



Mémoire de Master

Filière : Mathématiques

Spécialité : Recherche Opérationnelle

Thème

Modélisation et prévision des ventes au niveau de l'EURL
FLOREX EQUIPEMENTS par les modèles ARIMA

Présenté par :

– DAHMANI M'HAMED

Devant le jury composé de :

Président	<i>M^{me}</i> Khadidja BOUDANE	MAA	U. A/M/O Bouira.
Encadreur	<i>M^r</i> L'hadi BOUGHANI	MAA	U. A/M/O Bouira.
Examinatrice	<i>M^{me}</i> Lala Maghnia Alem	MAA	U. A/M/O Bouira.

2019/2020

Remerciements

Je tiens à remercier mon promoteur M.L'HADI BOUGHANI pour avoir accepté de diriger ce travail ainsi que pour ses conseils, sa patience et ses encouragements.

Je tiens à exprimer mes remerciements aux membres du jury qui ont accepté de juger notre travail, en l'occurrence MME. BOUDANE KHADIDJA présidente du jury et MME. ALEM LALA MAGHNIA examinatrice de ce travail.

Mes remerciements s'adressent enfin à tous ceux qui m'ont soutenu de près ou de loin.

Mecri à Tous

Résumé

L'objectif de ce mémoire est l'utilisation des techniques d'analyse et de modélisation des série temporelles (chronologiques) pour la prévision des ventes au niveau de l'entreprises FLOREX EQUIPEMENTS. On s'intéresse essentiellement aux modèles AR,MA,ARMA et ARIMA, On utilisant le logiciel R.

Mots clés : séries chronologiques, modélisation, AR, MA, ARMA, ARIMA, saisonnalité, tendance, résidus, bruit blanc

Abstract

The objective of this dissertation is the use of analytical techniques and time series modeling for forecasting sales at the FLOREX EQUIPEMENTS company. We are mainly interested in AR, MA, ARMA and ARIMA models, Using the R software.

Keywords : time series, modeling, AR, MA, ARMA, ARIMA, seasonality, trend, residuals, white noise

Table des figures

1.1	Grenaille acier angulaire	6
1.2	Unité de sablage 200 litres et casque de sablage	6
2.1	nombre mensuel de passagers aériens(usa)1949-1960	11
2.2	bénéfices trimestriels de Johnson and Johnson Share	11
2.3	la tendance de la série AirPassengers[6]	12
2.4	saisonnalité de la série AirPassengers[6]	12
2.5	la résidu de la série AirPassengers[6]	13
3.1	Tableau des données relatives aux ventes mensuelles 2017–2018 des produits cds, gaa et Us 200	23
3.2	Corrélogramme et corrélogramme partiel de la série Cds	25
3.3	Corrélogramme et corrélogramme partiel des résidus de la série Cds	26
3.4	26
A.1	Graphe de la série Gaa	29
A.2	Corrélogramme et corrélogramme partiel de la série Gaa	29
A.3	Les trois composantes de la série Gaa(saisonnalité, tendance e résiduelle)	29
A.4	Résultats de la commande auto.arima pour la série Gaa	30
A.5	Graphe, ACF et PACF des Résidus Gaa	30
A.6	Résultat du Test de Box et Pierce sur les Résidus Gaa	30
A.7	Prévision des ventes Gaa sur l’horison Janvier 2019-Juin 2019	30
A.8	Graphe de la série Us 200	30
A.9	ACF et PACF de la série Us 200	31
A.10	Les trois composantes de la série Us 200(saisonnalité, tendance e résiduelle)	31
A.11	Réultats de la commande auto.arima pour la série Us 200	32
A.12	Graphe, ACF et PACF des résidus Us 200	32
A.13	Résultat du Test de Box et Pierce sur les Résidus Us 200	32
A.14	Prévision des ventes Us 200 sur l’horison Janvier 2019-Juin 2019	32

