

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

*Université M'Hamed Bougara de BOUMERDES
faculté des Sciences
Département de Mathématique*



Mémoire de Master

RECHERCHE OPÉRATIONNELLE OPTIMISATION
ET MANAGMENT STRATÉGIQUE

UTILISATION DE LA METHODOLOGIE DE BOX-JENKINS POUR ETUDE PREVISIONNELLE

Présenté par : *M^{elle}* Belkhiri Assala et *M^{me}* Karaoui Kahina

Date : 07/07/2022.

Membres du Jury :

Présidente :	<i>M^{me}</i> Metref N.	M.A.A , Université U.M.B.B
Promotrice :	<i>M^{me}</i> Akli M.	M.A.A , Université U.M.B.B
Examinatrice :	<i>M^{me}</i> Bourennani K.	M.A.A , Université U.M.B.B
Examinatrice :	<i>M^{me}</i> tahar S.	M.A.A , Université U.M.B.B

TABLE DES MATIÈRES

Remerciements	5
Dédicace	7
Dédicace	8
Introduction	10
1 Processus stochastiques	12
1.1 Définition d'une série chronologique :	12
1.1.1 Stationnarité :	12
1.2 Caractéristique d'une série chronologique :	13
1.2.1 Moyenne et variance :	13
1.2.2 Fonction d'auto-covariance :	13
1.2.3 Fonction d'auto-corrélation Simple :	14
1.2.4 Fonction d'auto-corrélation partielle :	14
1.2.5 Processus bruit blanc :	16
1.3 Série chronologique :	16
1.4 Décomposition d'une série chronologique :	17
1.4.1 Le modèle additif :	17
1.4.2 Modèle multiplicatif :	17
1.4.3 Modèles mixtes (hybrides) :	18
1.4.4 Choix du modèle :	18
1.4.5 Estimation des variations saisonnières :	19
1.5 Les opérateurs linéaires :	20
1.5.1 Opérateur de retard B :	20
1.5.2 Opérateur d'avance F :	20
1.5.3 Opérateur de différence ordinaire :	20
1.5.4 Opérateur de différence saisonnière :	20
1.6 Typologie des modèles Auto Régressifs Moving Average (ARMA) et ses dérivées :	21
1.6.1 Modèle AR (Auto Régressif) :	21
1.6.2 Modèle Moving Average (MA) :	23
1.6.3 Modèle Auto Régressif Moyenne Mobile ARMA :	24
1.6.4 Modèle Auto Régressif Moyenne Mobile Saisonnier SARMA :	25
1.7 Les processus aléatoires non stationnaires :	25
1.7.1 processus TS :	25
1.7.2 processus DS :	26
1.8 Modèles non stationnaires ARIMA et SARIMA :	26
1.8.1 Modèle ARIMA :	26
1.8.2 Modèle SARIMA :	27

2	Méthodologie de Box-Jenkins	28
2.1	Introduction	28
2.2	Notions sur la méthodologie de Box-Jenkins	29
2.2.1	Familiarisation avec les données :	29
2.2.2	Analyse préliminaire :	29
2.2.3	Stationnarisation de la série :	29
2.3	Identification du modèle :	31
2.3.1	Critère d' Akaike (Akaike information critrion) :	32
2.3.2	Critère de SCHWARZ (BAYESIAN information critrion) :	32
2.3.3	Critère de Hannan-Quin (1979) :	32
2.4	Estimation du modèle :	32
2.5	Validation du modèle :	33
2.5.1	Test sur les paramètres :	33
2.5.2	Test sur les résidus :	33
2.6	Prévision :	35
2.6.1	Prévision d'un modèle AR(p)	36
2.6.2	Prévision d'un modèle MA(q) :	36
2.6.3	Prévision d'un modèle ARMA(p,q)	36
3	Application de la méthode de Box-Jenkins	37
3.1	Analyse et exploratoire des données	37
3.1.1	Analyse de la série ApproBrute	37
3.1.2	Analyse de la série Approvisionnement Pré Traitement	40
3.2	Identification du modèle	42
3.2.1	Identification de la série ApproBrute	42
3.2.2	Identification du la série ApproPreTrai	42
3.3	Estimation des paramètres	43
3.3.1	Estimation du la série ApproBrute	43
3.3.2	Estimation du la série ApproPreTrai	43
3.4	Validation du modèle :	44
3.4.1	Validation du la série ApproBrute	44
3.4.2	Validation du la série ApproPreTrai	46
3.5	Prévision :	49
3.5.1	Prévision du la série ApproBrute	49
3.5.2	Prévision du la série ApproPreTrai :	49
4	Lissage exponentiel	51
4.1	Introduction	51
4.2	Lissage exponentiel	52
4.2.1	Le lissage exponentiel simple	53
4.2.2	le lissage double	53
4.2.3	Le lissage exponentiel généralisé	53
4.2.4	Le lissage de Holt & Winaters	54
5	Application de lissage exponentiel	56
5.1	Etude de la série ApproBrute	56
5.1.1	Les prévisions	57
5.2	Etude de la série ApproPreTrai	57
5.2.1	Les prévisions	57
5.3	Conclusion	58
	Conclusion générale	59

Résumé	60
Bibliographie	61

TABLE DES FIGURES

2.1	Diagramme des étapes de la méthode de Box-Jenkins	29
2.2	Dickey-Fuller augmenté	31
3.1	Représentation graphique de la série brute	37
3.2	Corrélogramme de la série Appro Brute	38
3.3	Test sur la tendance ApproBrute	39
3.4	Test sur la constante ApproBrute	39
3.5	Représentation graphique de la série Approvisionnement Pré Traitement	40
3.6	Corrélogramme de la série Appro-Pre-Trai	40
3.7	Test sur la tendance	41
3.8	Test sur la constante	42
3.9	Critères d'information pour les deux modèles	42
3.10	Critères d'information pour les cinq modèles	43
3.11	Modèle ARMA(2,2)	43
3.12	Modèle ARMA(2,0)	44
3.13	Graphe des inverses des racines	44
3.14	Graphique des séries résiduelles, actuelles et estimées	45
3.15	Corrélogramme des résidus	45
3.16	Histogramme des résidus	46
3.17	Graphe des inverses des racines	47
3.18	Graphique des séries résiduelles, actuelles et estimées	47
3.19	Corrélogramme des résidus	48
3.20	Histogramme des résidus	48
3.21	Tableau de Prévision	49
3.22	Tableau de Prévision	49
3.23	Graphe prévision	50
5.1	Lissage ApproBrute	56
5.2	Prévision lissage ApproBrute	57
5.3	Lissage ApproPreTrai	57
5.4	Prévision lissage ApproPreTrai	58



Remerciements

Premièrement, nous remercions DIEU de nous avoir donné la chance de terminer nos études et d'avoir notre diplôme.

Ce travail a été effectué sous la direction scientifique de madame Meriem Akli, on la remercie profondément d'avoir accepté d'être la promotrice de notre mémoire et aussi pour sa disponibilité, ses nombreux conseils et de nous avoir orienté le long de la réalisation de ce travail.

Un grand merci aux membres de jury Mme Nadia Metref, Mme Souhila Tahar et Mme Kenza Bourennani, pour l'honneur qu'elles nous ont fait de bien vouloir accepter d'évaluer notre travail.

Enfin, merci à tout ceux qui ont contribué des près ou de loin à l'achèvement de ce travail mais qui n'ont pas été cités. Nous ne saurions jamais vous remercier assez pour l'aide que vous nous avez apporté..

Merci !

Belkhiri & Karaoui 



Dédicace

Je dédie ce travail :

A ma chère mère et à mon cher père qui n'ont jamais cessé de me soutenir et m'encourager durant toutes mes années d'études. Qu'ils trouvent ici le témoignage de ma profonde gratitude et
Reconnaissance

A mes frères, Adem, Oussama et Islam, et ma chère soeur Amina, et ma grand-mère Dalila et ma famille qui me donnent de l'amour et de la vivacité.

A tous ceux qui m'ont aidé de près ou de loin et ceux qui ont partagé avec moi les moments d'émotion lors de la réalisation de ce travail et qui m'ont chaleureusement supporté et encouragé tout au long de mon parcours.

A tous mes amies, Sara, Fouzia et ma binome Kahina qui m'ont toujours encouragé, et à qui je souhaite plus de succès.

Merci !

Belkhiri & Assala





Dédicace

Je dédie ce travail :

*A ma chère mère et à mon cher père qui n'ont jamais cessé de me soutenir et m'encourager durant mes années d'études. Qu'ils trouvent ici le témoignage de ma profonde gratitude et
Reconnaissance*

A mes frères, Billal, Amin, et mes chères soeurs Warda, et surtout Louiza qui m'a aidé à déterminer mes études, à mon mari et ma famille qui me donnent de l'amour et de la vivacité.

A tous ceux qui m'ont aidé - de près ou de loin et ceux qui ont partagé avec moi les moments d'émotion lors de la réalisation de ce travail et qui m'ont chaleureusement supporté et encouragé tout au long de mon parcours.

A tous mes amies, Sarah, Fouzia, Yousra, Fatma, Noudjoud, Ikhlil et ma binome Assala qui m'ont toujours encouragé, et à qui je souhaite plus de succès.

Merci !

Karaoui & Kahima



Notations et abriviations

ADF Test de Dickey Fuller Augmenté

AIC Critère d'information d'Akaike.

ARIMA Autorégressif moyenne mobile intégré

ARMA Autorégressif moyenne mobile

SARIMA Autorégressif Moyenne Mobile Saisonnie

BB Bruit blanc

BIC Critère d'information bayésien

BJ Quantité de Jarque-Bera

Cov Covariance

Ds Non stationnarité stochastique

E Espérance

FAC Fonction d'autocorrélation

FACP Fonction d'autocorrélation partielle

MA Moyenne mobile

MAE Erreur absolue moyenne

MAPE Ecart absolu moyen en pourcentage

MSE Erreur quadratique moyenne

RMSE Racine carrée de l'erreur quadratique moyenne

TS Non stationnarité déterministe

INTRODUCTION GÉNÉRALE :

L'étude des séries temporelles correspond à l'analyse statistique des observations régulièrement espacées dans le temps. Cette analyse regroupe un ensemble de techniques et de méthodes qui permettent la détection de la structure de corrélation entre les composants du processus générateur de la série à étudier. Cette structure sera exploitée pour faire des prédictions dans le futur. Elle trouve ses applications dans de nombreux domaines, à savoir, l'économétrie, la météorologie ou la démographie, pour n'en citer qu'une petite partie.

On s'intéresse à l'évolution au cours du temps d'un phénomène, dans le but de décrire, d'expliquer puis de prévoir celui-ci phénomène dans le futur. On dispose ainsi d'observations à de différentes dates, c'est-à-dire, d'une suite de valeurs numériques indexées par le temps. Auparavant, l'analyse se résumait à une tentative d'ajustement des observations par un modèle préétabli par le calcul de certains paramètres. Ces méthodes, d'analyse statistique s'appliquent lorsque la série est stationnaire ayant peu de fluctuations. Elles atteignent rapidement leurs limites quand la vraie loi de variation de la série ne suit plus le cheminement qu'on voulait lui imposer. Dans ce cas, on dit que la série est affectée d'un ou d'une combinaison des phénomènes suivants : la tendance, la saisonnalité et/ou les différentes variations aléatoires.

Parmi les modèles les plus utilisées en analyse des séries chronologiques, on trouve les modèles autorégressifs moyenne mobile connus sous l'appellation des modèles ARMA. Cette classe de modèles a connu un développement remarquable grâce à l'apparition du théorème de WOLD, avec la publication du fameux ouvrage de Box et Jenkins « time series analysis, forecasting and control », en 1970, une nouvelle méthodologie a fait apparition pour la construction et la modélisation adéquate des séries. Cette méthodologie commence par une identification ARMA, suivie de l'estimation des paramètres et la validation et se termine par une prévision.

Cette nouvelle approche développée par Box et Jenkins est une véritable philosophie car, elle permet d'obtenir des prévisions de bonnes qualités et relativement précises et offre l'opportunité de juger si une loi de variation est satisfaisante avec une précision gérable. Parfois le modèle issu de la méthodologie de Box-Jenkins n'est pas tout à fait correct ou fidèle au modèle ou au phénomène réel mais ça permet d'avoir des indications supplémentaires permettant d'identifier le modèle le plus adéquat et la loi la plus correcte.

La méthodologie de Box-Jenkins est constituée essentiellement de trois phases, à savoir : l'identification du modèle ARMA avec détermination de ses paramètres et la validation qui se fait par l'application d'un certain nombre de tests assurant la stationnarité du modèle.

Ce mémoire, fait l'objet d'une étude prévisionnelle des approvisionnements en matière première de la cimenterie de Meftah, par application de la méthodologie de Box-Jenkins et la méthode de lissage exponentiel. Le travail est composé de deux parties, la première est dite étude théorique, constituée de trois chapitres : Un rappel des processus stochastiques, la présentation de la méthodologie de Box-Jenkins et le lissage exponentiel. La deuxième partie représente les résultats numériques, des deux méthodes, selon les données de la cimenterie.

Dans le premier chapitre, nous exposons les concepts fondamentaux des processus stochastique plus particulièrement les modèles auto régressifs et leurs différentes variantes, à savoir, les modèles : AR, MA, ARMA et ARIMA en étudiant leurs propriétés (la stationnarité, les fonctions d'auto-correlation simples et partielles,...); tout en passant par la détermination de leurs paramètres.

Le second chapitre sera axé sur la présentation de la méthodologie de Box-Jenkins. On commence par l'exposition des différents paramètres influant sur la série chronologique et sa stationnarité, à savoir, la tendance et la saisonnalité, puis on entame les étapes de la méthodologie de Box-Jenkins qui sont la modélisation et la validation passant par estimation des paramètres.

Le troisième, est consacré à l'application numérique de la méthodologie de Box-Jenkins sur deux séries chronologiques. La première série représente l'approvisionnement mensuel de la cimenterie de Meftah en matière première dans son état réel. La seconde série est issue de la première, en appliquant un lissage par la moyenne mobile pour tous les pics et les creux dont l'écart dépasse 50% de la moyenne.

Quatrième chapitre, fait l'objet de présentation de la méthode du lissage exponentiel et différentes variantes ainsi que leurs paramètres.

Dans le cinquième chapitre, nous afficherons les résultats numériques de la méthode du lissage exponentiel.

1.1 Définition d'une série chronologique :

On appelle série chronologique, ou temporelle, une suite d'observations numériques d'un processus stochastique $\{X_t, t \in \mathbb{Z}\}$. Ce processus évolue aléatoirement dans le temps.

Une série chronologique est dite discrète si l'ensemble des instants d'observations est discret, ou dénombrable, et elle est dite continue si l'ensemble des instants d'observations est continu, non dénombrable.

Il existe une gamme important de processus aléatoire dans la littérature. L'intérêt de l'analyse des chroniques était focalisé initialement uniquement les processus stationnaires qui ont la particularité d'avoir les propriétés (moyenne, variance, covariance, ...) invariantes dans le temps.[7]

1.1.1 Stationnarité :

Avant de traiter une série chronologique, il convient d'étudier ses caractéristiques stochastiques. Si elles (espérance, variance et covariance) se trouvent modifiées dans le temps, la série chronologique est considérée comme non stationnaire, sinon le processus est dit stationnaire, ce qui ebrainsque la série l'est aussi.

Pour le processus stationnaire on distingue deux types de stationnarités, la forte ou la stricte et la stationnarité faible dite aussi d'ordre deux.

La stationnarité forte :

Le processus $\{X_t, t \in \mathbb{Z}\}$ est dit processus stationnaire au sens strict si des distributions de probabilités conjointes sont identiques avec $t_{i+h} \in \mathbb{Z}, \forall i$ avec $i = 1, \dots, n$

$$P[X_{t1} < X_1, \dots, X_{tn} < X_n] \quad \text{et} \quad P[X_{t1+h} < X_1, \dots, X_{tn+h} < X_n]$$

Autrement dit, un processus est strictement stationnaire si tous ses moments sont invariables dans le temps.

La stationnarité de second ordre :

Le processus $\{X_t, t \in \mathbb{Z}\}$ est dite faiblement stationnaire si :

- La moyenne du processus existe et ne dépend pas du temps : $E(X_t) = m, \forall t \in \mathbb{Z}$.
- La variance existe et elle est invariante dans le temps : $E(X_t^2) < \infty, \forall t \in \mathbb{Z}$
- L'autocovariance dépend uniquement de l'écart des temps, elle ne dépend pas de l'instant t : $Cov(X_t, X_{t+h}) = \gamma_h, \forall t \in \mathbb{Z}$

Remarque :

- ❖ D'après la définition de la stationnarité forte et de la stationnarité faible on constate que la forte implique la faible mais pas l'inverse.
- ❖ Les processus stationnaires d'ordre deux sont des processus générateurs de chronique sans tendance en moyenne et sans tendance en variance mais cela ne signifie pas que les séries temporelles ont une représentation graphique stable.

