3 Etape 3 : Validation

i) **Test de significativité des coefficients** : Les résultats de la  $p$ -valeur du test de Student renvoyés, en utilisant le package `caschrono`, sont enregistrés dans le tableau [3.2](#).

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\gamma}_1$
Modèle 1	0.0000	0.00000	0.000000	
Modèle 2	0.000000	0.000000	0.000000	0.511781
Modèle 3	0.00700	0.002931		0.000000

TAB. 3.2 – Résumé des résultats obtenus à partir du test de significativité des coefficients

On remarque que les coefficients estimés sont significatifs au seuil 5% ( $p$ -valeur  $< 0.05$ ), à l'exception du coefficient  $\hat{\gamma}_1$  pour le modèle 2. Il reste donc, à étudier le modèle 1 et le modèle 3.

ii) **Test sur le Bruit blanc** : La préoccupation principale du modèle est de s'assurer que les résidus sont normalement distribués avec une moyenne nulle et non corrélés. La sortie de la commande `tsdiag` est la figure [3.3](#).

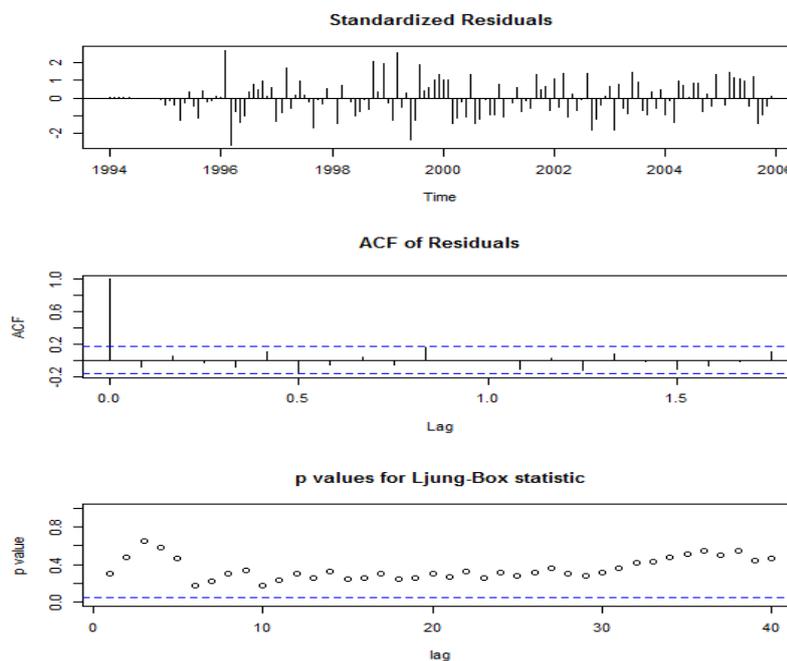


FIG. 3.3 – Analyse des résidus standardisés pour le modèle  $SARIMA(0, 1, 0)(3, 1, 0)_{12}$ .

- Le graphique (en haut) montre les résidus standardisés au fil du temps, et ils semblent être stationnaires sans composante saisonnière.
- L'autocorrélogramme (au milieu) indique que les résidus ont une forte corrélation pour  $h = 0$ , et une faible corrélation avec les versions décalées pour  $h > 0$ .

- Le graphique de la p-valeurs (en bas,) du test de Ljung et Box montre que l'hypothèse de blancheur des résidus est acceptée ( $p - valeur > 0.05$ ).
- Enfin, la sortie de la commande `qqnorm` donne le graphique QQ-plot de la loi normal (figure 3.4), il montre que la distribution des résidus suit de près la tendance linéaire des échantillons prélevés à partir d'une distribution normale standard, ce qui suggère que les résidus sont normalement distribués. Cela est confirmé par le test de Bera et Jarque à l'aide de la commande `jarque.bera.test` ( $p - valeur = 0.6454 > 0.05$ ).