Table des matières

Liste des figures	1
Introduction générale	4
1 Présentation de l'entreprise et de la fonction vente	5
1.1 Produit	5
1.2 Partenaires	7
1.3 Assistance	7
1.4 Evenements-nouveautés	7
1.5 Fonction vente	7
1.5.1 Définition de la fonction vente	7
1.5.2 La place de la force de vente dans la stratégie de l'entreprise	8
1.5.3 Objectifs de la fonction vente	8
2 Série temporelles	10
2.1 Série temporelles : Vocabulaire et exemples	10
2.2 Définition	10
2.3 Description d'une série temporelles	12
2.4 Objectifs principaux	13
2.5 Description schématique de l'étude complète d'une série temporelle	14
2.5.1 Correction des données	15
2.5.2 Observation de la série	15
2.5.3 Analyse de série à partir de ses composantes	15
2.5.4 Modélisation	15
2.5.5 Diagnostic du modèle	16
2.5.6 Prévision	16
2.6 Processus stationnaire de second ordre	16
2.6.1 Bruit blanc	17
2.6.2 Polynomes retard, avance et différence	17
2.7 Processus ARMA et ARIMA	18
2.7.1 Processus auto-régressifs d'ordre p (AR(p))	18
2.7.2 Processus moyenne mobile d'ordre q (MA(q))	19
2.7.3 Processus ARMA(p, q)	20
2.7.4 Processus ARIMA(p, d, q)	21
2.7.5 Processus SARIMA	21
3 Application à la prévision des ventes de Florex équipements	22
3.1 Modélisation des séries de données de ventes avec R	22
3.1.1 importation des données à partir d'un fichier Excel	22
3.1.2 Modélisation de la série Cds : Casque de sablage	24

3.1.3	Prévision de la série Cds	27
3.1.4	Modèles des séries Gaa e Us 200	27
3.2	Conclusion	27
	Conclusion générale	28
	A Graphes et figures des Résultats De Modélisation sous R	29
	Bibliographie	33

Introduction Générale

Beaucoup de problèmes de la vie courante implique des études via l'observation de données expérimentales relatives au facteurs régissant le système en quetsion. Ces données se présentent sous différentes formes et suivant chaque forme existe une théorie de traitement.

Les données représentées par l'observation de l'évolution d'un phénomène au cours du temps sont dites chroniques ou séries chronologiques (temporelles); elles sont d'une grande utilité pour les systèmes où l'estimation des valeurs futures de ce phénomène est d'une nécessité absolue (éviter une catastrophe ou sauver des vies).

La théorie des séries temporelles est appliqué de nos jours dans des domaines aussi variés que l'économétrie, la biologie ... etc. On s'intéresse à l'évolution au cours de temps d'un phénomène, dans le but de décrire, expliquer, modéliser, puis prévoir ce phénomène dans le futur.

Cette dernière offre plusieurs outils puissant d'analyse de données expérimentales relevée au cours du temps, telles que les modèles ARIMA, ARCH,... etc.

La prévision de vente est une tache que tout entreprise se doit de bien traiter et réaliser car son élaboration est lier à l'épanouissement de l'entreprise et joue un grand rôle dans la fonction du marketing.

Ce travail présente une étude de modélisation des ventes de Florex équipements, basée sur l'analyse des séries temporelles et un historique de données réelles relatives aux quantités vendues de trois produits : Casque de sablage, Grenaille acier angulaire et Unité de sablage 200 litres pris ur la période janvier 2017 juqu'à decembre 2018. Cette modélisation est réalisée par le logiciel R.

Dans ce mémoire on présente la problématique suivant :

quel serait le modèle théorique présenté par l'analyse des séries temporelles le mieux adapté à décrire l'évolution de la quantité vendue de chaqu'un des trois produits précédents.

La réponse à cette question exige une utilisation d'une certaine packages, fonctions et commandes sur le logiciel R.

Ce project comporte alors trois chapitres.

Le premier chapitre est consacré pour présenté l'entreprise FLOREX EQUIPEMENTS, ces produits, ces partenaires et l'assistance qu'elle offert.