1.2 Caractéristique d'une série chronologique :

1.2.1 Moyenne et variance :

Soit une série temporelle $\{X_t, t \in \mathbb{Z}\}; t = 1, \dots, n$ les expressions de la moyenne et de la variance sont respectivement :

$$\text{Moyenne : } E(X_t) = \frac{1}{n} \sum_{t=1}^n X_t$$

$$\text{Variance : } V(X_t) = \frac{1}{n} \sum_{t=1}^n [X_t - E(X_t)]^2$$

1.2.2 Fonction d'auto-covariance :

Définition

La fonction d'auto-covariance d'un processus aléatoire $\{X_t, t \in \mathbb{Z}\}$ mesure la covariance d'un couple de valeurs séparées par un intervalle de longueur h appelé retard. Elle fournit des informations sur variabilité de la série et sur les liaisons temporelles qui existent entre les différentes composantes de la série X_t . [2]

On appelle fonction d'auto-covariance du processus X_t la fonction γ suivante :

$$\forall h \in \mathbb{Z} : \gamma_h = Cov(X_t, X_{t+h}) = E[(X_t - E(X_t))(X_{t+h} - E(X_{t+h}))]$$

Propriétés :

La fonction d'auto-covariance vérifie les propriétés suivantes :

- ❖ $\gamma_0 = Cov(X_t, X_t) = E[(X_t - E(X_t))^2] = var(X_t) = \sigma_t^2 \geq 0$
- ❖ $|\gamma_h| \leq \gamma_0$
- ❖ γ est une fonction symétrique : pour tout $h \in \mathbb{N}, \gamma_{-h} = \gamma_h$

La fonction d'auto-covariance d'un processus aléatoire stationnaire $\{X_t, t \in \mathbb{Z}\}$ est estimée par la fonction d'auto-covariance empirique définie par :

$$r_h = \frac{1}{n} \sum_{t=h+1}^n (X_t - \bar{X})(X_{t-h} - \bar{X}) \quad \text{Avec :} \quad \bar{X} = \frac{1}{n} \sum_{t=1}^n X_t$$

1.2.3 Fonction d'auto-corrélation Simple :

Définition :

L'autocorrélation d'un processus stochastique est la corrélation de ce processus et lui-même avec un décalage dans le temps. Elle permet de révéler les irrégularités dans un processus stochastique, tel que les profils répétés, saisonnalité,...

On supposera par la suite que le processus $\{X_t, t \in \mathbb{Z}\}$ est stationnaire du second ordre.[9]

Le graphe représentatif de la fonction d'autocorrélation est appelé correlogramme. On appelle auto-corrélation simple du processus X_t la fonction ρ suivante :

$$\rho_h = \rho(X_t, X_{t-h}) = \frac{\text{cov}(X_t, X_{t-h})}{\sqrt{(\text{var}(X_t) * \text{var}(X_{t-h}))}} = \frac{\gamma_h}{\gamma_0}$$

Propriétés :

L'auto-corrélation simple est donc la fonction d'auto-covariance normalisée. Il vérifie des propriétés similaires, on a :

- 1) $\rho_0 = 1$.
- 2) $|\rho_h| \leq \rho_0, \forall h \in \mathbb{Z}$.
- 3) ρ est une fonction symétrique : $\forall h \in \mathbb{N} : \rho_h = \rho_{-h}$.

Dans le cas d'une série chronologique stationnaire, ou processus stationnaire, l'estimation de la fonction ρ_h est donnée par :

$$\hat{\rho}_h = \frac{\sum_{t=1}^{n-h} (X_t - \bar{X})(X_{t+h} - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2} \quad \text{Avec :} \quad \bar{X} = \frac{1}{n} \sum_{t=1}^n X_t$$

1.2.4 Fonction d'auto-corrélation partielle :

Définition :

Une seconde façon de mesurer l'influence d'un processus sur lui-même décalé dans le temps est de calculer la corrélation entre deux instants en enlevant la partie d'information entre ces deux instants.

Elle peut être calculée comme un rapport entre deux déterminations, comme indiqué la formule suivante :

$$\phi_{hh} = \frac{||\rho_h^*||}{||\rho_h||}$$

telle que :

$$P_h = \begin{pmatrix} 1 & \rho_1 & \cdots & \rho_{h-1} \\ \rho_1 & 1 & \cdots & \rho_{h-2} \\ \vdots & \cdots & 1 & \vdots \\ \rho_{h-1} & \rho_{h-2} & \cdots & 1 \end{pmatrix} ; \quad P_h^* = \begin{pmatrix} 1 & \rho_1 & \cdots & \rho_1 \\ \rho_1 & \vdots & \vdots & \rho_2 \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{h-1} & \cdots & \cdots & \rho_h \end{pmatrix}$$

$|P_h^*|$ = déterminant de la matrice P_h dans laquelle on a remplacé la dernière colonne par le vecteur $[\rho_1, \dots, \rho_h]$

On constate que la première valeur de l'autocorrélation partielle est égale à la première valeur de l'autocorrélation simple.[5]

Exemple :

Calculons ϕ_{11}, ϕ_{22} et ϕ_{33} :

On a :

$$\phi_{11} = \frac{|P_1^*|}{|P_1|} \quad \text{telque : } P_1 = [1]; P_1^* = [\rho_1]$$

Donc :

$$\phi_{11} = \frac{\rho_1}{1} = \rho_1$$

On constate que la première valeur de l'auto-corrélation partielle est égale à la première valeur de l'auto corrélation.

On a :

$$\phi_{22} = \frac{|P_2^*|}{|P_2|} \quad \text{telque : } P_2 = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix} ; \quad P_2^* = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{bmatrix}$$

Donc :

$$\phi_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

Pour ϕ_{33} :

$$\phi_{33} = \frac{|P_3^*|}{|P_3|} \quad \text{telque : } P_3 = \begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix} ; \quad P_3^* = \begin{bmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{bmatrix}$$

Donc :

$$\phi_{33} = \frac{\rho_1^3 - \rho_1 \rho_1 (1 - \rho_2) + \rho_3 (1 - \rho_1^2)}{1 - \rho_2^2 - 2\rho_1^2 (1 - \rho_2)}$$

On peut montrer que la fonction d'auto-corrélation partielle s'écrit :

$$\phi_{ii} = \begin{cases} \frac{\rho_i - \sum_{j=1, i=1} \phi_{i-1, \rho_{i-j}}}{1 - \sum_{j=1, i=1} \phi_{i-1, \rho_{i-j}}}, & i \geq 2 \\ \rho_1, & i = 1 \end{cases}$$

$\phi_{ij} = \phi_{i-1, j} - \phi_{ii} \phi_{i-1, i-j}$ avec : $j = 1, \dots, i-1$

Cet algorithme est connu sous le nom d'algorithme de Durbin.

1.2.5 Processus bruit blanc :

Définition :

On appelle processus bruit blanc $\{\varepsilon_t, t \in \mathbb{Z}\}$ une suite de variables X_t non corrélées et de moyenne nulle et variance σ_ε^2 , si les variables X_t sont normales donc c'est un bruit blanc gaussien.[3] Le processus $\{\varepsilon_t, t \in \mathbb{Z}\}$ est considéré comme centré si :

- ❖ $E(\varepsilon_t) = 0; \forall t \in \mathbb{Z}$.
- ❖ $Var(\varepsilon_t) = \sigma^2; \forall t \in \mathbb{Z}$.
- ❖ $Cov(\varepsilon_t, \varepsilon_{t-h}) = \begin{cases} \sigma_\varepsilon^2 & , h = 0 \\ 0 & , \text{sinon} \end{cases}$

Remarque :

- Si $\{\varepsilon_t\}$ est un bruit blanc faible ; on notera $\{\varepsilon_t\} \rightsquigarrow N(0, \sigma^2)$.
- Si $\{\varepsilon_t\}$ est un bruit blanc fort ; on notera $\{\varepsilon_t\} \rightsquigarrow IID(0, \sigma^2)$.

1.3 Série chronologique :

Avant d'entamer l'étude d'une série chronologique, il est nécessaire de définir ses différentes composantes. La première étape de l'étude d'une série temporelle est la représentation graphique. Cette étape donne un aperçu global du phénomène en question. Il révèle des indicateurs précieux pour le choix du modèle adéquat.

Ainsi, à partir du graphe, on peut détecter les irrégularités ou les anomalies qui apparaissent sous forme de croissance, rupture et / ou fluctuation. L'examen d'une série met en relief quatre composantes fondamentales.

- ❖ La tendance, ou trend noté(T_t) : Représente l'évolution à long terme de la série étudiée. Elle traduit le comportement « moyen » de la série.
- ❖ La composante saisonnière (S_t) : Elle correspond à un phénomène qui se répète à des intervalles de temps relativement réguliers, ou périodique. En général, la série est affectée d'une saisonnalité lorsqu'elle modélise un processus saisonnier comme la variation de la température, le taux de glycémie dans le sang. . . .
- ❖ La composante cyclique (C_t) : Il s'agit d'un phénomène qui se répète à des intervalles de temps réguliers comme la saisonnalité sauf que là, la longueur des intervalles du cycle est beaucoup plus importante que la longueur des intervalles des saisons, d'une part et d'autre, dans un cycle on peut trouver la saisonnalité ainsi que la tendance.
- ❖ La composante résiduelle ou bruit, noté (ε_t) : Elle correspond à des fluctuations Irrégulières, généralement, de faible intensité mais de nature aléatoire. On parle aussi d'aléas.

1.4 Décomposition d'une série chronologique :

Une série chronologique se décompose en deux parties :

- Une partie déterministe qui est représenté par une composante tendancielle D_t et/ou une composante saisonnière S_t .
- Une partie aléatoire à moyenne nulle stationnaire, noté W_t .

Selon la composition des composantes de la série, nous distinguons plusieurs modèles :

1.4.1 Le modèle additif :

Nous considérons dans cette section une série $\{X_t, t \in \mathbb{Z}\}$ admettant une décomposition additive :

$$X_t = Z_t + S_t + W_t, \quad t = 1, \dots, n$$

Où D_t est la composante tendancielle, S_t la composante saisonnière et W_t représente l'erreur ou l'écart du modèle.

Graphiquement, l'amplitude des variations est constante autour de la tendance, cela veut dire que la composante saisonnière est constante sur chaque période P :

$$S_{t+P} = S_t, \quad \forall t$$

Cela revient à dire que l'effet saisonnier est à moyenne nulle sur une période, ce qui signifie :

$$\sum_{i=1}^P S_i = 0$$

Les erreurs sont des variables aléatoires centrées. On considère, souvent, un bruit blanc, c'est-à-dire une suite de variables aléatoires réelles telles que :

$$E(W_t) = 0 \quad \text{et} \quad E(W_t W_{t'}) = \sigma^2 \delta_{tt'}$$

Dans un modèle additif, on suppose que les trois composantes : tendance, variations saisonnières et variations résiduelles sont indépendantes les unes des autres.

1.4.2 Modèle multiplicatif :

Soit $\{X_t, t \in \mathbb{Z}\}$, une chronique admettant une décomposition multiplicative :

$$X_t = D_t(1 + S_t)(1 + W_t), \quad t = 1, \dots, n$$

La composante saisonnière vérifie :

$$\sum_{i=1}^P S_i = 1$$

Dans ce modèle, l'amplitude de la série n'est plus constante au cours du temps, elle varie proportionnellement à la tendance D_t à un bruit près.

Graphiquement, l'amplitude des variations (saisonnières) varie.

1.4.3 Modèles mixtes (hybrides) :

Il représente les différentes combinaisons des modèles additifs et multiplicatifs, d'écrit l'équation suivante :

$$X_t = (D_t + S_t)(1 + W_t), \quad t = 1, \dots, n$$

1.4.4 Choix du modèle :

Avant toute modélisation ou étude approfondie d'une série chronologique, on doit déterminer son type, si elle représente un modèle additif, multiplicatif ou mixte. Pour faire la distinction entre ces modèles, on peut procéder par la méthode graphique ou la méthode analytique.

Méthode graphique :

1. Méthode du profil : on utilise le graphique des courbes superposées
 - Si ces 2 droites sont à peu près parallèles, le modèle est dit additif.
 - les pics et creux s'accroissent, alors le modèle est dit multiplicatif.
2. Méthode de la bande : on utilise le graphe de la série et la droite passant par le minimum et celle passant par les maximums.
 - Si ces 2 droites sont à peu près parallèles alors le modèle est additif.
 - Si ces 2 droites ne sont pas parallèles alors le modèle est multiplicatif.

Méthode analytique :

On calcule les moyennes et les écarts types pour chacune des périodes considérées par la droite des moindres carrés : $\sigma = a\bar{x} + b$.

Si a est nul, le modèle est additif, sinon c'est le modèle multiplicatif.

Analyse de la tendance :

- 1) Ajustement tendanciel par la méthode des moindres carrés : si la tendance a l'allure d'une fonction linéaire, on ajuste la tendance par la méthode des moindres carrés d'où l'expression analytique de la tendance D_t en fonction de t .
- 2) Ajustement tendanciel non linéaire : il suffit juste de faire un changement de variable approprié pour se ramener à un ajustement linéaire.
- 3) La méthode de moyenne mobile (MM) : cette technique est utilisée pour éliminer ou amortir les moments cycliques, saisonniers et résiduels.

Le principe est de construire une nouvelle série obtenue en calculant des moyennes arithmétiques successives de longueur ' P ', fixées à partir des données originales. Chacune des moyennes obtenues correspondra au milieu de la période P pour laquelle la moyenne vient d'être calculée.

- Si p est impair : $p = 2m + 1$:

$$MM_p(t) = \frac{1}{P} \sum_{i=1}^m X_{t+i}$$

- Si p est pair : $p = 2m$:

$$MM_p(t) = \frac{1}{p} \left[\frac{X_{t-m}}{2} + \sum_{i=-m+1}^{m-1} X_{t+1} + \frac{X_{t+m}}{2} \right]$$

1.4.5 Estimation des variations saisonnières :

Cas d'un modèle additif :

- On calcule des données sans tendance : $X_t - D_t$
- On calcule la moyenne des données sans tendance du mois j sur les années ceci par chacun des p saisonniers.

D'où :

$$S_j = \frac{1}{n} \sum_{i=1}^n (X_{i,j} - D_{i,j})$$

- On calcule la moyenne des S_j :

$$\bar{S} = \frac{1}{p} \sum_{j=1}^p S_j$$

- Si $\bar{S} \neq 0$, on corrige les S_j :

$$S'_j = S_j - \bar{S}$$

Cas d'une modèle multiplicatif :

- On calcule X_t/D_t
- On calcule la moyenne des données sans tendance du mois notée j sur les années noté n par chacun des p saisonniers.

$$S_j = \frac{1}{n} \sum_{i=1}^n \frac{X_{ij}}{D_{ij}}$$

- On calcule la moyenne des S_j :

$$\bar{S} = \frac{1}{p} \sum_{j=1}^p S_j$$

- Si $\bar{S} \neq 1$, on corrige les S_j par :

$$S'_j = \frac{S_j}{\bar{S}}$$

Conclusion :

Pour décomposer une série chronologique, on doit commencer par :

- Tracer son graphique ;
- Estimer la tendance D_t ;
- Choisir un modèle de compositions (additif ou multiplicatif) ;

- Estimer les variations saisonnières ;

1.5 Les opérateurs linéaires :

1.5.1 Opérateur de retard B :

L'opérateur de retard B est un opérateur linéaire qui décale le processus d'une unité de temps vers le passé : $BX_t = X_{t-1}$.

Si on applique n fois cet opérateur, on décale le processus n unités de temps :

$$B(B(\dots B(X_t)\dots)) = B^n X_t = X_{t-n}$$

On posera $B^0 = 1$. Si l'on applique le polynôme de retard $\phi(B)$ défini comme suit : $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ à une série X_t , on a :

$$\phi(B)X_t = X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p}$$

Où ϕ_i sont des nombres ou coefficients réels $i = 1, \dots, p$.

1.5.2 Opérateur d'avance F :

L'opérateur d'avance F est un opérateur linéaire défini de la manière suivante : $FX_t = X_{t+1}$ plus généralement $F^n X_t = X_{t+n}$.