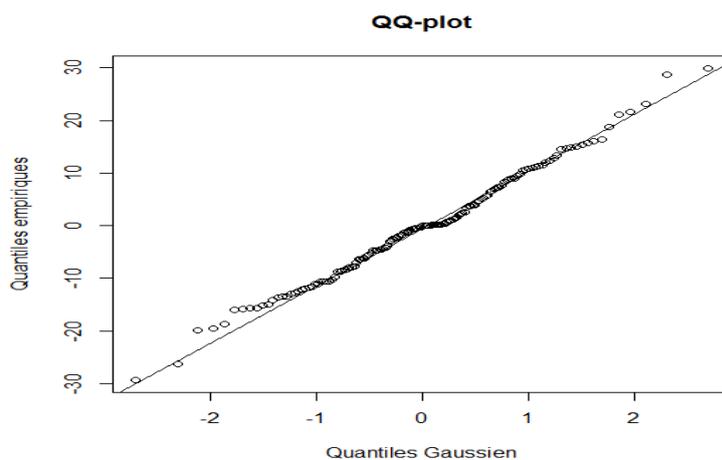


FIG. 3.4 – QQ-plot des résidus du modèle :  $SARIMA(0, 1, 0)(3, 1, 0)_{12}$  et la loi normale.

### 3.4 Etape 4 : Choix du modèle

Revenant à l'analyse des valeurs de l'AIC et le BIC résumées dans le tableau 3.3, et qui sont fournies à l'aide de la commande `Arima` (du package `forecast`).

	AIC	BIC
Modèle 1	1032.05	1043.55
Modèle 3	1054.24	1065.74

TAB. 3.3 – Résumé des résultats obtenus pour les valeurs des critères AIC et BIC

Les deux critères d'information AIC et BIC préfèrent le modèle1,  $SARIMA(0, 1, 0)(3, 1, 0)_{12}$  :

$$(1 - \beta_1 L^s - \beta_2 L^{2s} - \beta_3 L^{3s})(1 - L)(1 - L^s) X_t = c + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2).$$

De plus, la trajectoire de la série  $\Delta(\Delta_{12})X_t$  (figure 3.1) oscille davantage autour de 0, ce qui

signifie que la constante  $c$  est nulle. Par conséquent, le meilleur modèle ajusté est :

$$(1 + 0.9133L^{12} + 0.8146L^{24} + 0.6002L^{36})(1 - L)(1 - L^{12})X_t = \varepsilon_t, \varepsilon_t \sim \mathcal{BB}(0, \sigma_\varepsilon^2).$$

### 3.5 Etape 5 : Prévision

Finalement, selon le modèle choisi, on peut faire des prévisions futures de la série "milk" sur une période à court terme, à l'aide de la commande `forecast`, avec un intervalle de prévision à 95%. Les résultats de prévision sur une période de quatre ans (2006 – 2009) sont illustrées dans la figure 3.5. À partir de cette représentation, nous voyons que les valeurs de prévision sont assez proches aux valeurs réelles de la série lors des trois dernières années : 2003, 2004 et 2005. On constate que le modèle  $SARIMA(0, 1, 0)(3, 1, 0)_{12}$  décrit bien les données de production laitière aux États-Unis.