Le deuxième chapitre est réservé a l'étude des séries temporelles, où on introduit quelques concepts de base nécessaires à la compréhension de ses outils, par la suite ses différents modèles MA, AR, ARMA et ARIMA.

Le dernier chapitre est consacré à la présentation des résultats obtenus par le logiciel R lors de la modélisation des séries temporelles relatives aux données réelles de l'évolution de la quantité vendue des trois produits précédents dans le temps.

Chapitre 1

Présentation de l'entreprise et de la fonction vente

Créée en 1998, FLOREX EQUIPEMENTS, Leader algérien dans le domaine des Traitements de Surfaces, est spécialisée dans le Conseil, la Vente, la Formation, la Mise en Service et la Réparation d'Equipements et de Machines destinés aux activités de Traitements de Surfaces, situé a Draria Wilaya d'Alger.

1.1 Produit

L'entreprise Florex Equipements offre une diversité de produits et installations à la disposition de sa clientèle. Elle se fixe l'objectif de satisfaire à la demande du territoire national. Les principaux produits fournis par Florex Equipements sont :

- Les Sableuses.
- Hydrogommeuses.
- Cabines de sablage.
- Grenailleuses à turbines
- Abrasifs de sablage grenailage.
- Ponceuses, surfaceuses et raboteuses pour sols en béton.
- Pistolets de métallisation.
- Pistolet de peinture finition.
- Pistolet pour émaillage.
- Pompes airless de peinture anticorrosion.
- Pompes airless de peinture bâtiment.
- Cabines de peinture.
- Pistolets de poudrage électrostatique.
- Poudres époxy thermodurcissables.
- Cabines de poudrage et fours de cuisson.
- Lignes de poudrage complètes.
- Accrochages et masquages pour poudrage et peinture.
- Appareils de mesure pour traitement de surfaces.
- Appareils de mesure pour laboratoires : pharma, agro, cosmétique, chimie, . . .
- Microbilles de broyage pour fabrication de peinture.
- Microbilles et paillettes décoratives pour peinture bâtiment, cosmétique, . . .
- Traceurs routiers.
- Appareils de contrôle pour traçage routier.

- Pompes haute pression pour nettoyage ou hydrocurage.
- Pompes haute pression pour décapage.
- Pompes haute pression pour démolition.
- Pompes haute pression pour démarquage routier.
- Pompes haute pression pour essais hydrostatiques.
- Pompes haute pression pour découpe à jet d'eau Abrasif.
- garnet pour découpe à jet d'eau.
- Nettoyeurs d'échangeurs / faisceaux routier.
- Extracteurs d'échangeurs / faisceaux routier.
- Pincés de levage d'échangeurs / faisceaux.

Exemples :



FIGURE 1.1 – Grenaille acier angulaire



FIGURE 1.2 – Unité de sablage 200 litres et casque de sablage

1.2 Partenaires

Florex Equipement a beaucoup de partenaires très actifs dans le marché national tel que :

- Peinemann Equipment,
- Labomat instruments et spécialités,
- Zhentner testing instruments . . . etc.

1.3 Assistance

FLOREX EQUIPEMENTS assure la Livraison, le Montage, la Mise en Route et le Service après vente de tous ses équipements, sur l'ensemble du territoire national.

Elle assure également la Réparation de tous ses équipements et dispose de la Pièce de Rechange nécessaire.

Elle dispose de Techniciens Agréés qui sont en mesure d'intervenir rapidement et efficacement, aussi bien en ses ateliers que sur les sites de ses clients.

FLOREX EQUIPEMENTS est en mesure de délivrer une Formation au profit des entreprises souhaitant avoir un personnel certifié et qualifié dans le domaine du Traitement de Surfaces.

1.4 Evenements-nouveautés

FLOREX EQUIPEMENTS a participé à la 21e édition du BATIMATEC du 22 au 26 Avril 2018, qui s'est déroulée comme les années précédentes au Palais de la SAFEX – Pins Maritimes – Wilaya d'Alger. Le BATIMATEC est devenu l'événement incontournable pour le secteur du bâtiment en Algérie, il reste une référence pour les acteurs de ce secteur.