Ces opérateurs sont inversibles :

$$F^{-1} = B \quad \text{et} \quad B^{-1} = F.$$

1.5.3 Opérateur de différence ordinaire :

L'opérateur Δ fait la différence entre le processus et sa version décalée d'un unité de Temps. Cet opérateur se construit en utilisant l'opérateur retard B :

$$\Delta X_t = X_t - X_{t-1} = X_t - BX_t \Leftrightarrow \Delta = 1 - B$$

On définit le d^{ime} opérateur de différence ordinaire par :

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

De façon générale, l'opérateur Δ^d élimine les tendances polynomiales de degré d .

1.5.4 Opérateur de différence saisonnière :

L'opérateur $\Delta^s = (1 - B^s)$ Associé à un processus $\{X_t, t \in \mathbb{Z}\}$ élimine une saisonnalité de période $s, s \in \mathbb{N}$. Soit le processus $\{\Delta_s X_t, t \in \mathbb{Z}\}$ tel que :

$$\Delta_s X_t = X_t - X_{t-s} = (1 - B^s)X_t$$

On définit le d^{eme} opérateur de différence saisonnière par :

$$\Delta_s^d = (1 - B^s)^d X_t$$

1.6 Typologie des modèles Auto Régressifs Moving Average (ARMA) et ses dérivées :

Nous allons présenter une famille de processus aléatoires qui sont censés recouvrir une large gamme d'évolutions possibles de la série chronologiques : les processus autorégressifs et les processus de moyenne mobile.

1.6.1 Modèle AR (Auto Régressif) :

Définition :

Dans le processus autorégressif d'ordre p , l'observation présente X_t , est générée par une moyenne pondérée des observations passées jusqu'à la p^{eme} période sous la forme suivante :

$$AR(1) : X_t = \phi_1 X_{t-1} + \varepsilon_t + c$$

$$AR(2) : X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t + c$$

⋮

$$AR(p) : X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t + c$$

Où $\phi_1, \phi_2, \dots, \phi_p$ sont des paramètres à estimer positifs ou négatifs, ε_t est un processus aléatoire non nécessairement gaussien.

Nous pouvons ajouter à ce processus une constante qui ne modifie en rien les propriétés stochastiques.[1] Les équations peuvent aussi s'écrire à l'aide de l'opérateur de décalage B , comme c'est indiqué par l'équation suivante :

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t = \varepsilon_t + c \Leftrightarrow \phi(B) X_t = \varepsilon_t + c$$

Théorème 1.6.1 (condition de stationnarité). *Une condition nécessaire et suffisante pour que le processus soit stationnaire est que les racines du polynôme caractéristique $\phi(Z)$ soient en valeur absolue inférieures à 1, c'est-à-dire que les racines sont à l'intérieur du cercle unitaire, avec :*

$$\phi(Z) = Z^p - \phi_1 Z^{p-1} - \dots - \phi_p Z^{p-p}$$

Théorème 1.6.2 (condition d'inversibilité). *Un processus $AR(p)$ est toujours inversible.*

Caractéristique d'un processus AR(p) :

L'espérance d'un processus AR(p) :

Soit $X_t AR(p)$ stationnaire, on a :

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

$$D'où : E(X_t) = E(c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t)$$

$$E(X_t) = c + \phi_1 E(X_{t-1}) + \phi_2 E(X_{t-2}) + \dots + \phi_p E(X_{t-p}) + E(\varepsilon_t) \text{ avec } E(\varepsilon_t) = 0$$

$$E(X_t) = c + E(X_t) [\phi_1 + \phi_2 + \dots + \phi_p]$$

$$E(X_t) [1 - \phi_1 - \phi_2 - \dots - \phi_p] = c$$

donc :

$$E(X_t) = \frac{c}{1 - \phi_1 - \phi_2 - \dots - \phi_p} = \frac{c}{1 - \sum_{i=1}^p \phi_i}$$

Exemple :

$$\text{soit } : X_t = 0.8 + 0.7X_{t-1} + 0.25X_{t-2} + \varepsilon_t \text{ avec } : \varepsilon_t \rightsquigarrow BB(0, \sigma^2)$$

$$X_t \rightsquigarrow AR(2)$$

$$E(X_t) = \frac{0.8}{1 - 0.7 - 0.25} = 1$$

caractéristiques de la fonction d'autocovariance d'un AR(p) :

La fonction d'autocovariance γ_h d'un processus AR(p) $\{X_t, t \in \mathbb{Z}\}$ satisfait une relation récurrente de la forme :

$$\gamma = \begin{cases} \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma_\varepsilon^2 & h = 0 \\ \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_{h-p} & h > 0 \end{cases}$$

Avec : $\gamma_h = \gamma_{-h}, \forall h \in \mathbb{Z}$

Preuve :

On considère la définition de X_t :

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

Les autocorrélations d'un processus AR(p) peuvent être calculées en multipliant chaque nombre de l'équation (1) par $X_{t-h}, (h > 0)$, on pose que $E(\varepsilon_t) = 0$ et $c = 0$ on aura donc :

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

En prenant ensuite l'espérance des variables, on obtient la relation suivante :

$$E(X_t X_{t-h}) = \gamma_h = \phi_1 E(X_{t-1} X_{t-h}) + \phi_2 E(X_{t-2} X_{t-h}) + \dots + \phi_p E(X_{t-p} X_{t-h}) + E(X_{t-h} \varepsilon_t)$$

puisque $E(\varepsilon_t X_{t-h}) = 0$ car X_{t-h} ne dépend que des ε_{t-h-i} avec $i \geq 0$ de la même façon

$$\gamma_0 = \phi_1 E(X_{t-1} X_t) + \phi_2 E(X_{t-2} X_t) + \dots + \phi_p E(X_{t-p} X_t) + E(X_t \varepsilon_t)$$

$$\gamma_0 = \phi_1 \gamma_{h-1} + \phi_2 \gamma_{h-2} + \dots + \phi_p \gamma_{h-p} + E(X_t \varepsilon_t)$$

$$\gamma_0 = \phi_1 \gamma_{h-1} + \phi_2 \gamma_{h-2} + \dots + \phi_p \gamma_{h-p} + \sigma_\varepsilon^2$$

Car $E(\varepsilon_t X_t) = \sigma_\varepsilon^2$ et puisque X_t peut s'écrire sous la forme d'une somme pondérée des chocs passés (théorème de Wold, 1954) :

$$X_t = \phi(B)^{-1} \varepsilon_t = \sum_{i=0}^{\infty} \theta_i \varepsilon_{t-i} \quad \text{avec : } \theta_0 = 1$$

Caractéristique de la fonction d'auto-corrélation d'un processus AR(p) :

On sait par définition que :

$$\rho_h = \frac{\gamma_h}{\gamma_0}$$

La fonction d'auto-corrélation, notée ρ_h , d'un processus AR(p) $\{X_t, t \in \mathbb{Z}\}$ satisfait une relation récurrente de la forme :

$$\phi(B) \rho_h = 0 \Leftrightarrow \rho_h = \begin{cases} 1 & \text{si } h = 0 \\ \phi_1 \rho_{h-1} + \phi_2 \rho_{h-2} + \dots + \phi_p \rho_{h-p} & \text{si } \forall h \in \mathbb{Z}^* \end{cases}$$

Ces relations sont connues sous le nom des équations de Yule-Walker.

L'auto-corrélation d'ordre h est déterminée par une équation aux différences homogènes d'ordre h dont on peut avoir la solution générale.

Caractéristiques des corrélogrammes :

Il a été démontré que le corrélogramme simple d'un processus AR(p) est caractérisé par une décroissance géométrique de ses termes de type :

$$p_h = p^h$$

Le corrélogramme partiel a ses p premiers termes différents de 0.

1.6.2 Modèle Moving Average (MA) :

Définition :

Dans le processus moyenne mobile d'ordre q , chaque observation X_t , est générée par une moyenne pondérée d'aléas jusqu'à la q^{eme} période.

$$\text{MA}(1) : X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

$$\text{MA}(2) : X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

⋮

$$\text{MA}(q) : X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

Où $\theta_1, \theta_2, \dots, \theta_q$ sont des paramètres pouvant être positifs ou négatifs, ε_t est un aléa gaussien. [8]

L'équation MA(q) peut s'écrire :

$$(1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \varepsilon_t = X_t \iff \theta(B) \varepsilon_t = X_t$$

Théorème 1.6.3 (condition de stationnarité). *Le processus $\{X_t, t \in \mathbb{Z}\}$ moyenne mobile d'ordre q est toujours stationnaire car il est une combinaison linéaire finie d'un processus stationnaire ε_t .*

Théorème 1.6.4 (condition d'inversibilité). *Une condition nécessaire et suffisante pour que le processus moyenne mobile soit inversible est que les racines du polynôme caractéristique $\theta(Z)$ soient en valeur absolue inférieure à 1.*

avec :

$$\theta(Z) = Z^q - \theta_1 Z^{q-1} - \dots - \theta_q Z^{q-p}$$

caractéristique de la fonction d'autocovariance MA(q) :

la fonction d'autocovariance γ_h d'un processus MA(q) $X_t, t \in \mathbb{Z}$ définie par $E[X_t X_{t-h}]$ est donnée par la relation :

$$\gamma_h = \begin{cases} (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma_\varepsilon^2 & h = 0 \\ (-\theta_h + \theta_1 \theta_{h+1} + \dots + \theta_{q-h} \theta_q) \sigma_\varepsilon^2 & 0 < h \leq q \\ 0 & h > q \end{cases}$$

preuve :

Pour obtenir ces résultats, il suffit de rappeler que $E(\varepsilon_t \varepsilon_{t-i}) = 0$.

si $i \neq 0$ et $E(\varepsilon_t^2) = \sigma_\varepsilon^2, \forall i$ on a :

$$\gamma_h = E[(\varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q})(\varepsilon_{t-h} - \theta_1 \varepsilon_{t-h-1} - \dots - \theta_q \varepsilon_{t-h-q})]$$

Si les conditions d'inversibilité sont vérifiées et en développant cette expression on retrouve le résultat général énoncé ci-dessus.

si $X_t \rightsquigarrow MA(q)$ inversible X_t admet la représentation AR(∞) :

$$\varepsilon_t = \theta(B)^{-1} X_t = \sum_{i=0}^{\infty} \beta_i X_{t-i} \quad \text{telque} \quad \beta_0 = 1, \sum |\beta_i| < \infty$$

Caractéristiques de la fonction d'autocorrélation d'un MA(q) :

En divisant la fonction d'autocovariance par la variance on obtient le système :

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}$$

la fonction d'autocorrélation :

$$\rho_h = \frac{\gamma_h}{\gamma_0} = \begin{cases} 1 & h = 0 \\ \frac{-\theta_h + \theta_1\theta_{h+1} + \dots + \theta_{q-h}\theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} & 0 < h \leq q \\ 0 & h > q \end{cases}$$

La fonction d'autocorrélation d'un processus MA(q) s'annule à l'ordre $q + 1$.

Caractéristiques de la fonction d'autocorrélation partielle d'un processus MA(q) :

Pour calculer les autocorrélation partielles d'un processus MA(q) $\{X_t, t \in \mathbb{Z}\}$ définie par :

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

On utilise l'algorithme de Durbin.

Caractéristiques des corrélogrammes :

- Le corrélogramme simple à ses seuls q premiers termes différentes de zéro .
- Le corrélogramme partiel se comporte comme une exponentielle ou une sinusoidale amortie .

1.6.3 Modèle Auto Régressif Moyenne Mobile ARMA :**Définition :**

Les modèles ARMA(p,q) sont donc représentatifs d'un processus généré par une combinaison des valeurs passées et des erreurs passées. Ils sont définis par l'équation :

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \varepsilon_t$$

Qui est équivalente à :

$$\phi(B) X_t = \theta(B) \varepsilon_t$$

où $\phi(B)$ et $\theta(B)$ sont des polynômes de degrés p et q respectivement, ϕ_i et θ_j , $i = 1, \dots, p; j = 1, \dots, q$ sont des coefficients à estimer, et ε_t est un bruit blanc.[4]

Théorème 1.6.5. (condition de stationnarité et d'inversibilité) : Un processus ARMA(p,q) est stationnaire, si le polynôme $\phi(Z)$ a toutes ses racines en valeurs absolues inférieures à 1, et inversible si toutes les racines du polynôme $\theta(Z)$ sont en valeurs absolues inférieures à 1.

Caractéristiques des corrélogrammes :

Les corrélogrammes simples et partiels sont, par voie de conséquence, un mélange des deux corrélogrammes des processus AR et MA purs (fonction exponentielles sinusoidales amorties). Il s'avère ainsi plus délicat d'identifier ces processus à partir de l'étude des fonctions d'autocorrélation empirique.

Condition d'utilisation :

- Les modèles AR, MA et ARMA ne sont représentatifs que de chronique :
- Stationnaires en tendance.
 - Corrigées des variations saisonnières.

1.6.4 Modèle Auto Régressif Moyenne Mobile Saisonnier SARMA :**Définition :**

Un processus stationnaire $\{X_t, t \in \mathbb{Z}\}$ satisfait une représentation ARMA saisonnière ou (SARMA), notée $ARMA_{s,s}, (p, q)$, si :

$$\sum_{j=0}^p \phi_{js} X_{t-js} = c + \sum_{j=0}^q \phi_{js'} \varepsilon_{t-js'}$$

Avec : $c \in \mathbb{R}, \forall j < p, \phi_j \in \mathbb{R}, \forall j < q, \phi_j \in \mathbb{R}, \phi_0 = \theta_0 = 1$ et ε_t i.i.d $(0, \sigma_\varepsilon^2)$

et s désigne la période de la composante AR et s' désigne la période de saisonnalité de la composante MA.[6]

1.7 Les processus aléatoires non stationnaires :

Les chroniques économiques sont rarement des réalisations de processus aléatoire stationnaire. Pour analyser la non stationnarité, deux types de processus sont à identifier :

- Les processus *TS* (Trend Stationary) qui représentent une non-stationnarité de type déterministe.
- Les processus *DS* (Differency Stationary) pour les processus non stationnaires aléatoires.

1.7.1 processus TS :

Un processus TS s'écrit : $X_t = f_t + \varepsilon_t$ où f_t est une fonction polynôme du temps, linéaire ou non linéaire, et ε_t un processus stationnaire. Le processus *TS* le plus simple et le plus répandu est représenté par une fonction polynomiale 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 ε_t est un bruit blanc (gaussien ou non), les caractéristiques de ce processus sont alors :

- 1) $E(X_t) = a_0 + a_1 t + E[\varepsilon_t] = a_0 + a_1 t$
- 2) $V(X_t) = 0 + v[\varepsilon_t] = \sigma_\varepsilon^2$
- 3) $cov(X_t, X_{t'}) = 0$ pour $t \leq t'$

Ce processus TS est non stationnaire car $E(X_t)$ dépend du temps. connaissant a_0 et a_1 , le processus X_t , peut être stationnarisé en retranchant la valeur de X_t en t la valeur estimée.

Dans ce type de modélisation, l'effet produit par un choc (ou par plusieurs chocs aléatoires) à un instant t est transitoire. Le modèle étant déterministe, la chronique retrouve son mouvement de long terme qui est ici la droite de la tendance.

1.7.2 processus DS :

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

$$(1 - B)^d X_t = \delta + \varepsilon_t \Leftrightarrow X_t = X_{t-1} + \delta + \varepsilon_t$$

L'introduction de la constante σ dans le processus DS permet de définir deux processus différents :

- $\delta = 0$: le processus DS est dit sans dérivées, il s'écrit :

$$X_t = X_{t-1} + \varepsilon_t$$

Comme ε est un bruit blanc, le processus DS porte le nom du modèle de marche au hasard ou de la marche aléatoire (Random Walk Model), dite aussi la marche du toxicomane. Il est très fréquemment utilisé pour analyser l'efficacité des marchés financiers.

- $\delta \neq 0$, le processus porte le nom de processus DS avec dérivées, il s'écrit :

$$X_t = X_{t-1} + \delta + \varepsilon_t$$

1.8 Modèles non stationnaires ARIMA et SARIMA :

En pratique, l'hypothèse de stationnarité des processus n'est généralement pas vérifiée ce qui rend leur étude délicate, compte tenu du fait que la plupart des méthodes d'analyse supposent cette hypothèse réalisée.

Pour aborder l'étude des séries non stationnaires du type ARIMA et SARIMA, On peut intégrer certains types de non-stationnarité en élargissant le modèle ARMA.