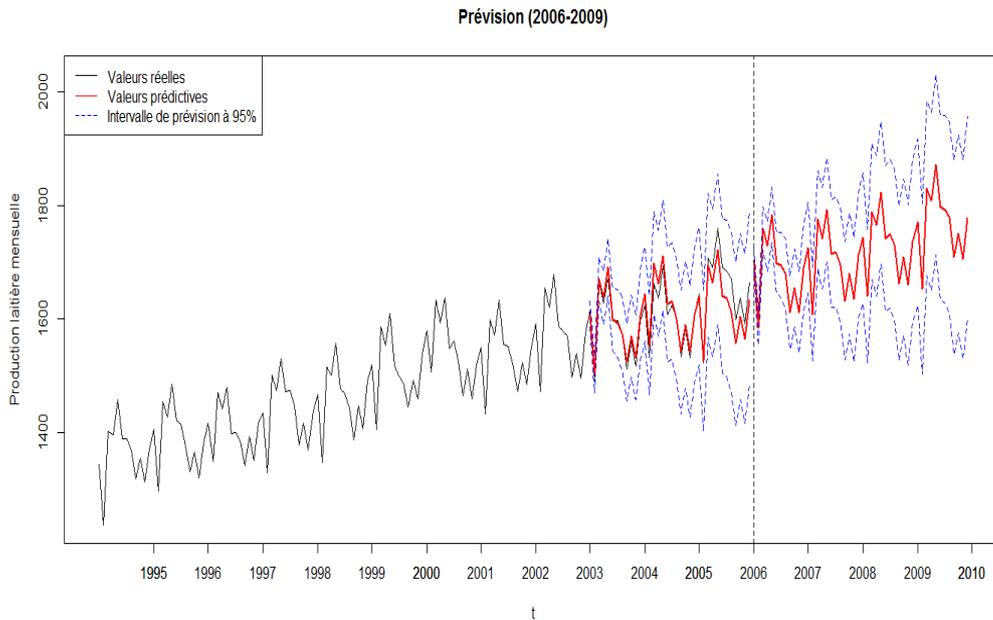


FIG. 3.5 – Représentation graphique des valeurs prédictives de la série "milk" sur une période de 4 ans à l'aide du modèle  $SARIMA(0, 1, 0)(3, 1, 0)_{12}$ .

## 3.6 Mise en œuvre sous R

Dans ce chapitre, les résultats ont été obtenus par l'implémentation du code R ci-dessous.

---

```
library(TSA)

data(milk)

X=milk

plot.ts(X,main="Modèle additif",xlab="Temps",ylab=expression(X[t]))

decomp=decompose(X,type="additive")

plot(cbind(Observations=decomp$x,Tendance=decomp$trend,Saisonnalité=decomp$seasonal,
          Résidu=decomp$random),main="Modèle d'une série temporelle",xlab="Temps")

acf(milk,lag.max=50,main="Autocorrélogramme",xlab="h")

library(tseries)

adf.test(X,alternative="stationary",k=12)

Xdd=diff(diff(X,lag=12),difference=1)

plot.ts(Xdd,main="Série stationnaire",ylab=expression(Delta(Delta[12])(X[t])))

abline(h=0,lty=2)

library(astsa)

acf2(Xdd,max.lag=100,main=expression(Delta(Delta[12])(X[t])))

library(forecast)

modèle1=Arima(X,order=c(0,1,0),list(order=c(3,1,0),period=12),include.mean=FALSE)
modèle2=Arima(X,order=c(0,1,0),list(order=c(3,1,1),period=12),include.mean=FALSE)
modèle3=Arima(X,order=c(0,1,0),list(order=c(2,1,1),period=12),include.mean=FALSE)

library(caschrono)

t_stat(modèle1)

t_stat(modèle2)

t_stat(modèle3)

tsdiag(modèle1,gof.lag=40)

qqnorm(residuals(modèle1),xlab="Quantiles Gaussien",ylab="Quantiles empiriques",
       main="QQ-plot")

qqline(residuals(modèle1))

jarque.bera.test(residuals(modèle1))
```

```
pred_model=forecast(modèle1,h=48,level=95)
pred=pred_model$mean
pred_l=ts(pred_model$lower,start=c(2006,1),frequency=12)
pred_u=ts(pred_model$upper,start=c(2006,1),frequency=12)
Xtronc=window(X,end=c(2002,12))
modeltronc=Arima(Xtronc,order=c(0,1,0),list(order=c(3,1,0),period=12),include.mean=FALSE)
predmod1tronc=forecast(modeltronc,h=36,level=95)
predtronc=predmod1tronc$mean
predtroncl=ts(predmod1tronc$lower,start=c(2003,1),frequency=12)
predtroncu=ts(predmod1tronc$upper,start=c(2003,1),frequency=12)
ts.plot(X,pred,pred_l,pred_u,predtronc,predtroncl,predtroncu,xlab="t",
        ylab="Production laitière mensuelle",col=c(1,2,4,4,2,4,4),lty=c(1,1,2,2,1,2,2),
        lwd=c(1,2,1,1,2,1,1),main="Prévision (2006-2009)")
legend("topleft",c("Valeurs réelles","Valeurs prédictives","Intervalle de prévision à 95%"),
      col=c(1,2,4),lty=c(1,1,2))
abline(v=2006,lty=2)
axis(1,xaxp=c(1995,2010,15))
```