FLOREX EQUIPEMENTS a participé à la 10e édition du Salon ALGER INDUSTRIES qui s'est déroulée du 03 au 07 Octobre 2016 – Palais de la SAFEX – Pins Maritimes- Wilaya d'Alger.

FLOREX EQUIPEMENTS était présente avec son fournisseur PEINEMANN qui s'est tenu à l'Esplanade de l'Hôtel HILTON – Alger – Algérie, du 08 au 11 mars 2016.

FLOREX EQUIPEMENTS était présente du 14 au 17 octobre 2008 à la 2e édition du Salon de l'Industrie qui a eu lieu au Palais de la SAFEX – Pins Maritimes – ALGER Une large gamme de produits destinés aux applications industrielles a été exposée lors de ce salon sur le stand Florex.

1.5 Fonction vente

1.5.1 Définition de la fonction vente

La force de vente d'une entreprise se compose de l'ensemble des personnes qui ont pour mission principale de vendre les produits de l'entreprise au moyen de contacts directes avec les clients potentiels, les distributeurs ou les prescripteurs de ces produits.

Une force de vente est bien plus que la somme de commerciaux qui la composent, elle dispose d'une organisation et d'une structure qui définissent les tâches de chacun des commerciaux.

La fonction des vendeurs ne se limite pas uniquement à la vente des produits ou services, mais elle la dépasse vers l'amélioration de l'image de marque de l'entreprise, la collecte d'information et la fidélisation des clients. D'autres auteurs parlent souvent de représentants et non uniquement de vendeurs, car ces derniers sont des "envoyés spéciaux" chargés de démontrer aux clients et aux concurrents la force et les atouts des entreprises qu'ils représentent[10].

1.5.2 La place de la force de vente dans la stratégie de l'entreprise

L'action d'une force de vente ne peut être isolée, bien au contraire, elle s'inscrit dans un ensemble d'actions commerciales.

Elle adapte le plan marketing en fonction des clients, aide à la préparation de la décision, la participation à la formation, et à la réalisation des objectifs de l'entreprise, alors que le marketing intervient pour accroître les responsabilités de la force de vente, en l'incitant à coopérer à la politique de l'entreprise, à apporter de l'information, ainsi que d'éventuelles idées de nouveaux produits.

La participation de la force de vente au processus d'information, de proposition et d'évaluation des ajustements proposés par le responsable montre son rôle dans la phase d'évaluation d'une stratégie commerciale.

Pour atteindre une productivité intéressante de sa force de vente, l'entreprise doit avant tout avoir une direction stratégique de ses vendeurs, c'est-à-dire qu'elle doit implanter la démarche commerciale (prévision des ventes, affectation des ressources...) dans une vraie démarche stratégique de marketing mixe de l'entreprise (politique de produit, de prix, de distribution et de communication). Elle fait l'offre au client final, ou au client intermédiaire, elle distribue le produit, ou le fait distribué, elle est aussi un agent de communication privilégié puisqu'elle permet le contact direct d'homme à homme.

1.5.3 Objectifs de la fonction vente

La principale tâche pour la force de vente est de vendre. Vendre c'est un art, et posséder son art c'est maîtriser l'ensemble des techniques permettant l'accomplissement d'une oeuvre, vendre requiert des connaissances approfondies comme dans tout autre métier.

Donc l'acte de vente suppose la maîtrise successive de la prospection, de la démonstration, de la réponse aux objections, de la conclusion et du suivi.

Le bon vendeur n'est plus seulement celui qui vend, mais celui qui sait aussi installer une relation durable avec les clients.

Lorsque les vendeurs croient au succès de leurs produits, leur argumentation est plus efficace, ils jouent alors un rôle important dans la représentation directe de l'entreprise.