La tendance ou la saisonnalité.

1.8.1 Modèle ARIMA :

On a vu que si un processus X_t admet une tendance polynomial de degré d , le processus différencié d fois est stationnaire :

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

Le modèle ARIMA revient à appliquer un modèle ARMA sur le processus différencié :

$$Y_t = ARMA(p, q) \Leftrightarrow X_t = ARIMA(p, d, q)$$

L'équation d'un modèle ARIMA(p,d,q) est donc donnée par :

$$\phi(B)(1 - B)^d X_t = \theta(B)\varepsilon_t$$

où ϕ et θ sont deux polynômes de degrés respectifs p et q .

- Le 'I' de ARIMA signifie 'integrated' comme réciproque de la différenciation 'd' est un entier positif.
- ε_t est un bruit blanc ainsi, ajuster un processus ARIMA(p,d,q) sur la série non stationnaire X_t est équivalent à ajuster un processus ARIMA(p,q) sur la série stationnaire $(1 - B)^d X_t$.
- Les processus ARMA apparaissent comme un cas particulier des processus ARIMA dans lesquels $d = 0$.

1.8.2 Modèle SARIMA :

La saisonnalité est un autre facteur de non-stationnarité. On a vu qu'une façon simple d'éliminer une saisonnalité de période s consiste à appliquer l'opération Δ_s .

$$Y_t = \Delta_s X_t - X_{t-s} = (1 - B^s)X_t$$

De façon générale, on peut supposer que l'influence des choses se transmet entre dates distantes d'un nombre entier de période selon un processus ARIMA(P,D,Q) :

$$\phi_s(B^s)\Delta_s^D X_t = \phi_s(B^s)\varepsilon_t$$

Où ε_t est un bruit blanc.

Un tel modèle est noté SARIMA(p,d,q) \times (P,D,Q)_s et son équation générale est :

$$\phi(B)\phi_s(B^s)\Delta^d\Delta_s^D X_t = \theta(B)\theta_s(B^s)\varepsilon_t$$

Où ϕ , ϕ_s , θ et θ_s sont des polynômes des degrés respectifs p, P, q, Q .

L'identification des paramètres (p, d, q) et (P, D, Q) se fait de la manière suivante : les paramètres non stationnaires (p, d, q) sont identifiés comme dans le cas d'un processus ARMA c'est-à-dire en examinant la coupure des fonctions d'auto-corrélations simples et partielles.

De même pour les paramètres saisonniers (P, D, Q) , on examine la coupure des fonctions d'autocorrélation partielles et d'autocorrélation à des ordres multiples de S respectivement (S étant la période). Il est noté que D est le nombre de fois où l'opérateur de différence saisonnière d'ordre S a été appliqué. De même pour les paramètres saisonniers (P, D, Q) , on examine la coupure des fonctions d'autocorrélation partielles et d'autocorrélation à des ordres multiples de S respectivement (S étant la période). Il est noté que D est le nombre de fois où l'opérateur de différence saisonnière d'ordre S a été appliqué.

2.1 Introduction

La partie autorégressive d'un processus, notée AR, est constituée par une combinaison linéaire finie des valeurs passées du processus en étude. La partie moyenne mobile, notée MA, est constituée d'une combinaison linéaire finie en fonction de t des valeurs passées d'un bruit blanc. Wold(1954) montre que les modèles ARMA de permettent de représenter la plupart des processus stationnaire.

L'approche de Box-Jenkins (1970) consiste à étudier systématiquement les séries chronologiques à partir de leurs caractéristiques afin de déterminer, dans la famille des modèles ARIMA, celui le plus adapté à représenter le phénomène en question, en fait elle représente une similitude très prononcée avec la méthode scientifique qui rappelle nous, consiste à formuler la supposition sous forme d'un modèle et à le mettre à l'épreuve et à réviser le modèle ARMA connu.

On détermine mécaniquement les prévisions à court terme des valeurs futures du processus en étude. Par fois, il est nécessaire représenter la tendance et la saisonnalité, dans ce cas, on étend la classe du modèle aux modèles ARIMA et SARIMA

La démarche suivie par la méthode de Box-Jenkins s'effectue en trois étapes de toutes les méthodes d'analyse prévisionnelle.. Le schémat suivant illustre les différentes étapes de la méthodologie de Box-Jenkins :

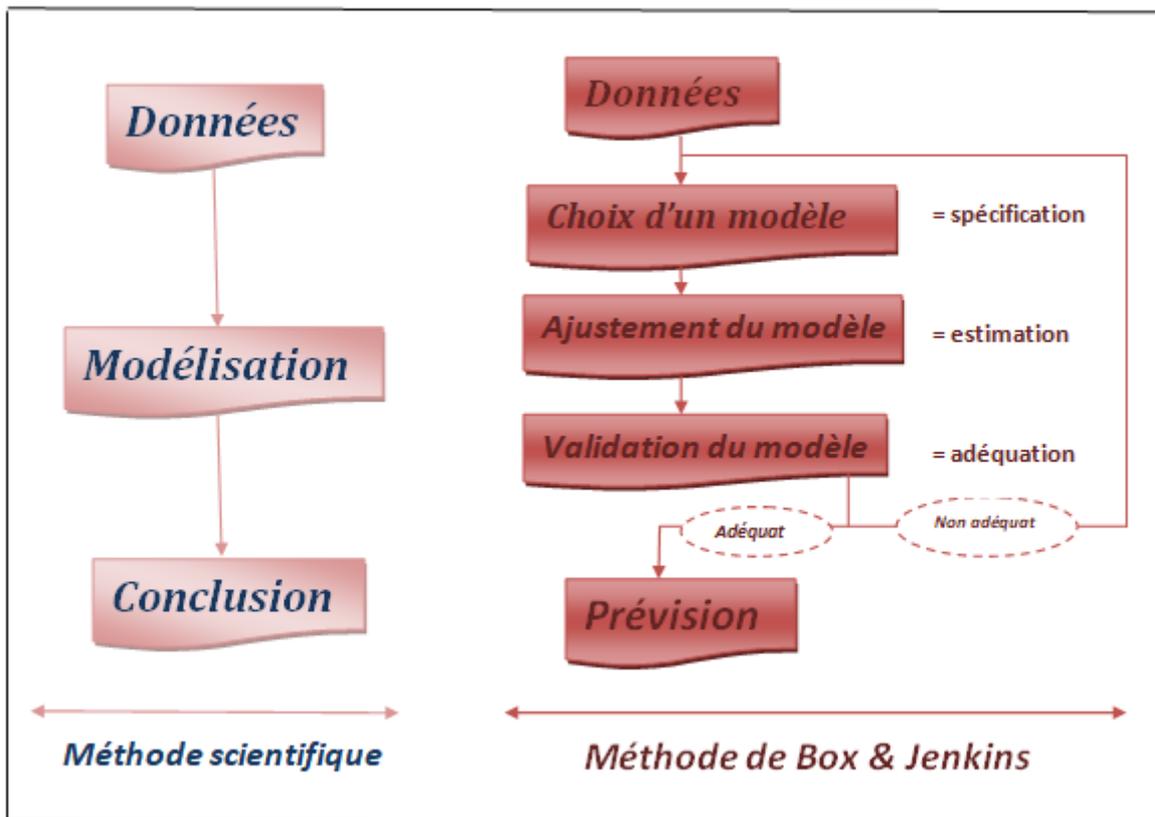


FIGURE 2.1 – Diagramme des étapes de la méthode de Box-Jenkins

2.2 Notions sur la méthodologie de Box-Jenkins

2.2.1 Familiarisation avec les données :

Cette étape permet à l'utilisateur de s'informer sur le domaine dont il prélève les données à savoir : leur nature (journalières, hebdomadaires, mensuelle, annuelle...) les éléments influents et les objectifs (ponctuels ou par intervalle).

La familiarisation avec les données se fait par l'examen graphique de la série qui peut révéler des changements structurelles, tel que :

- L'existence d'une saisonnalité ou de tendance.
- L'existence de point de rupture

2.2.2 Analyse préliminaire :

Cette étape fournit à l'utilisateur des outils lui permettant de corriger les valeurs aberrantes, compléter celle manquantes et enfin transformer les données et cela pour se ramener à un modèle stationnaire, car la méthodologie de Box-jenkins est conditionnée par la stationnarité de la série.

2.2.3 Stationnarisation de la série :

L'examen graphique souvent ne suffit pas pour mettre en évidence une saisonnalité, donc pour confirmer la non stationnarité de la série on utilise les tests suivants :

a. Test de saisonnalité (Fisher dit aussi ANOVA) :

Ce test permet de détecter la saisonnalité ou pour confirmer son existence lorsque on l'a déjà suspectée sur le graphe.

$$\begin{cases} H_0 = & \text{« La série est non saisonnière »} \\ H_1 = & \text{« La série est saisonnière »} \end{cases}$$

On considère :

n : Le nombre d'années,

p : Le nombre d'observations dans d'année.

Le principe est de calculer la valeur $F_0 = Var_p / Var_n$ puis de comparer à la valeur tabulée $F_{v_1 v_2}^\alpha$. Avec : $V_1 = (p - 1)$, $V_2 = (n - 1)(p - 1)$ degré de liberté (pour la saisonnalité).

$$V_1 = (n - 1), V_2 = (n - 1)(p - 1) \text{ degré de liberté (pour la tendance).}$$

Si $F_0 > F_{v_1 v_2}^\alpha$ alors on rejette H_0 et on considère qu'il y a saisonnalité, sinon on dit que la série est non saisonnière.

b. Test de la tendance (Dickey Fuller) :

le test de Dickey-Fuller (DF) permet de mettre en évidence le caractère stationnaire ou non d'une série par la détermination d'une tendance de type déterministe (TS) ou aléatoire (DS).

Modèle [3] : « Test sur la tendance »

$$\Delta X_t = \theta X_{t-1} + c + \beta t + \sum_{j=1}^d \theta_j \Delta X_{t-j} + \varepsilon_t$$

L'hypothèse du test est :

$$\begin{cases} H_0 : & t = 0 \\ H_1 : & t \neq 0 \end{cases}$$

Modèle [2] : « Test sur la constance »

$$\Delta X_t = \theta X_{t-1} + c + \sum_{j=1}^d \theta_j \Delta X_{t-j} + \varepsilon_t$$

L'hypothèse du test est :

$$\begin{cases} H_0 : & C = 0 \\ H_1 : & C \neq 0 \end{cases}$$

Modèle [1] : « Test sur la racine unitaire »

$$\Delta X_t = \theta X_{t-1} + \sum_{j=1}^d \theta_j \Delta X_{t-j} + \varepsilon_t$$

L'hypothèse du test est :

$$\begin{cases} H_0 : & \theta = 0 \\ H_1 : & \theta \neq 0 \end{cases}$$

modèle le plus susceptibles de représenter la série.

La sélection d'un modèle parmi l'ensemble des modèles proposés se fait à l'aide de quelques critères d'information qui permettent de ce fait de réduire considérablement le champ de recherche.

2.3.1 Critère d'Akaike (Akaike information criterion) :

Le critère d'information d'Akaike (1969) est utilisé assez fréquemment et, ce que ne gâche rien, cet indicateur est polyvalent. Grâce à lui, on évalue la bonne adéquation d'un modèle et surtout on peut comparer plusieurs modèles entre eux.

Dans le cas d'une estimation par la méthode des moindres carrés, ce critère mesurera l'écart entre les résidus et distribution gaussienne, par la sélection des ordres p et q , qui minimise la fonction suivante :

$$AIC = \log(\sigma^2) + \frac{2(p+q)}{N}$$

2.3.2 Critère de SCHWARZ (BAYESIAN information criterion) :

Le critère d'information bayésien est un critère d'information dérivé du critère d'information d'Akaike proposé par Gideon Schwarz en 1978. La différence entre ces deux critères réside par le fait que les ordres p et q estimés par ce dernier convergent sûrement vers les vrais ordres.

En effet :

$$SC = \log(\sigma^2) + (p+q) \frac{\log N}{N}$$

2.3.3 Critère de Hannan-Quin (1979) :

$$HQ(p, q) = \log(\sigma^2) + (p+q)c \log\left[\frac{\log n}{n}\right]$$

où c est une constante à spécifier supérieure à 2.

Ce critère l'avantage de plus pénaliser les modèles où les paramètres sont en surnombre comparativement à l'AIC.

Remarque :

D'autres critères existent, les « critères de pouvoir prédictif », Ils sont utilisés pour vérifier les performances des prévisions.

2.4 Estimation du modèle :

Il s'agit, dans cette étape, d'estimer les paramètres des modèles ARMA sélectionnés à l'étape d'identification. Les méthodes d'estimation diffèrent selon le type du processus diagnostiqué.

Dans le cas d'un modèle AR, nous pouvons appliquer une méthode des moindres carrés ou bien nous pouvons utiliser les relations existantes entre l'autocorrélation et les coefficients des modèles (équations de Yule-Walker).

Par contre l'estimation des processus MA ou ARMA repose généralement sur des techniques d'estimation fondées sur la maximisation d'une fonction de vraisemblance.

2.5 Validation du modèle :

Quand on finit d'estimer les paramètres du modèle, on examine les résultats de cette estimation. Cette étape est crucial car elle procède à la certification de ses paramètres pour se prononcer sur leur aptitude réelle à être utilisés pour établir d'éventuelles prévisions.

Il existe deux types de tests pour valider un modèle :

2.5.1 Test sur les paramètres :

La première étape de la validation consiste à vérifier les coefficients du modèle qui doivent être significativement différents de zéro. Si l'un des coefficients n'est pas significativement différent de zéro, il convient d'envisager une nouvelle spécification éliminant l'ordre du modèle AR et MA non valide.

Pour cela on calcule la statistique de Student donnée par : $t_{\phi_p} = \frac{\phi_p}{\sigma_{\phi_p}}$

On a :

$$H_0 = ARMA(p-1, q) \text{ vs } H_1 = ARMA(p, q)$$

Si $|t_{\phi_p}| < t_{(1-\frac{\alpha}{2})}$ (valeur tabulée de Student) : on accepte l'hypothèse H_0 sinon on rejette et on retient le modèle ARMA(p,q).

2.5.2 Test sur les résidus :

Le principe de ce test est de vérifier l'hypothèse d'autocorrélation des perturbations aléatoires ε_t (les résidus entre les valeurs observées et les valeurs estimées doivent se comporter comme un bruit blanc). Dans le cas où la moyenne n'est pas nulle $E(\varepsilon_t) \neq 0$, il convient d'ajouter une constante au modèle.

a. Tests de Box-Pierce et Ljung-Box :

Le test de Box-Pierce « portemanteau » permet d'identifier les processus sans mémoire (suite de variables aléatoires indépendantes entre elles). Nous devons donc identifier :

$$\text{cov}(X_t, X_{t+h}) = 0 \quad \text{ou encore} \quad \rho_h = 0 \quad \forall h$$

Soit les hypothèses : $\begin{cases} H_0 : \rho_h = 0 \quad \forall h \\ H_1 : \text{« Il existe au moins } \rho_i \text{ significativement différent de 0 »} \end{cases}$

Pour effectuer ce test, on a recours à la statistique Q qui est donnée par :

$$Q = n \sum_{h=1}^k \hat{\rho}_h^2$$

k : Nombre de retards

n : Nombre d'observations

$\hat{\rho}_h^2$: Autocorrélation empirique d'ordre h

Q : suit asymptotiquement une loi χ^2 de k degrés de liberté

Si la statistique Q est supérieure au χ^2 dans la table au seuil $(1 - \alpha)$ et k degrés de liberté, les résidus ne forment pas un bruit blanc, la spécification du modèle est donc incomplète et il manque au

moins un ordre au processus.