---

# Conclusion

L'analyse d'une série temporelle ne constitue que rarement une fin en soi. Le véritable objectif poursuivi est la prévision à fin de prendre des décisions sur la stratégie qui sera suivie à l'avenir. Ce mémoire de master, vise à donner un aperçu sur la méthode de Box et Jenkins, qui constitue l'une des méthodes les plus utilisées dans le domaine de l'analyse et de la prévision des séries temporelles.

Le principe de cette méthode est de prédire les valeurs futures des séries temporelles à partir de leurs valeurs passées, en utilisant une approche itérative en cinq étapes, pour arriver à un modèle approprié de type : AR, MA, ARMA, ARIMA ou SARIMA, qui permet d'obtenir des résultats et des prédictions précis.

Une étude pratique a été réalisée, en appliquant la méthode de Box et Jenkins pour modéliser une série temporelle réelle qui représente la production laitière mensuelle aux États-Unis.

En conclusion, le meilleur modèle retenu est  $SARIMA(0, 1, 0)(3, 1, 0)_{12}$  et c'est à partir de ce dernier que les prévisions ont été faite pour une période de quatre ans.

# Bibliographie

- [1] Anderson, T. W. (1971). *The Statistical Analysis of Time Series*. Wiley, New York.
- [2] Adhikari, R., & Agrawal, R. K. (2013). *An Introductory Study on Time Series Modeling and Forecasting*.
- [3] Box, G. E. P., & Jenkins, G. M. (1970). *Time Series Analysis : Forecasting and Control*, Holden-Day, San Francisco. First Edition.
- [4] Box, G. E. P., & Jenkins, G. M. (1976). *Time Series Analysis Forecasting and Control*. Revised Edition.
- [5] Brockwell, P. J., & Davis, R. A. (1987). *Time Series : Theory and Methods*. Springer New York.
- [6] Brockwell, P. J., & Davis, R. A. (1991). *Time Series : Theory and Methods*, 2nd edition. Springer New York.
- [7] Brockwell, P. J., & Davis, R. A. (1996). *Introduction to Time Series and Forecasting*, 2nd edition. Springer New York.
- [8] Charpentier, A. (1916). *Cours de séries temporelles théorie et applications, Volume 1 : Introduction à la théorie des processus en temps discret Modèles ARIMA et méthode de Box & Jenkins*.
- [9] Charpentier, A. (2012). *Modèles de prévision Séries temporelles*. Paris, France.
- [10] Chatfield, C. (2004). *The Analysis of Time Series An introduction*. Sixth edition.
- [11] Chatfield, C., & Hall, C. R. C. (2000). *Time-Series Forecasting*.
- [12] Cryer, D. J., & Chan, K. S. (2008). *Time Series Analysis With Application in R*. Second Edition. Springer.
- [13] Dauxois, J. Y. (2017). *Introduction à l'Étude des Séries Temporelles*.