Pour cela, les objectifs assignés à la force de vente doivent prendre en considération la nature des marchés visés par l'entreprise et le positionnement recherché sur chaque marché. Ces objectifs varient d'une entreprise à une autre et sont fixés par les hommes du marketing; sont généralement ceux de la firme elle même.

Une des taches importantes dans la fonction vente es la gestion de la prévision des ventes en se basant sur l'historiques des données enregistrées quotidiennement, mensuellement . . . etc.

Cette dernière tache affecte directement le bénéfice de l'entreprise, donc la maitrise de celle ci est d'une nécessité absolue pour toute entreprise.

L'entreprise Florex equipments n'échappe pas à cette tache, et dans le chapitre qui suit on présente un outil très puissant pour la prévision de ses ventes et la maitrise de cette fonction. ce

dernier est offert par l'analyse de séries chronologiques.

Chapitre 2

Série temporelles

2.1 Série temporelles : Vocabulaire et exemples

2.2 Définition

La théorie des séries temporelles (ou chronologiques) abordée dans ce projet est appliquée de nos jours dans des domaines aussi variés que l'économétrie, la médecine 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, expliquer puis prévoir ce phénomènes dans le future. On dispose ainsi d'observations à des dates différentes, c'est à dire d'une suite de valeurs numériques indicées par le temps.

Exemple : l'évolution du nombre de voyageurs utilisant le train, le croisement relatif mensuel de l'indice des prix ou encore l'occurrence d'un phénomène naturel (comme le nombre de taches solaires).

Cette suite d'observations d'une famille de variables aléatoires réelles notées $(X_t)_{t \in T}$ est appelée **série temporelle** (ou chronologique). ou l'ensemble T est appelé **espace temps** qui peut être :
- **discret** : Dans ce cas $T \subset \mathbb{Z}$.

Les dates d'observations sont le plus souvent équidistantes : par exemple relevés mensuels, trimestriels. . . .

Ces dates équidistantes sont alors indexées par des entiers : $1, 2, \dots, n$ et n est le nombre d'observations. On dispose donc des observations des variables X_1, X_2, \dots, X_n issues de la famille $(X_t)_{t \in T}$ ou $T \subset \mathbb{Z}$ (le plus souvent $T = \mathbb{Z}$). Ainsi si h est l'intervalle de temps séparant deux observations et t_0 l'instant de la première observations, on a le schéma suivant :

$$\begin{array}{cccc} t_0 & t_0 + h & \dots & t_0 + (n-1)h \\ & \Downarrow & \Downarrow & \Downarrow \\ & & & \\ X_{t_0} & X_{t_0+h} & \dots & X_{t_0+(n-1)h} \\ & \Downarrow & \Downarrow & \Downarrow \\ & & & \\ x_1 & x_2 & \dots & x_n \end{array}$$

- **continu** (signal radio, résultat d'un électrocardiogramme. . .). L'indice de temps est à valeurs

dans un intervalle de \mathbb{R} et on dispose (au moins potentiellement) d'une infinité d'observations issues d'un processus $(X_t)_{t \in T}$ où T est un intervalle de \mathbb{R} . Un tel processus est dit à temps continu. Les méthodes présentées dans ce cadre sont différentes de celles pour les séries chronologiques à temps discret.

Dans ce mémoire, on considère uniquement des séries temporelles $(X_t)_{t \in \mathbb{Z}}$ à **temps discret** et **unidimensionnels** : chaque observation x_t est un réel. On peut également s'intéresser à des séries temporelles multidimensionnelles, c'est à dire telles que X_t soit un vecteur de \mathbb{R}^d . [6]

Exemple 2. 1 On peut étudier l'évolution du nombre de passagers qui empruntent le transport aérien (la série AirPassengers extrait de R data sets) et les bénéfices trimestriels de "Johnson and Johnson Share" (la série JohnsonJohnson de R data sets) :