Nous pouvons également utiliser une autre statistique, dont les propriétés asymptotiques sont meilleures, dérivée de la première qui est le Q' de Ljung-Box :

$$Q' = n(n+2) \sum_{h=1}^K \frac{\hat{\rho}_h^2}{n-h}$$

Qui est aussi distribuée selon à k degrés de liberté et dont les règles de décisions sont identiques au précédent.

b. Test Durbin-Watson :

Les modèles ajustés à des séries chronologiques manifestent parfois un certain degré de corrélation entre les valeurs successives des erreurs. En termes probabilistes, cela signifie que les erreurs sont auto-corrélées, ou encore qu'une erreur produite en $t-1$ a une influence sur l'erreur en t . Le test de Durbin et Watson (1951) permet de détecter l'auto-corrélation des résidus pour un ordre un (corrélation entre ε_t et ε_{t-1}) sous la forme :

$$\varepsilon_t = \rho \varepsilon_{t-1} + V_t \text{ o } V_t \rightarrow N(0; \sigma_v^2)$$

Ce test permet de tester l'auto-corrélation d'ordre 1 sous l'hypothèse que les résidus sont Gaussiens. Donc il teste l'hypothèse nulle :

$$\begin{cases} H_0 : \rho = 0 & \text{« Absence d'auto-corrélation à l'ordre 1 des résidus »} \\ H_1 : \rho \neq 0 & \text{« Présence d'autocorrélation à l'ordre 1 des résidus »} \end{cases}$$

La statistique de Durbin Watson est obtenue après estimation du modèle par les méthodes des moindres carrés ordinaires, notée DW, est donnée par :

$$DW = \frac{\sum_{t=2}^n (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^n \varepsilon_t^2}$$

supposons que $t \rightarrow \infty$; on aura : $\sum_{t=2}^n (\varepsilon_{t-1})^2 = \sum_{t=2}^n \varepsilon_t^2$

Et par conséquent, en développant l'expression (1), nous obtenons :

$$DW = 2 - 2\hat{\varepsilon}$$

$$\text{Si } \hat{\varepsilon} = \begin{cases} 1 & DW = 0 & \text{« il existe une auto corrélation positive »} \\ -1 & DW = 4 & \text{« il existe une auto corrélation négative »} \\ 0 & DW = 2 & \text{« indique l'absence d'auto-corrélation »} \end{cases}$$

c. Test de normalité :

Pour vérifier si le processus des résidus $\{\varepsilon_t, t \in \mathbb{Z}\}$ forme un bruit blanc gaussien, plusieurs tests peuvent être utilisés, mais Le test le plus fréquent qui permet de vérifier la normalité d'une distribution statistique est celui de Jarque et Berra (1984).

Ce dernier est fondé sur la notion de Skewness (moment d'ordre 3 et asymétrie) et Kurtosis (moment d'ordre 4, aplatissement). Ces deux tests permettront par la suite le calcul des intervalles de confiance prévisionnels et aussi d'effectuer les tests de Student sur les paramètres.

Les tests de Skewness et de Kurtosis :

Soit μ_k le moment empirique d'ordre k du processus $\{\varepsilon_t, t \in \mathbb{Z}\}$:

$$\mu_k = \frac{1}{n} \sum_{t=1}^n (\hat{\varepsilon}_t - \bar{\varepsilon})^k$$

Le coefficient de Skewness est défini par : $S_k^{1/2} = \frac{\mu_3}{\mu_2^{3/2}}$

Le coefficient de Kurtosis est défini par : $K_u = \frac{\mu_4}{\mu_2^2}$

Si la distribution est normale et le nombre d'observations grand ($n > 30$) alors :

$$S_k^{1/2} \rightarrow N(0; \sqrt{\frac{6}{n}}) \quad K_u \rightarrow N(3; \sqrt{\frac{24}{n}})$$

On construit alors les statistiques centrées réduites correspondantes à $S_k^{1/2}$ et K_u

$$v_1 = \frac{|S_k^{1/2} - 0|}{\sqrt{\frac{6}{n}}} \quad \text{et} \quad v_2 = \frac{|K_u - 3|}{\sqrt{\frac{24}{n}}}$$

que l'on compare à 1.96 (valeur de la loi normale au seuil de 5%)

Soient les hypothèses :

H_0 : $v_1 = 0$ (asymétrie) et $v_2 = 0$ (aplatissement normal), dans le cas contraire, l'hypothèse de normalité est rejetée.

Test de Jaque-Berra :

Il s'agit d'un test qui regroupe les résultats précédents. On construit la statistique :

$$JB = \frac{n}{6} S_k + \frac{n}{24} (K_u - 3)^2$$

suit une loi χ^2 à deux degrés de liberté.

Donc si $JB > \chi_{1-\alpha}^2(2)$, on rejette l'hypothèse de normalité des résidus au seuil de $\alpha(0.05)$

2.6 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

$$\hat{X}_{t+h} = E(X_{t+h} / X_t, X_{t-1}, \dots, X_1)$$

2.6.1 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 $\widehat{X}_{t+h} = E(X_{t+h}/X_t, X_{t-1}, \dots, X_1)$

Donc $\widehat{X}_{t+h} = \phi_1 X_t + \dots + \phi_p X_{t+1-p}$

De façon analogue $X_{t+h} = \phi_1 X_{t+h-1} + \dots + \phi_p X_{t+h-p} + \varepsilon_{t+h}$

Donc $\widehat{X}_{t+h} = E(X_{t+h}/X_t, X_{t-1}, \dots, X_1)$

De façon récursive

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

2.6.2 Prédiction d'un modèle MA(q) :

Soit le modèle MA(q) $X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} = \Theta(L)\varepsilon_t$

La prédiction optimale à la date $t + 1$ est $\widehat{X}_{t+1} = \theta_1 \varepsilon_t + \dots + \theta_q \varepsilon_{t+1-q}$

De façon analogue

$$\widehat{X}_{t+h} = \begin{cases} \theta_h \widehat{X}_t + \dots + \theta_q \widehat{X}_{t+h-q} & \text{pour } h \leq q \\ 0 & \text{pour } h > q \end{cases}$$

2.6.3 Prédiction 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

$$\widehat{X}_{t+h} = \sum_{i=1}^p \phi_i \widehat{X}_{t+h-i} + \sum_{j=1}^q \theta_j \widehat{\varepsilon}_{t+h-j}$$

Où

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

3.1 Analyse et exploratoire des données

Nous nous intéressons à appliquer la méthode de Box-Jenkins sur des données réelles notées "mouvement sable" de la Cimenterie MEFTAH (scmi), les données considérées sont mensuelles et la période retenue pour l'étude est de janvier 2014 à décembre 2021, notée ApproBrut, vue la volatilité de la série d'origine nous avons construit une nouvelle série notée ApproPreTrai.

Avant de construire le modèle prévisionnel, nous commençons par l'analyse graphique des séries puis l'analyse de la stationnarité à l'aide d'un ensemble de tests.

3.1.1 Analyse de la série ApproBrute

Etude de la stationnarité

Représentation graphique :

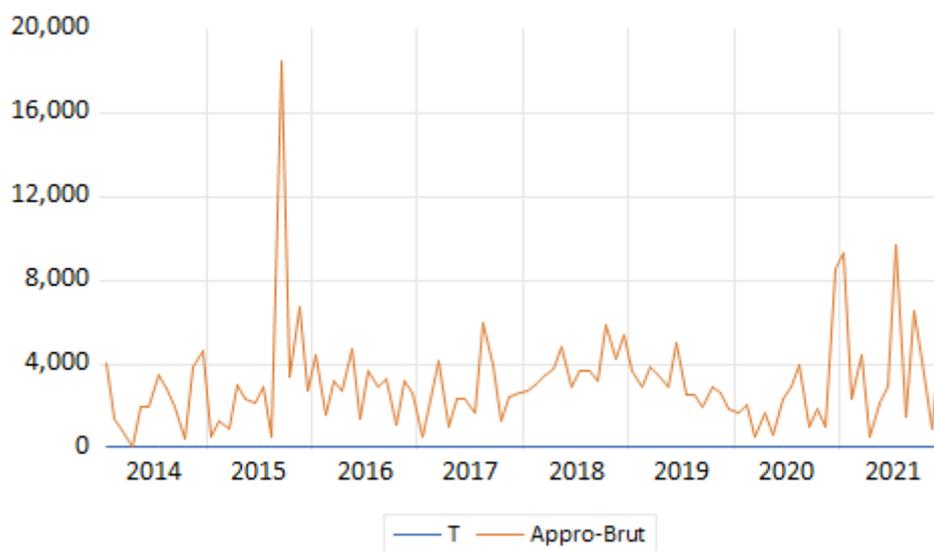


FIGURE 3.1 – Représentation graphique de la série brute

Le corrélogramme de la série ApproBrute montre que la série est stationnaire.

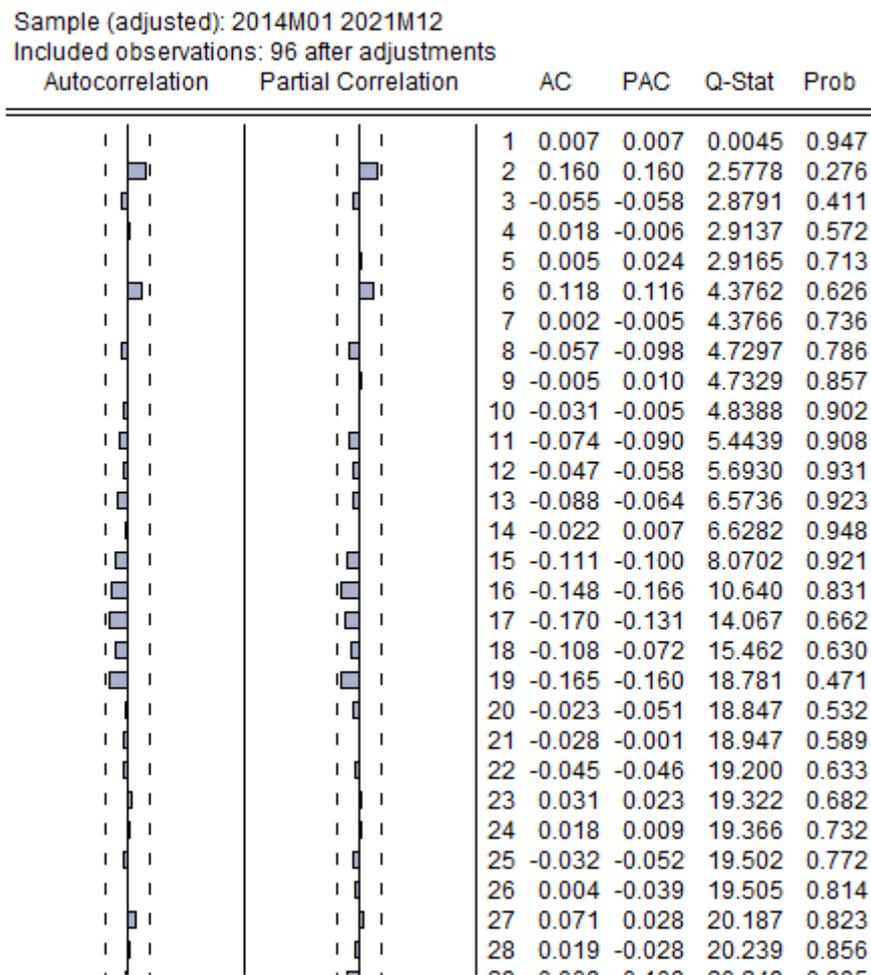


FIGURE 3.2 – Corrélogramme de la série Appro Brute

Pour vérifier, on utilise le test de racine unitaire (test ADF) pour confirmer la stationnarité :

Test de racine unitaire (Dickey-Fuller) :

Pour vérifier l'existence de la tendance et déterminer son type on applique le test de la racine unitaire (Dickey-Fuller) qui nécessite tout d'abord la sélection du nombre de retards p , de sorte à minimiser les critères d'informations (AIC et BIC) et maximise R-squared (pour notre cas $p = 2$), puis on estime le modèle [3] :

Modèle [3] : test sur la tendance

$$\Delta \text{ApproBrute}_t = \phi \text{ApproBrute}_{t-1} + C + \beta + \sum_{j=1}^2 \phi_j \text{ApproBrute}_{t-j} + \varepsilon_t$$

avec ε_t processus stationnaire

Le principe consiste à tester la nullité du coefficient de la tendance β .

On voit que la valeur empirique de la statistique de Student (t-statistic) relative à la tendance (@TREND) qui est égale à (0.715369) est inférieure aux valeurs tabulées (3.53, **2.79**, 2.38 resp. aux seuils 1%, **5%**, 10%). Donc on accepte l'hypothèse nulle. La série n'est pas affectée de tendance déterministe. Le modèle [3] est remis en cause, il faut recommencer ce test à partir du modèle [2]

Null Hypothesis: APPROB has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 2 (Fixed)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-5.217531	0.0002
Test critical values:		
1% level	-4.059734	
5% level	-3.458856	
10% level	-3.155470	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(APPROB)
Method: Least Squares
Date: 06/21/22 Time: 23:13
Sample (adjusted): 2014M04 2021M12
Included observations: 93 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
APPROB(-1)	-0.904942	0.173443	-5.217531	0.0000
D(APPROB(-1))	-0.091476	0.149472	-0.611992	0.5421
D(APPROB(-2))	0.067093	0.106432	0.630384	0.5301
C	2503.654	692.9446	3.613065	0.0005
@TREND("2014M01")	6.818921	9.532027	0.715369	0.4763

FIGURE 3.3 – Test sur la tendance ApproBrute

Modèle [2] : test sur la constante

$$\Delta ApproBrute_t = \phi ApproBrute_{t-1} + C + \sum_{j=1}^2 \phi_j ApproBrute_{t-j} + \varepsilon_t$$

avec ε_t processus stationnaire

On teste la nullité de la constante C pour infirmer ou confirmer la présence de la constante.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(APPROB)
Method: Least Squares
Date: 06/25/22 Time: 21:35
Sample (adjusted): 2014M04 2021M12
Included observations: 93 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
APPROB(-1)	-0.882892	0.170213	-5.186986	0.0000
D(APPROB(-1))	-0.107759	0.147323	-0.731446	0.4664
D(APPROB(-2))	0.059962	0.105673	0.567429	0.5719
C	2769.586	583.1847	4.749072	0.0000

FIGURE 3.4 – Test sur la constante ApproBrute

On remarque que la valeur empirique de la statistique de Student (t-statistic) relative à la constante (C) égale à (4.749072) est supérieure aux valeurs critiques tabulées (3.22, **2.54**, 2.17 resp. aux seuils de 1%, **5%**, 10%), alors on accepte l'hypothèse H_1 la constante est significative, notre série est affecte par une constant, on le confirme par la probabilité (0.0000) inférieure à 0.05, donc notre série est stationnaire.

3.1.2 Analyse de la série Approvisionnement Pré Traitement

Etude de la stationnarité

Représentation graphique :

Ce graphe représentant la série de ApproBrute après lissage

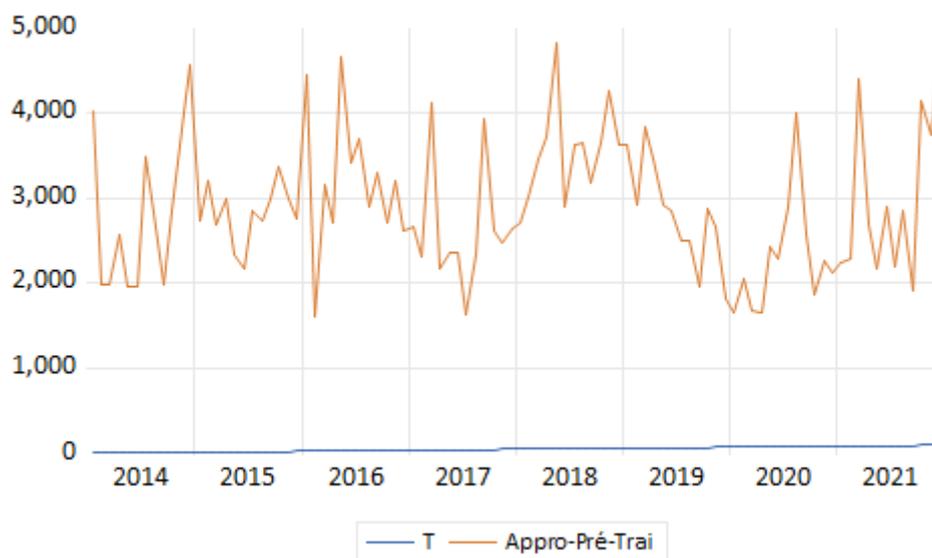


FIGURE 3.5 – Représentation graphique de la série Approvisionnement Pré Traitement

Le graphe de la série ApproPreTrai montre que la série est stationnaire.