- [14] Ferrara, L. *Séries Temporelles Avancées Polycopié de Cours*.
- [15] Gaude, Y. (2016). *Introduction aux séries temporelles, tendance et composante saisonnière*.
- [16] Desbois, D. (2014). *Une introduction à la méthodologie de Box et Jenkins : l'utilisation de modèles ARIMA avec SPSS*.
- [17] Gouriéroux, C., & Monfort, A. (1997). *Time Series and Dynamic Models*.
- [18] Shumway, R. H., & Stoffer, D. S. (2006). *Time Series Analysis and Its Applications With R Exapmles*, 2nd edition. Springer New York.
- [19] Shumway, R. H., & Stoffer, D. S. (2017). *Time Series Analysis and Its Applications With R Examples*. Fourth Edition. Springer.
- [20] Tebbs, J. M. (2013). *Forecasting and Time Series : Lecture Note*. University of South Carolina.
- [21] Von Sachs, R., & Von Bellegem, S. (2005). *Séries Chronologiques*. 4e édition. Université de Louvain.
- [22] Wei, W. W. S. (2006). *Time Series Analysis Univariate and Multivariate Methods*. Second Edition.
- [23] Yahia, D. (2005). *Séries temporelles et test d'adéquation pour un modèle GARCH(1,1)*. Mémoire de Magister. Université de Biskra.
- [24] Yahia, D., & Benameur, S. (2020). *Notes de Cours : Séries Chronologiques*. Université de Biskra.

# Annexe A : Abréviations et Notations

<b>Symbole</b>	<b>Signification</b>
$ACF$	: Autocorrélogramme.
$PACF$	: Autocorrélogramme partielle.
$AR$	: Autorégressif.
$ARMA$	: Autorégressif moyenne mobile.
$ARIMA$	: Autorégressif moyenne mobile intégré.
$SARIMA$	: Autorégressif moyenne mobile intégré saisonnier.
$\mathcal{BB}$	: Bruit Blanc.
$\xrightarrow{\mathcal{D}}$	: Converge en distribution.
$\xrightarrow{\mathcal{P}}$	: Converge en probabilité.
$corr(.)$	: Corrélation.
$E(. .)$	: Espérance conditionnelle.
$E(.)$	: Espérance mathématique (Moyenne).
$M(L)$	: Filtre.
$\rho_X(.)$	: Fonction d'autocorrélation de $X$ .
$\phi_X(.)$	: Fonction d'autocorrélation partielle de $X$ .
$\gamma_X(.)$	: Fonction d'autocovariance de $X$ .
$H_1$	: Hypothèse alternative.
$H_0$	: Hypothèse nulle.
$iid$	: Indépendante et identiquement distribuée.

$i.e$	: in other words.
$\chi_v^2$	: Loi de Khi-deux avec $v$ degré de liberté.
$\mathcal{N}(\mu, \sigma^2)$	: Loi normale, d'espérance $\mu$ et de variance $\sigma^2$ .
$\mathcal{R}_p$	: Matrice des corrélations.
$MA$	: Moyenne mobile (Moving average).
$F$	: Opérateur d'avance.
$\Delta$	: Opérateur de différentiation.
$L$	: Opérateur de retard.
$\varepsilon_t$	: Résidu (Erreur).
$S_t$	: Saisonnalité.
$X_t$	: Série temporelle.
$BJ$	: Statistique de Bera et Jarque.
$Q_{LB}$	: Statistique de Ljung et Box.
$Z_t$	: Tendance.
$ADF$	: Test de Dickey-Fuller augmenté.
$DF$	: Test de Dickey-Fuller simple.
$ \cdot $	: Valeur absolue.
$var(\cdot)$	: Variance.

# Annexe B : Packages et Commandes

Pour les simulations, nous avons utilisés le logiciel R, et plus précisément les commandes et les packages que nous allons montrer ci-dessous :

## Les packages :

- ✓ `TSA` : pour obtenir les données de la série (milk).
- ✓ `tseries` : pour effectuer le test de Dickey-Fuller augmenté.
- ✓ `astsa` : pour analyser des séries temporelles statistiques appliquées.
- ✓ `forecast` : pour estimer les paramètres d'un modèle, selon la méthode du maximum de vraisemblance.
- ✓ `caschrono` : pour appliquer le test de Student.
- ✓ `graphics` : pour utiliser certaines des capacités graphiques de R.