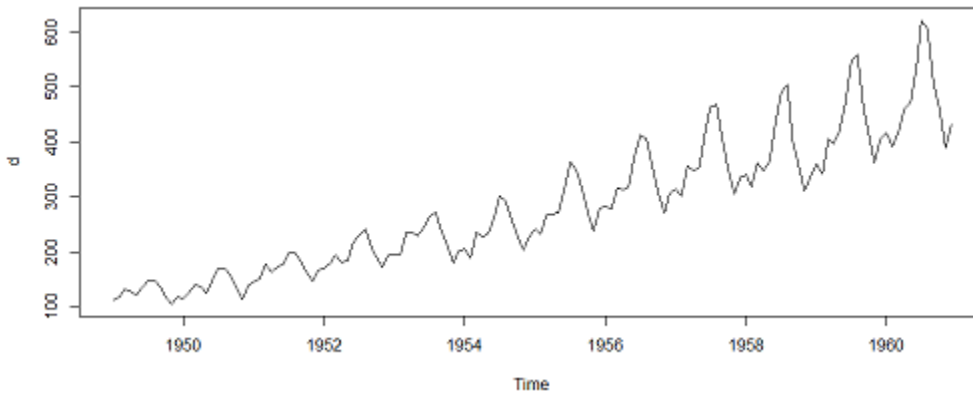


FIGURE 2.1 – nombre mensuel de passagers aériens(usa)1949-1960

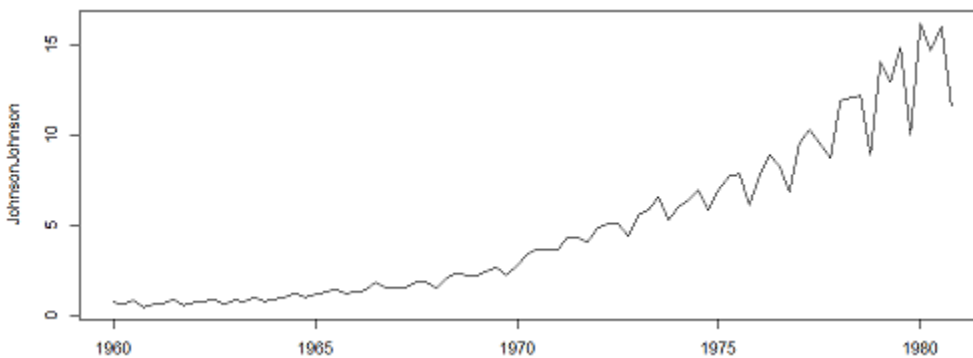


FIGURE 2.2 – bénéfices trimestriels de Johnson and Johnson Share

2.3 Description d'une série temporelles

On considère qu'une série temporelle (X_t) est la résultante de différentes composantes fondamentales :

-La **tendance** (ou trend) (Z_t) représente l'évolution à long terme de la série étudiée. Elle traduit le comportement "moyen" de la série, par exemple la tendance de la série précédente "AirPassengers" est représentée dans la figure suivante [6] :

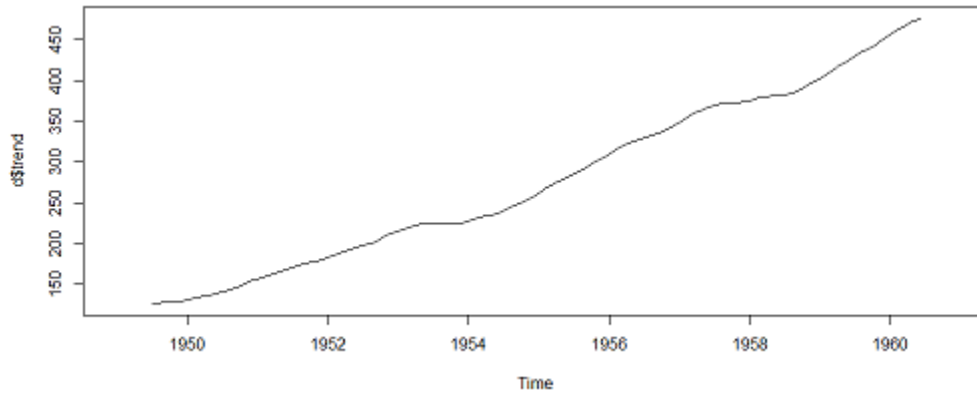


FIGURE 2.3 – la tendance de la série AirPassengers[6]

-La **saisonnalité** (ou composante saisonnière) (S_t) correspond à un phénomène qui se répète à l'intervalle de temps régulier (périodiques). En général, c'est un phénomène saisonnier d'où le terme de variations saisonnières.