Sample (adjusted): 2014M01 2021M12
Included observations: 96 after adjustments

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.269	0.269	7.1923	0.007	
2	0.238	0.178	12.853	0.002	
3	0.104	0.004	13.944	0.003	
4	0.152	0.095	16.294	0.003	
5	-0.014	-0.096	16.315	0.006	
6	0.125	0.116	17.945	0.006	
7	0.111	0.081	19.244	0.007	
8	0.098	0.009	20.279	0.009	
9	-0.020	-0.078	20.321	0.016	
10	-0.003	-0.042	20.323	0.026	
11	-0.099	-0.090	21.411	0.029	
12	-0.190	-0.168	25.467	0.013	
13	-0.173	-0.073	28.851	0.007	
14	-0.085	0.011	29.673	0.008	
15	-0.131	-0.058	31.679	0.007	
16	-0.099	-0.015	32.840	0.008	
17	-0.051	0.032	33.145	0.011	
18	-0.133	-0.085	35.271	0.009	
19	-0.211	-0.108	40.735	0.003	
20	-0.139	-0.006	43.139	0.002	
21	-0.147	-0.062	45.866	0.001	
22	-0.139	-0.051	48.310	0.001	
23	-0.142	-0.081	50.920	0.001	
24	0.017	0.064	50.956	0.001	
25	-0.103	-0.097	52.365	0.001	
26	-0.039	0.007	52.565	0.002	
27	-0.056	-0.015	52.992	0.002	

FIGURE 3.6 – Corrélogramme de la série Appro-Pre-Trai

On voit que le corrélogramme simple décroît ce qui caractérise les séries n'est pas affectées de tendance. Le premier terme des autocorrélations partiels est très élevé par rapport aux autres termes, il est significativement différent de zéro (7.1923), ce qui veut dire qu'il y a une grande influence de

ApproPreTrai($t - 1$) sur ApproPreTrai (t). Pour la saisonnalité, encore une fois elle n'est pas visible sur le corrélogramme partiel. Donc l'évocation d'autres techniques d'analyse des séries temporelles plus appropriées à l'identification de ces caractéristiques non stationnaires s'impose, à savoir le test de Fisher et de Dickey-Fuller. Pour infirmer ou confirmer l'existence de la tendance et/ou de la saisonnalité.

Test de racine unitaire (Dickey-Fuller) :

Pour confirmer l'existence de la tendance et déterminer son type on applique le test de la racine unitaire (Dickey-Fuller) qui nécessite tout d'abord la sélection du nombre de retards p , de sorte à minimiser les critères d'informations (AIC et BIC) et maximise R-squared (pour notre cas $p = 2$), puis on estime le modèle [3] :

Modèle [3] : test sur la tendance

$$\Delta \text{ApproPreTrai}_t = \phi \text{ApproPreTrai}_{t-1} + C + \beta + \sum_{j=1}^2 \phi_j \text{ApproPreTrai}_{t-j} + \varepsilon_t$$

avec ε_t processus stationnaire

Le principe consiste à tester la nullité du coefficient de la tendance β .

Null Hypothesis: APPROPRETRAI has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 2 (Fixed)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.694909	0.0275
Test critical values:		
1% level	-4.059734	
5% level	-3.458856	
10% level	-3.155470	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(APPROPRETRAI)
Method: Least Squares
Date: 06/23/22 Time: 14:55
Sample (adjusted): 2014M04 2021M12
Included observations: 93 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
APPROPRETRAI(-1)	-0.540863	0.146381	-3.694909	0.0004
D(APPROPRETRAI(-1))	-0.220264	0.137937	-1.596850	0.1139
D(APPROPRETRAI(-2))	-0.009256	0.107897	-0.085785	0.9318
C	1616.742	468.8896	3.448022	0.0009
@TREND("2014M01")	-0.549979	2.922211	-0.188206	0.8511

FIGURE 3.7 – Test sur la tendance

On voit que la valeur empirique de la statistique de Student (t-statistic) relative à la tendance (@TREND) qui est égale à (-0.188206) est inférieure aux valeurs tabulées (3.53, 2.79, 2.38 resp aux seuils 1%, 5%, 10%). Donc on accepte l'hypothèse nulle. La série n'est pas affectée de tendance déterministe. Le modèle [3] est remis en cause, il faut recommencer ce test à partir du modèle [2]

Modèle [2] : test sur la constante

$$\Delta \text{ApproPreTrai}_t = \phi \text{ApproPreTrai}_{t-1} + C + \sum_{j=1}^2 \phi_j \text{ApproPreTrai}_{t-j} + \varepsilon_t$$

avec ε_t processus stationnaire

On teste la nullité de la constante C pour infirmer ou confirmer la présence de la constante.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
APPROPRETRAI(-1)	-0.536775	0.143974	-3.728290	0.0003
D(APPROPRETRAI(-1))	-0.223478	0.136132	-1.641627	0.1042
D(APPROPRETRAI(-2))	-0.011690	0.106537	-0.109730	0.9129
C	1578.101	419.2612	3.764005	0.0003

FIGURE 3.8 – Test sur la constante

On remarque que la valeur empirique de la statistique de Student (t-statistic) relative à la constante (C) égale à (3.764005) est supérieure aux valeurs critiques tabulées (3.22, 2.54, 2.17 resp. aux seuils de 1%, 5%, 10%), alors on accepte l'hypothèse H_1 la constante est significative, notre série est affecte par une constant, on le confirme par la probabilité (0.0003) inférieure à 0.05, donc notre série est stationnaire.

3.2 Identification du modèle

3.2.1 Identification de la série ApproBrute

La série ApproBrute est stationnaire, on cherche un modèle ARMA(p , q) qui représente cette série. Pour identifier le processus qui représente au mieux notre série, on examine les autocorrélation simples (pour déterminer le nombres de retards $q = 1, 2, 3, \dots$ de MA) et partielles (pour déterminer le nombres de retard $p = 1, 2, 3, 4, \dots$ d'un processus AR).

L'examen de ce corrélogramme (figure 3.2) montre qu'il existe des pics important pour le terme AR et le terme MA. Alors on a estimé deux modèles candidats à représenter la série : ARMA(1, 1), ARMA(2, 2). Un seul doit être choisi.

Modèle	AIC	BIC	R-squared
ARMA(1,1)	2,5326	2,63945	0,021671
ARMA(2,2)	2,465472	2,625744	0,138246

FIGURE 3.9 – Critères d'information pour les deux modèles

Sachant que le meilleur modèle est celui qui maximise le R-squared et minimise le AIC et le BIC alors le modèle le plus adéquat est le ARMA(2, 2).

3.2.2 Identification du la série ApproPreTrai

La série ApproPreTrai est stationnaire, on cherche un modèle ARMA(p , q) qui représente cette série. Pour identifier le processus qui représente au mieux notre série, on examine les autocorrélation simples (pour déterminer le nombres de retards $q = 1, 2, 3, \dots$ de MA) et partielles (pour déterminer le nombres de retard $p = 1, 2, 3, 4, \dots$ d'un processus AR).

L'examen de ce corrélogramme (figure 3.6) montre qu'il existe des pics important pour le terme AR et le terme MA. Alors on a estimé deux modèles candidats à représenter la série : ARMA(1, 0), ARMA(2, 0), ARMA(1, 1), ARMA(0, 1), ARMA(0, 2). Un seul doit être choisi.

Sachant que le meilleur modèle est celui qui maximise le R-squared et minimise le AIC et le BIC alors le modèle le plus adéquat est le ARMA(2, 0).

Modèle	AIC	BIC	R-squared
ARMA(1,0)	16,15076	16,23089	0,07932
ARMA(2,0)	16,13692	16,24376	0,111414
ARMA(1,1)	16,14007	16,24692	0,108565
ARMA(0,1)	16,1758	16,2593	0,055532
ARMA(0,2)	16,15488	16,26173	0,094735

FIGURE 3.10 – Critères d'information pour les cinq modèles

3.3 Estimation des paramètres

3.3.1 Estimation du la série ApproBrute

Le tableau suivant contient l'estimation du modèle ARMA(2,2) qui a été choisi :

Dependent Variable: LOG(APPROB)
Method: ARMA Maximum Likelihood (BFGS)
Date: 06/25/22 Time: 18:44
Sample: 2014M01 2021M12
Included observations: 96
Convergence achieved after 31 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7.777145	0.119750	64.94481	0.0000
AR(1)	0.304735	0.133458	2.283376	0.0248
AR(2)	-0.718477	0.119899	-5.992344	0.0000
MA(1)	-0.181025	0.066326	-2.729312	0.0076
MA(2)	0.946708	0.072060	13.13781	0.0000
SIGMASQ	0.596841	0.078253	7.627099	0.0000

R-squared	0.138246	Mean dependent var	7.775394
Adjusted R-squared	0.090370	S.D. dependent var	0.836587
S.E. of regression	0.797891	Akaike info criterion	2.465472
Sum squared resid	57.29670	Schwarz criterion	2.625744
Log likelihood	-112.3427	Hannan-Quinn criter.	2.530257
F-statistic	2.887622	Durbin-Watson stat	2.121312
Prob(F-statistic)	0.018289		

Inverted AR Roots	.15+.83i	.15-.83i
Inverted MA Roots	.09+.97i	.09-.97i

FIGURE 3.11 – Modèle ARMA(2,2)

3.3.2 Estimation du la série ApproPreTrai

Le tableau suivant contient l'estimation du modèle ARMA(2,0) qui a été choisi :

Dependent Variable: APPROPRETRAI
Method: ARMA Maximum Likelihood (BFGS)
Date: 06/25/22 Time: 19:26
Sample: 2014M01 2021M12
Included observations: 96
Convergence achieved after 5 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2921.424	142.9862	20.43152	0.0000
AR(1)	0.235785	0.101782	2.316575	0.0227
AR(2)	0.193713	0.114752	1.688098	0.0948
SIGMASQ	547971.1	91508.72	5.988184	0.0000

R-squared	0.111414	Mean dependent var	2898.395
Adjusted R-squared	0.082438	S.D. dependent var	789.4103
S.E. of regression	756.1719	Akaike info criterion	16.13692
Sum squared resid	52605224	Schwarz criterion	16.24376
Log likelihood	-770.5720	Hannan-Quinn criter.	16.18011
F-statistic	3.845077	Durbin-Watson stat	1.938996
Prob(F-statistic)	0.012135		

Inverted AR Roots	.57	-.34
-------------------	-----	------

FIGURE 3.12 – Modèle ARMA(2,0)

3.4 Validation du modèle :

3.4.1 Validation de la série ApproBrute

Tests sur les paramètres :

D'après le tableau d'estimation nous constatons que :

- 1) Tous les coefficients du modèle sont significativement différents de zéros, car la t-statistique correspondante est en valeur absolue supérieure à 1.96, ce qui est confirmé par la probabilités de nullité des coefficients qui sont inférieurs à 0.05
- 2) Les inverses des racines (Invertes AR Roots) fournis par Eviews sont en module inférieur à 1.

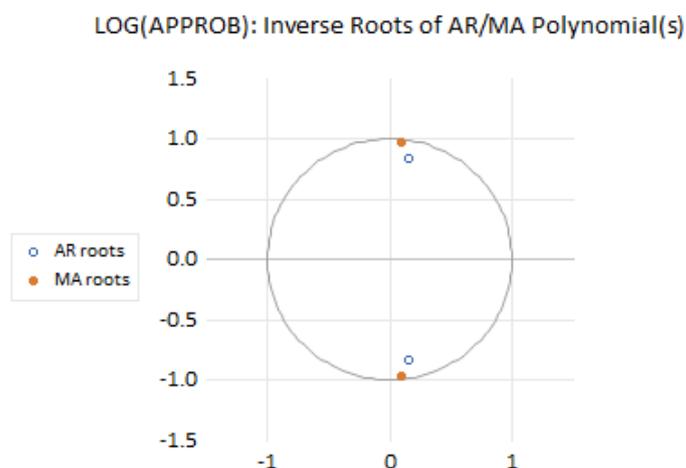


FIGURE 3.13 – Graphe des inverses des racines

Tests sur les résidus : La représentation graphique des séries résiduelles, actuelles et estimées (figure 3.14) montre que le modèle à bien expliqué la série. En effet, le coefficient R-squared de cette estimation est relativement important (égale à 0,138246).

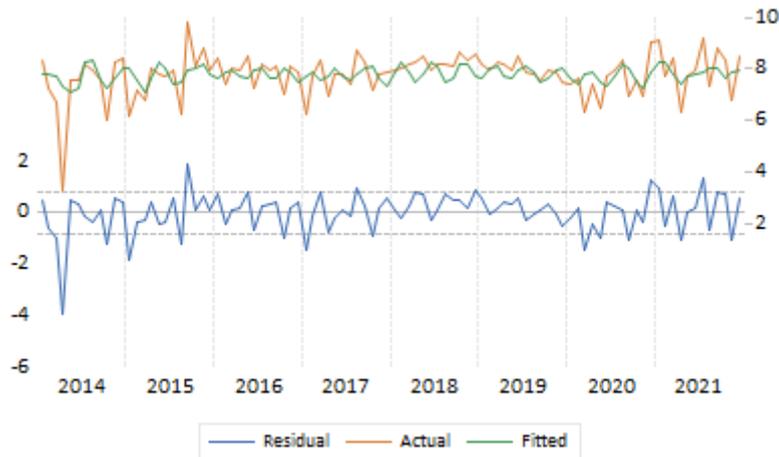


FIGURE 3.14 – Graphique des séries résiduelles, actuelles et estimées

Examen du corrélogramme des résidus :

Il convient maintenant d’analyser les résidus à partir de leurs fonctions d’auocorrélation et d’appliquer des tests tels que :

- 1) Test de bruit blanc (Ljung-Box).
- 2) Test de normalité des résidus (test de Jaque-Berra).
- 3) Test de détection d’autocorrélation d’ordre 1 des erreurs (Durbin-Watson).

Sample (adjusted): 2014M01 2021M12
 Q-statistic probabilities adjusted for 4 ARMA terms

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.065	-0.065	0.4221	
		2 0.001	-0.003	0.4222	
		3 0.036	0.036	0.5505	
		4 0.106	0.112	1.7078	
		5 0.023	0.038	1.7613	0.184
		6 0.163	0.169	4.5310	0.104
		7 0.025	0.043	4.5961	0.204
		8 -0.090	-0.100	5.4542	0.244
		9 0.114	0.085	6.8652	0.231
		10 0.025	-0.002	6.9353	0.327
		11 -0.030	-0.042	7.0344	0.425
		12 -0.053	-0.077	7.3543	0.499
		13 0.026	-0.012	7.4292	0.593
		14 0.011	0.035	7.4439	0.683
		15 -0.136	-0.162	9.5912	0.567
		16 -0.007	-0.033	9.5964	0.651
		17 -0.085	-0.063	10.451	0.657
		18 -0.078	-0.081	11.186	0.671
		19 -0.186	-0.197	15.412	0.422
		20 0.053	0.028	15.766	0.469
		21 -0.065	0.025	16.294	0.503
		22 -0.019	0.023	16.339	0.569

FIGURE 3.15 – Corrélogramme des résidus

Test de Ljung-Box :

Nous remarquons que la statistique Box-Pierce(Q-Stat) est inférieure à la valeur théorique de χ_{h-k}^2 pour tout h , en particulier pour $h = 19, k = 4$;

On a $Q\text{-Stat} = 15.412 < \chi_{15}^2 = 24.9958$ au seuil 5%.

Dans ce cas on peut conclure que les résidus forment un bruit blanc puisque tous les termes sont significativement nuls.

Test de normalité :

Les tests sont effectués à partir des valeurs empiriques des coefficients de Skewness, Kurtosis et la statistique de Jaque-Berra données par le logiciel Eviews 12, ce qui nous donne l'histogramme suivant :

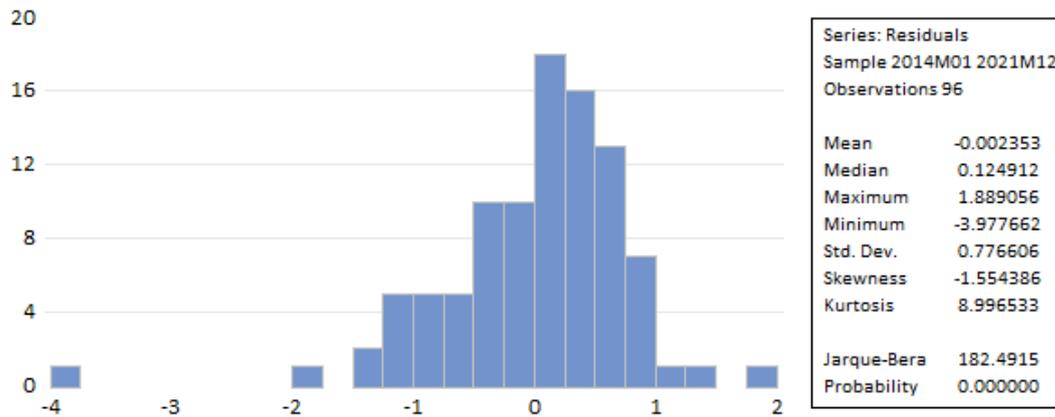


FIGURE 3.16 – Histogramme des résidus

Test de Skewness :

$$(S_k)^{1/2} = -1.554386 \rightarrow V1 = \frac{|(S_k)^{1/2} - 0|}{\sqrt{\frac{6}{96}}} = 6.217544 > 1.96$$

Test de Kurtosis :

$$(K_u) = 8.996533 \rightarrow v_2 = \frac{|k_u - 4|}{\sqrt{\frac{24}{96}}} = 9.993066 > 1.96$$

Ainsi on rejette l'hypothèse de normalité, ce qui est confirmé par la statistique de Jaque-Berra = 16.21061 > 5,9915 ; donc les résidus forment un bruit blanc non gaussien.