## Les commandes :

- ✓ `plot.ts (plot)` : pour représenter la série temporelle graphiquement.
- ✓ `diff(diff(serie,lag=s,difference=d))` : pour appliquer l'opérateur de différentiation  $\Delta^d \Delta_s$ .
- ✓ `decompose(serie,type=c("additive","multiplicative"))` : pour extraire les composantes d'une série temporelle.
- ✓ `adf.test` : pour effectuer le test de Dickey-Fuller augmenté.
- ✓ `acf` : pour tracer l'ACF de la série temporelle.
- ✓ `acf2` : pour tracer simultanément l'ACF et le PACF sur le même échelle.
- ✓ `jarque.bera.test` : pour appliquer le test de Bera et Jarque de normalité.
- ✓ `Arima(serie,ordre=c(p,d,q),list=(ordre=c(P,D,Q),periode=s))` : pour estimer les paramètres d'un modèle  $SARIMA(p, d, q)(P, D, Q)_s$ .

- ✓ `tsdiag` : pour tracer des diagnostics des résidus de séries temporelles.
- ✓ `qqnorm` : pour tracer le graphique QQ plot pour une loi normale.
- ✓ `qqline` : pour ajouter une ligne qui passe par les premier et troisième quartiles sur le graphique QQ plot (la droite de Henry).
- ✓ `abline` : pour ajouter une ou plusieurs lignes droites à travers le tracé courant.
- ✓ `window` : pour extraire le sous ensemble de l'objet X observé entre les instants de début et de fin.
- ✓ `axis` : pour ajouter un axe au tracé actuel, permettant de spécifier le côté, la position, les étiquettes et d'autres options.
- ✓ `ts.plot` : pour tracer plusieurs séries temporelles sur un même tracé.
- ✓ `legend` : pour ajouter des légendes aux graphiques.

## Résumé

Ce mémoire est essentiellement consacré à la méthode de Box et Jenkins. Cette méthode en cinq étapes vise à formuler un modèle qui permet de représenter des séries temporelles afin de faire des prévisions des valeurs futures.

Tout d'abord, nous introduisons quelques concepts de base liés aux séries temporelles. S'ensuit la présentation des modèles linéaires, à savoir le modèle AR, MA, ARMA, ARIMA et SARIMA. Puis nous présentons la méthode de Box et Jenkins. Nous terminons ce travail par une application de cette méthode afin de déterminer le modèle le plus approprié pour prédire une série de données réelles, en utilisant le logiciel R.

**Mots clés :** Modèle linéaire - Méthode de Box et Jenkins - Prévision - Série temporelle - Stationnarité.

## ملخص

خصصت هاته المذكرة بشكل أساسي لطريقة بوكس وجينكينز . تهدف هذه الطريقة المكونة من خمس مراحل إلى صياغة نموذج يسمح بتمثيل سلاسل زمنية من أجل إجراء تنبؤات بالقيم المستقبلية.

أولاً، نقدم بعض المفاهيم الأساسية المتعلقة بالسلسلة الزمنية، يلي ذلك عرض لنماذج خطية، وهي AR ، MA ، ARMA ، ARIMA و SARIMA . ثم نعرض طريقة بوكس وجينكينز . ننهي هذا العمل من خلال تطبيق هذه الطريقة لتحديد النموذج الأنسب للتنبؤ بسلسلة من البيانات الحقيقية، باستخدام برنامج R.

**الكلمات المفتاحية :** النموذج الخطي- طريقة بوكس و جينكينز - التنبؤ - السلسلة الزمنية - الثبات .

## Abstract

This memory is essentially devoted to the Box and Jenkins method. This five-step method aims to formulate a model that can represent time series in order to make predictions of future values.

First, we introduce some basic concepts related to time series. This is followed by the presentation of linear models, namely the AR, MA, ARMA, ARIMA and SARIMA model. Then we present the Box and Jenkins method. We finish this work by applying this method in order to select the most appropriate model to predict a series of real data, using the R software.

**Keywords :** Linear model - Box and Jenkins method - Forecasting - Time serie – Stationarity.