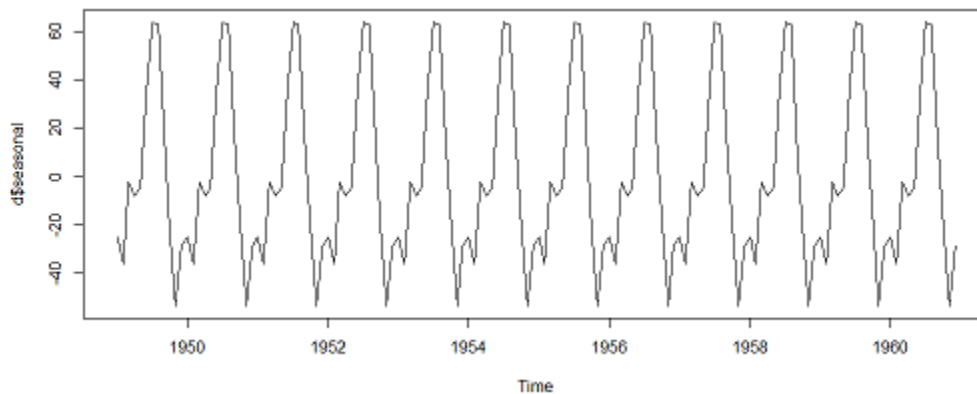


FIGURE 2.4 – saisonnalité de la série AirPassengers[6]

-La **composante résiduelle** (ou bruit de résidu) (ϵ_t) correspond à des fluctuations irrégulières, en général de faible intensité mais de nature aléatoire. On parle aussi d'aléas.



FIGURE 2.5 – la résidu de la série AirPassengers[6]

Les modèles présentés dans ce chapitre tiennent compte de ces trois composantes (tendance, saisonnalité et composante résiduelle).

-Des (grèves, conditions météorologiques exceptionnelles, crash financier) peuvent notamment intervenir appelés *cycles* en terminologie des séries temporelles.

-Une autre composante parfois étudiée de manière spécifique a trait au *phénomène cyclique* : c'est souvent le cas en climatologie et en économie (exemple : récession et expansion ...). Il s'agit d'un phénomène qui se répète mais contrairement à la saisonnalité sur des durées qui ne sont pas fixes et généralement plus longues. Sans informations spécifiques, il est généralement très difficile de dissocier tendance de cycle.

2.4 Objectifs principaux

L'étude d'une série temporelle permet *d'analyser*, de *décrire* et d'*expliquer* un phénomène au cours du temps et d'en tirer des conséquences pour des prises de décision (marketing ...).

cette étude permet aussi de faire un *contrôle*, par exemple pour la gestion des stocks, le contrôle d'un processus chimique ...

L'objectif principal de l'étude d'une série temporelle est la *prévision* qui consiste à prévoir les valeurs futures x_{n+h} , ($h = 1, 2, 3, \dots$) de la série temporelle à partir de ses valeurs observées jusqu'au temps n : x_1, x_2, \dots, x_n . La prévision de la série temporelle au temps $t + h$ est notée $\hat{x}_n(h)$. En générale elle est différente de la valeur réelle x_{n+h} que prend la série au temps $n + h$. Pour mesurer cette différence on définira l'*erreur de prédiction* par la différence $\hat{x}_n(h) - X_{n+h}$ en moyenne avec l'idée que plus h est grand plus l'erreur est grande.