Test de Durbin-Watson :

De la statistique de Durbin-Watson on a ($1,78 < DW = 2.121312 < 2,22$), on constate que les résidus ne sont pas corrélés. (Corrélation d'ordre 1)

3.4.2 Validation de la série ApproPreTrai

Tests sur les paramètres :

D'après le tableau d'estimation nous constatons que :

- 1) Tous les coefficients du modèle sont significativement différents de zéros, car la t-statistique correspondante est en valeur absolue supérieure à 1.96, ce qui est confirmé par la probabilité de nullité des coefficients qui sont inférieures à 0.05

2) Les inverses des racines (Invertes AR Roots) fournis par Eviews sont en module inférieur à 1.

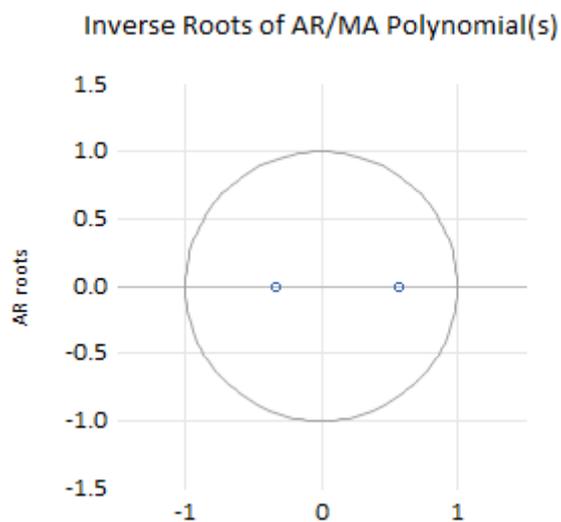


FIGURE 3.17 – Graphe des inverses des racines

Tests sur les résidus :

La représentation graphique des séries résiduelles, actuelles et estimées (figure 3.18) montre que le modèle a bien expliqué la série. En effet, le coefficient R-squared de cette estimation est relativement important (égale à 0,111414).

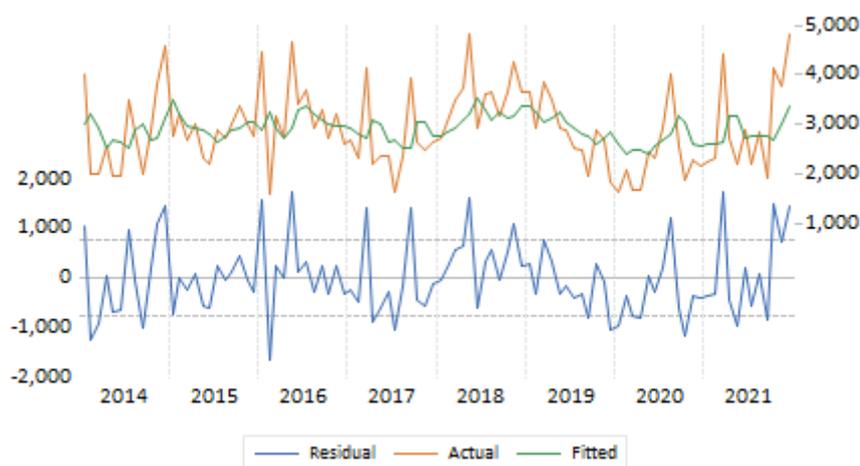


FIGURE 3.18 – Graphique des séries résiduelles, actuelles et estimées

Examen du corrélogramme des résidus :

Il convient maintenant d'analyser les résidus à partir de leurs fonctions d'autocorrélation et d'appliquer des tests tels que :

- 1) Test de bruit blanc (Ljung-Box).
- 2) Test de normalité des résidus (test de Jaque-Berra).
- 3) Test de détection d'autocorrélation d'ordre 1 des erreurs (Durbin-Watson).

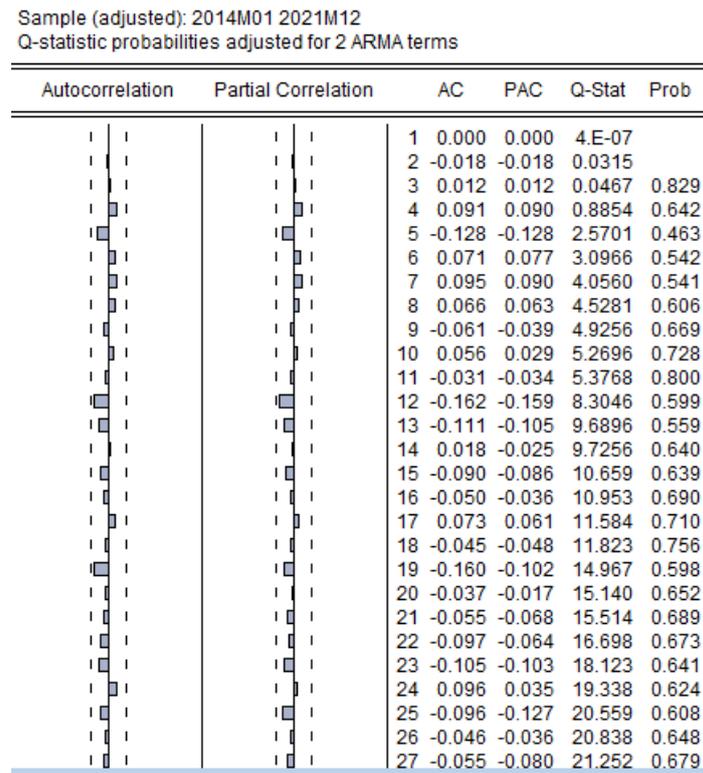


FIGURE 3.19 – Corrélogramme des résidus

Test de Ljung-Box :

Nous remarquons que la statistique Box-Pierce(Q-Stat) est inférieure à la valeur théorique de χ^2_{h-k} pour tout h , en particulier pour $h = 19, k = 2$;
On a $Q\text{-Stat} = 15.412 < \chi^2_{15} = 24.9958$ au seuil 5%.

Dans ce cas on peut conclure que les résidus forment un bruit blanc puisque tous les termes sont significativement nuls.

Test de normalité :

Les tests sont effectués à partir des valeurs empiriques des coefficients de Skewness, Kurtosis et la statistique de Jaque-Berra données par le logiciel Eviews 12, ce qui nous donne l’histogramme suivant :

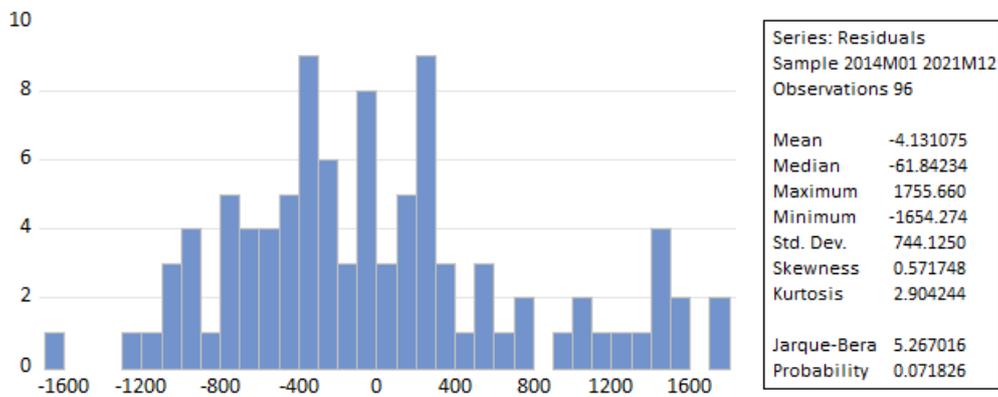


FIGURE 3.20 – Histogramme des résidus

Test de Skewness :

$$(S_k)^{1/2} = 0.571748 \rightarrow V1 = \frac{|(S_k)^{1/2} - 0|}{\sqrt{\frac{6}{96}}} = 2.286992 > 1.96$$

Test de Kurtosis :

$$(K_u) = 2.904244 \rightarrow v_2 = \frac{|k_u - 2|}{\sqrt{\frac{24}{96}}} = 1.808488 > 1.96$$

Ainsi on rejette l'hypothèse de normalité, ce qui est confirmé par la statistique de Jaque-Berra = 4.09548 < 5,9915; donc les résidus forment un bruit blanc gaussien.

Test de Durbin-Watson

De la statistique de Durbin-Watson on a ($1,78 < DW = 1.938996 < 2,22$), on constate que les résidus ne sont pas corrélés. (Corrélation d'ordre 1)

3.5 Prédiction :

3.5.1 Prédiction de la série ApproBrute

Table de prévisions et diagramme séquentiel de la série ApproBrute et de la série prévue. Les prévisions sont calculées pour une durée de cinq (5) mois allant du 12 M 2021 au 05 M 2022

Date	Prédiction
01 M 2022	2385
02 M 2022	2385
03 M 2022	2385
04 M 2022	2385
05 M 2022	2385

FIGURE 3.21 – Tableau de Prédiction

3.5.2 Prédiction de la série ApproPreTrai :

Table de prévisions et diagramme séquentiel de la série brut Approb et de la série prévue. Les prévisions sont calculées pour une durée de cinq (5) mois allant du 12 M 2021 au 05 M 2022

Date	Prédiction
01 M 2022	3500
02 M 2022	3510
03 M 2022	3200
04 M 2022	3300
05 M 2022	3050

FIGURE 3.22 – Tableau de Prédiction

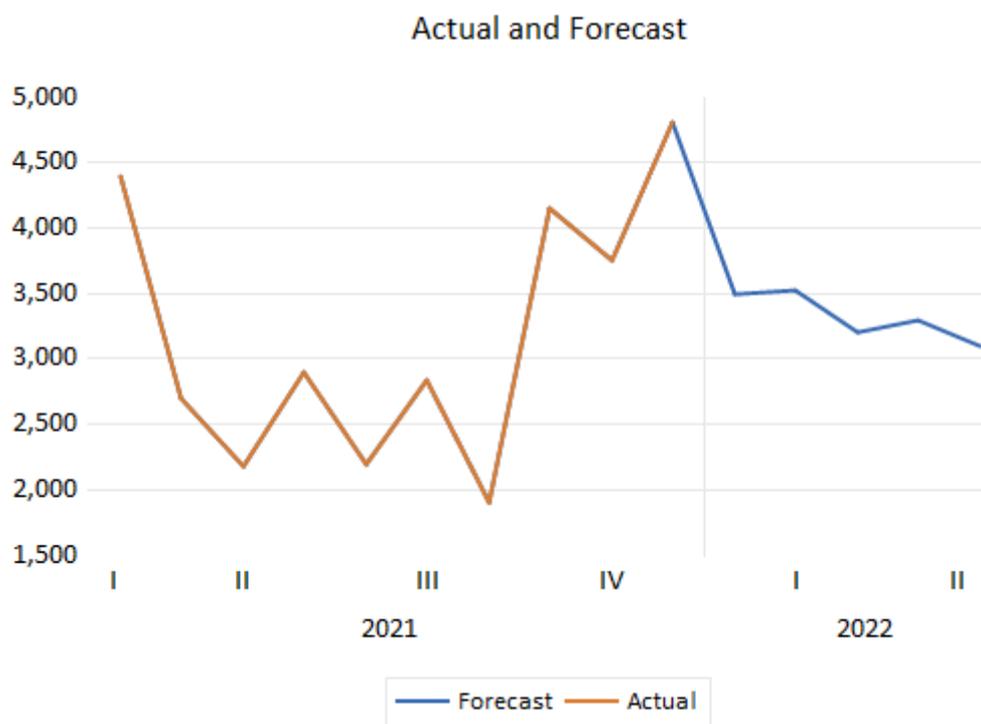


FIGURE 3.23 – Graphe prévision

4.1 Introduction

Le lissage est relativement l'une des méthodes les plus récentes¹ dans le développement de la pratique statistique. Il a été introduit par Holt et Winter en 1957 puis développé d'avantage par Brown en 1962.

La méthode du lissage a profondément marqué les méthodes de prévisions à cours terme.

Elle permet, en effet, le calcul rapide des prévisions pour un grand nombre de chroniques. Le concept de base de la méthode est simple, il apparaît comme une moyenne pondérée entre la dernière prévision et la dernière observation. La formule fondamentale est la suivante :

$$\widehat{X}_t(h) = (1 - \alpha) \sum_{j=1}^{t-1} \alpha^j X_{t-j}$$

$$\widehat{X}_t(h) = \alpha X_{t-1} + \alpha(1 - \alpha)X_{t-2} + \alpha(1 - \alpha)^2 X_{t-3} + \dots + \alpha(1 - \alpha)^j X_{t-j}$$

$\widehat{X}_t(h)$: Prévision à effectuer pour l'instant t

X_{t-j} : Valeur de la série à l'instant $t - j$

h : Horizon de prévision ($h \in \mathbb{N}^*$)

Trois formules de lissage peuvent dériver de celle qui précède.

Lissage classique

Il est donné par la formule suivante :

$$\widehat{X}_t(h) = (1 - \alpha) \sum_{j=0}^{t-1} \alpha^j X_{t-j}$$

Le lissage "exponentiel"

Le lissage exponentiel n'est autre que la somme pondérée de valeurs passées, avec une structure de système de pondération exponentiellement décroissant dans le temps (d'où le nom de la méthode). Il est donné par la formule suivante :

$$\widehat{X}_t = \alpha X_{t-1} + \alpha(1 - \alpha)X_{t-2} + \alpha(1 - \alpha)^2 X_{t-3} + \dots$$

1. G.Mélard p 139

$$\widehat{X}_t = \sum_{i=0}^{+\infty} \alpha(1-\alpha)^i X_{t-i}$$

Lissage à récupération d'erreurs

Il montre que la méthode adapte la nouvelle prévision en fonction d'une fraction de l'erreur de la prévision pour être moderne, cette formule montre que l'algorithme du lissage exponentiel rétro propage la dernière erreur de prévision. Il est donné par la formule suivante :

$$\widehat{X}_t = \widehat{X}_{t-1} + \alpha[X_{t-1} - \widehat{X}_{t-1}] = \widehat{X}_{t-1} + \alpha[\varepsilon(t)]$$

Ces deux dernières formules montrent l'existence d'une dualité entre un modèle autorégressif et modèle adaptatif avec une seule erreur de prévision. « Cette dualité est largement exploitée dans les méthodes récentes »².

Notre étude se restreint uniquement au lissage exponentiel.

4.2 Lissage exponentiel

Choix de la constante du lissage

La méthode du lissage exponentiel débute par la détermination du paramètre α , il convient de choisir une valeur, comprise entre $[0, 1]$, pour ce paramètre. Ce choix est très important car il conditionne la prévision future à travers le degré de pondération que l'on effectue au passé récent et au passé lointain et ce pour assurer une bonne prévision.

Le choix cité précédemment s'effectue selon un certain nombre de critères, citons deux d'entre eux :

Critère1 :

Il est plus intuitif, il vise la minimisation de l'erreur de la prévision moyenne, appelé Erreur absolue moyenne :

$$\sum_t |\varepsilon_t| = \sum_t |X_t - \widehat{X}_t|$$

Critère2 :

Le plus utilisé en pratique, son objectif est de pénaliser les erreurs en les mettant au carré. Il est dit : Erreur quadratique moyenne (MSE en terminologie anglaise), la meilleure valeur de α est celle qui minimise la fonction suivante :

$$\sum_t |\varepsilon_t|^2 = \sum_t (X_t - \widehat{X}_t)^2$$

Remarques

- En général l'influence du paramètre α s'affaiblit lorsque la plage devient assez grande.
- Lorsque α est proche de 0, la pondération s'étale sur un grand nombre de termes du passé, la mémoire du phénomène étudié est forte ou autrement dit profonde et la prévision est peu réactive aux dernières observations ;

2. S.Thiria p 106

- Lorsque α est proche de 1, les observations les plus récentes ont un poids prépondérant sur les termes anciens, la mémoire du phénomène est faible et le lissage est très réactif aux dernières observations.

Selon la nature de la chronique le lissage exponentiel se répartit en trois types :

4.2.1 Le lissage exponentiel simple

Ce type de lissage s'applique uniquement pour les séries qui ne sont affectées ni de tendance ni de composantes saisonnières.

La formule du lissage est donnée par :

$$\widehat{X}_t = \alpha_t + (1 - \alpha)\widehat{X}_{t-1} \quad 0 \leq \alpha \leq 1$$

Initialisation(pour $t = 1$) :

$$\widehat{X}_t = X_t$$

D'autres initialisations existent, la plus intéressante³ est la moyenne \bar{X} de la série.

Quand au choix de α , remarquons les deux cas particuliers suivants :

- Si $\alpha = 0$, alors toutes les prévisions sont identiques.
- Si $\alpha = 1$, la prévision est égale à la dernière valeur observée.

Remarque

Le lissage exponentiel simple est adapté seulement au cas où la série peut être ajusté par une droite horizontale au voisinage de t , dans le cas contraire on fait appel au lissage exponentiel double.

4.2.2 le lissage double

Cette méthode n'est autre que la généralisation de la méthode précédente, elle s'applique pour toutes les séries qui présentent une tendance déterministe et qui s'ajustent par une droite quelconque au voisinage de t

$$(Y_s = \alpha_1 + (s - t)\alpha_2)$$

Cela induit la prévision donnée par la fonction suivante :

$$\widehat{X}_t(k) = \widehat{\alpha}(t) + k\widehat{\beta}(t)$$

ou k : horizon

$$\text{et } \begin{cases} \widehat{\alpha}(t) = \widehat{\alpha}(t-1)\widehat{\beta}(t) + (1 - \alpha^2)[X_t - \widehat{X}_{t-1}(1)] \\ \widehat{\beta}(t) = \widehat{\beta}(t-1) + (1 - \alpha^2)[X_t - \widehat{X}_{t-1}(1)] \end{cases}$$

4.2.3 Le lissage exponentiel généralisé

Parfois on se trouve face à des séries qui ne s'ajustent pas par des droites au voisinage de t mais par des fonctions beaucoup plus compliquées (particulièrement des fonctions ayant des composantes périodiques pour les séries affectés d'une saisonnalité). La difficulté de cette généralisation était de conserver la simplicité des méthodes de calcul. Ce problème a été résolu par Brown en introduisant la notion de vecteurs de fonctions à matrice de transition fixe. La prévision est donnée par :

$$X_t(k) = f'(k)\widehat{\alpha}(t)$$

ou $f'(t)$ est la transposée de $f(t)$ qui est la matrice de transition fixe, et $\widehat{\alpha}$ est un minimisant une certaine fonction.

3. Car la série converge souvent vers sa moyenne

4.2.4 Le lissage de Holt & Winters

Le lissage exponentiel de Holt & Winters sans saisonnalité La méthode de lissage exponentiel de Holt, encore appelée méthode de Holt & Winters sans saisonnalité (1960), est une méthode plus flexible que la méthode du lissage double dans la mesure où elle fait intervenir deux constantes α et β au lieu d'une constante α , même si le principe est fondé sur la même hypothèse. Elle comporte deux lissages exponentiels simultanés.

- Lissage de la moyenne a avec le coefficient α , $\alpha \in [0, 1]$;
- Lissage de la tendance b avec le coefficient β , $\beta \in [0, 1]$. Les formules sont données par :
- Lissage de la moyenne :

$$\alpha_t = \alpha X_t + (1 - \alpha)(a_{t-1} + b_{t-1})$$

$$a_{t-1} + b_{t-1} : \text{représente la nouvelle moyenne lissée en } t .$$
- Lissage de la tendance :

$$b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1}$$

Avec $\begin{cases} X_t & \text{la valeur observée de la série en } t. \\ a_t & \text{moyenne lissée de la série en } t. \end{cases}$

Initialisation (pour $t = 1$) :

Initialisation de la moyenne $a_1 = X_1$

Initialisation de la tendance $b_1 = 0$

prévision calculée pour un horizon h :

$$\hat{X}_{t+h} = a_t + b_t h$$

Le lissage exponentiel de Holt et Winters avec saisonnalité

La méthode de lissage de Winters, n'est autre que la méthode de Holt généralisée par l'introduction d'une composante saisonnière, de ce fait elle est appelée aussi méthode de Holt Winters avec saisonnalité, cette composante saisonnière peut être introduite de deux façons (additive ou multiplicative), en fait elle consiste en trois lissages exponentiels simultanés.

On définit donc trois paramètres, notés α , β , γ :

α : Coefficient de lissage de la moyenne, $\alpha \in [0, 1]$

β : Coefficient de lissage de la tendance, $\beta \in [0, 1]$

γ : Coefficient de lissage de la saisonnalité, $\gamma \in [0, 1]$.

❖ Modèle de Winters additif :

Soit le modèle
$$X_t = a_t + \beta_t + S_t + \varepsilon_t$$

- Lissage de la moyenne : $a_t = \alpha(X_t - S_{t-p}) + (1 - \alpha)(a_{t-1} + \beta_{t-1})$
- Lissage de la tendance : $\beta_t = \beta(\alpha_t - a_{t-1}) + (1 - \beta)\beta_{t-1}$
- Lissage de saisonnalité : $S_t = \gamma(X_t - a_t) + (1 - \gamma)S_{t-p}$
 Initialisation (pour la 1^{re} première année) :

- Initialisation de la saisonnalité : $S_t = X_t/\bar{X}$ pour $t = 1, \dots, p$
- Initialisation de la moyenne : $a_p = \bar{X}$
- Initialisation de la tendance : $\beta_p = 0$

Prévision pour un horizon h :

Les prévisions pour un horizon h se font selon que h soit supérieur ou inférieur à p qui représente la période

$$\begin{aligned} \text{Si } 1 \leq h \leq p & \quad \text{alors} \quad \widehat{X}_{t+h} = (a_t + h\beta_t) + S_{t-p+h} \\ \text{Si } P+1 \leq h \leq 2p & \quad \text{alors} \quad \widehat{X}_{t+h} = (a_t + h\beta_t) + S_{t-2p+h} \end{aligned}$$

❖ Modèle de Winters multiplicatif :

$$\text{soit le modèle} \quad X_t = (a_t + \beta_t t)S_t + \varepsilon_t$$

- Lissage de la moyenne : $a_t = \alpha(X_t/S_{t-p}) + (1 - \alpha)(a_{t-1} + \beta_{t-1})$
- Lissage de la tendance : $a_{1t} = \beta(a_t - a_{t-1}) + (1 - \beta)\beta_{t-1}$
- lissage de saisonnalité : $S_t = \gamma(X_t/a_t) + (1 - \gamma)S_{t-p}$

On utilise S_{t-p} car S_t n'est pas encore connue, avec :

X_t : valeur observée de la série en t .

p : Périodicité des données ;

a_t : Moyenne lissée de la série en t .

β_t : Pente estimée en t .

Initialisation (pour $t = 1$) :

- Initialisation de la saisonnalité : $S_t = X_t/\bar{X} \quad t = 1, \dots, p$
- Initialisation de la moyenne : $a_t = \bar{X}$
- Initialisation de la tendance : $\beta_p = 0$

Les prévisions pour un horizon h :

$$\begin{aligned} \text{Si } 1 \leq h \leq p & \quad \text{alors} \quad \widehat{X}_{t+h} = (a_t + h\beta_t)S_{t-p+h} \\ \text{Si } P+1 \leq h \leq 2P & \quad \text{alors} \quad \widehat{X}_{t+h} = (a_t + h\beta_t)S_{t-2p+h} \end{aligned}$$

Remarque

En pratique, une fois la tendance et/ou la saisonnalité détecté par le test d'ANOVA et/ou le test de Dickey-Fuller on doit spécifier la manière dont la saisonnalité s'introduit dans le modèle. Pour cela on doit tracer le graphe de la série brute sur laquelle on doit relier les deux plus grands pics par une droite dite supérieure et les deux plus petits creux par une droite dite inférieure, ensuite :

Si les deux droites (inférieure et supérieure) sont parallèles alors on se trouve devant un modèle additif, sinon c'est un modèle multiplicatif.

Le chapitre suivant fera l'objet d'application du lissage exponentiel sur les deux séries dont on dispose.

Introduction

On tient à préciser que ce chapitre vient juste après celui de Box Jenkins, car pour déterminer le type de lissage à appliquer (le lissage simple) on doit connaître les caractéristique stochastiques de chacune des séries à savoir la tendance et la saisonnalité. Et ceci se déduit de l'étude préliminaire de la méthodologie de Box Jenkins. Les résultats sont les suivants :

La série ApproBrute est affectée une saisonnalité mais n'est pas affectée une tendance déterministe. La série ApproPréTrai est affectée une saisonnalité mais n'est pas affectée une tendance déterministe

5.1 Etude de la série ApproBrute

La série ApproBrute est affectée de saisonnalité mais n'est pas affecté une tendance déterministe, donc on effectue la méthode de lissage simple saisonnière pour des prévisions de cinq mois. Les résultats donnés par evIEWS se résument dans le tableau 5.1.

Sample: 2014M01 2021M12		
Included observations: 96		
Method: Holt-Winters Additive Seasonal		
Original Series: APPRO_BRUT		
Forecast Series: APPRO_SM		
<hr/>		
Parameters:	Alpha	0.0000
	Beta	0.0100
	Gamma	0.0000
	Sum of Squared Residuals	4.92E+08
	Root Mean Squared Error	2264.591
<hr/>		

FIGURE 5.1 – Lissage ApproBrute

α , β et γ représentent respectivement :

Le coefficient du lissage de la moyenne ($\alpha = 0.0000$)

Le coefficient du lissage de la tendance ($\beta = 0.0100$)

Le coefficient du lissage de la saisonnalité ($\gamma = 0.0000$)

La racine moyenne des carrés résiduels est donnée par $RMSE = 2200.383$

5.1.1 Les prévisions

La Fig5.2 représente les valeurs prévues pour une date future commençant du

End of Period Levels:	Mean	4214.755
	Trend	21.71903
	Seasonals: 2021M01	370.6108
	2021M02	-892.3215
	2021M03	-365.1535
	2021M04	-1024.733
	2021M05	-333.4153
	2021M06	-467.6806
	2021M07	683.1179
	2021M08	-154.4899
	2021M09	1884.905
	2021M10	-561.6929
	2021M11	-66.83947
	2021M12	927.6928

FIGURE 5.2 – Prévision lissage ApproBrute

5.2 Etude de la série ApproPreTrai

La série ApproPreTrai est affectée de saisonnalité mais n'est pas affecté une tendance déterministe, donc on effectue la méthode de lissage simple saisonnière pour des prévisions de cinq mois. Les résultats donnés par eviews se résument dans le tableau 5.3.

Sample: 2014M01 2021M12		
Included observations: 96		
Method: Holt-Winters Additive Seasonal		
Original Series: APPRO_PRE_TRAI		
Forecast Series: APPRO_SM		
Parameters:	Alpha	0.1300
	Beta	0.0000
	Gamma	0.0000
	Sum of Squared Residuals	1.16E+08
	Root Mean Squared Error	1099.590

FIGURE 5.3 – Lissage ApproPreTrai

α , β et γ représentent respectivement :

Le coefficient du lissage de la moyenne ($\alpha = 0.1300$)

Le coefficient du lissage de la tendance ($\beta = 0.0000$)

Le coefficient du lissage de la saisonnalité ($\gamma = 0.0000$)

La racine moyenne des carrés résiduels est donnée par $RMSE = 1099.590$

5.2.1 Les prévisions

La Fig :5.4 représente les valeurs prévues pour une date future commençant du

End of Period Levels:	Mean	2473.589
	Trend	2.677639
	Seasonals:	
	2021M01	-49.47243
	2021M02	-406.2133
	2021M03	139.9960
	2021M04	-500.5422
	2021M05	209.8170
	2021M06	-172.6770
	2021M07	330.8330
	2021M08	-8.937568
	2021M09	-0.138540
	2021M10	-196.0314
	2021M11	74.60014
	2021M12	578.7663

FIGURE 5.4 – Prévision lissage ApproPreTrai

5.3 Conclusion

Ce qu'on remarque est que la prévision est relativement faible pour $j + 1$, celles de $j + 2, \dots, j + 5$ s'éloignent de la vraie valeur.

CONCLUSION GÉNÉRALE

L'objectif de notre étude est d'utiliser la méthodologie de Box-Jenkins, afin de modéliser un processus stochastique qui génère par la suite les deux séries ApproBrute et ApproPreTrai, et qui permet entre autre d'entrevoir ou de prévoir des valeurs futures des séries en fonction de ses valeurs passé.

La méthode en question est une procédure itérative constituée de quatre étapes : l'identification du modèle, l'estimation du modèle, la validation et la prévision du modèle établie. Parmi ces quatre étapes, le processus d'identification est considéré comme l'étape la plus délicate et la plus complexe où l'analyste rencontre le plus de difficultés. La meilleure façon d'éviter cet obstacle est de comprendre la méthodologie et de l'appliquer à des problèmes du monde réel.

C'est la raison pour laquelle, on juge la pratique comme primordiale afin d'acquérir l'expérience dont l'analyste a besoin pour devenir efficace dans la construction du modèle.

Les deux série ApproBrute, ApproPreTrai, semblent à l'origine non stationnaire, et le test de Dicky Fuller a certifié cette hypothèse. Alors des ajustements, ou différentiations saisonnières, ont été appliqués à ces séries, afin de les rendre stationnaires, car la méthodologie de Box-Jenkins est fondée sur la stationnarité des séries.

Par la suite, le processus d'identification du modèle a généré plusieurs modèles de marche aléatoire pour les deux séries. Une fois le modèle le plus adapté a été choisis, la statistique du test porte-manteau a été utilisée pour analyser l'adéquation globale du modèle estimé.

Après la description de la théorie de la méthode du lissage exponentiel. Nous avons appliqué la méthode du lissage exponentiel. Les deux chapitres restant feront l'objet d'une étude multi-variée afin de détecter les relations existantes entre les différentes séries.

Enfin, le modèle adéquat a été utilisé pour prévoir les valeurs futures de la série chronologique.

L'analyse d'une série temporelle ne constitue que rarement une fin en soi. Le véritable objectif poursuivi est la prévision afin de prendre des décisions sur la stratégie qui sera suivie à l'avenir. Ce mémoire de master, vise à donner un aperçu sur la méthodologie 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 des 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 précises.

Une étude pratique a été réalisée, en appliquant la méthodologie de Box et Jenkins pour modéliser une série temporelle réelle qui représente le mouvement du sable de la cimenterie de Meftah.

En conclusion, le meilleur modèle retenu est ARMA(2,2) et c'est à partir de ce dernier que les prévisions ont été faites pour une période de cinq mois.

- [1] Philippe C Besse and Hervé Cardot. Approximation spline de la prévision d'un processus fonctionnel autorégressif d'ordre 1. *Canadian Journal of Statistics*, 24(4) :467–487, 1996.
- [2] Pierre Bidan. Commande diffusive d'une machine électrique : une introduction. In *ESAIM : Proceedings*, volume 5, pages 55–68. EDP Sciences, 1998.
- [3] Denis Bosq. Processus linéaires vectoriels et prédiction. *Comptes Rendus Mathématique*, 337(2) :115–118, 2003.
- [4] G Carayannis. Nouvelles conceptions pour la résolution récurrente des équations yule-walker dans le cas des modèles arma (auto-régressifs et à moyenne mobile). *Bulletins de l'Académie Royale de Belgique*, 63(1) :363–378, 1977.
- [5] Serge Degerine. *Fonction d'autocorrélation partielle et estimation autorégressive dans le domaine temporel*. PhD thesis, Université Joseph-Fourier-Grenoble I, 1988.
- [6] Guy Melard et al. Méthode d'identification récurrente pour la modélisation de données chronologiques économiques. *Actes du Salon International de l'Informatique de la Télématique*, 1989.
- [7] JE PATUREL and B KOUAME. Caractérisation de fluctuations dans une série chronologique par applications de tests statistiques etude bibliographique. 1994.
- [8] Bernard Picinbono. Positivité d'une covariance de type ma (2) ou ma (3). *Traitement du Signal*, 33(4), 2016.
- [9] S Ranganathan. Binary collisions in fluids : II : velocity auto correlation function. *Canadian journal of physics*, 61(12) :1655–1659, 1983.