L'intervalle de précision défini par les valeurs $\hat{x}_n^1(h)$ et $\hat{x}_n^2(h)$, est susceptible de contenir la valeur inconnue x_{n+h} .

La qualité de la prévision pourra être mesurée en se basant sur %80 des observations, puis en simulant une prévision sur les 20% des observations restantes. Cette technique est aussi utile pour :
-mesurer l'effet d'un phénomène accidentel(erreur, ...)

Un autre problème intéressant est la *détection de ruptures résultantes*, par exemple d'un changement économique (politique). Ces ruptures peuvent être de deux ordres : une rupture de niveau (par exemple la crise économique mondiale des années 2007 – 2012). la prévision de ces

dates de rupture est bien évidemment très importante.

Il existe encore d'autres objectifs immédiats relatifs à l'étude des séries chronologiques. Par exemple si deux séries sont observées, on peut se demander quelle influence elles exercent l'une sur l'autre. En notant X_t et Y_t les deux séries en question, on examine s'il existe par exemple des relations de type

$$Y_t = a_1 X_{t+1} + a_3 X_{t+3}$$

. D'abord on vérifie la **causalité** c'est à dire quelle variable (dans cet exemple (X_t)) va expliquer l'autre ((Y_t)), ce qui amène à la deuxième question, celle du **décalage temporel** : si une influence de (X_t) sur (Y_t) existe, avec quel délai et pendant combien de temps la variable explicative ((X_t)) influence-t-elle la variable expliquée ((Y_t)) ?.

Un autre problème macro économétrique, est de déterminer les relations à long terme des autres relations à court terme.

2.5 Description schématique de l'étude complète d'une série temporelle

Comme nous venons de le voir, l'un des objectifs principaux de l'étude d'une série temporelle est la **prévision** des valeurs futures de cette série. Pour cela, on a besoin de connaître ou tout au moins de **modéliser** le mécanisme de production de la série temporelle.

Notons que les variables X_t ne sont plus souvent ni indépendantes (on peut s'attendre en effet à ce que des observations relativement proches dans le temps soient liées) ni identiquement distribuées (dans la plupart des cas, le phénomène évolue, se modifie au cours du temps ce qui entraîne que les variables le décrivant ne sont pas équidistribuées). Cela nécessite des méthodes statistiques de traitement et de modélisation spécifiques puisqu'en particulier dans un cadre standard (celui de la description d'un échantillon) les méthodes statistiques classiques sont basées sur des hypothèses d'indépendance.

Schématiquement, les principales étapes de traitement d'une série temporelle sont les suivantes :

1. correction des données
2. observation de la série
3. analyse de la série à partir de ses composantes
4. modélisation (avec un nombre fini de paramètres)
5. diagnostic du modèle - ajustement au modèle
6. prédiction (prévision)

2.5.1 Correction des données

Avant de lancer une étude d'une série temporelle, il est souvent nécessaire de traiter, modifier les données brutes. Par exemple : -évaluation de données manquantes, remplacement de données accidentelles, ... etc.

-découpage en sous-séries.

-standardisation afin de se ramener à des intervalles de longueur fixe.

-transformation des données : on peut être parfois amenés à utiliser des données transformées pour des raisons diverses.

2.5.2 Observation de la série

Une règle générale en statistique descriptive consiste à commencer par regarder les données avant d'effectuer le moindre calcul. Ainsi, une fois la série corrigée et pré traitée, on trace son graphique. L'observation de ce graphique est souvent une aide à la modélisation de la série temporelle et permet de se faire une idée des différentes composantes de la série temporelle.

2.5.3 Analyse de série à partir de ses composantes

On étudie les composantes du modèle les unes après les autres.

2.5.4 Modélisation

Un *modèle* est une image simplifiée de la réalité qui vise à traduire les mécanismes de fonctionnement du phénomène étudié et permet de mieux les comprendre. Un modèle peut être qu'un autre pour décrire la réalité et bien sûr plusieurs questions se posent alors : comment mesurer cette qualité ? comment diagnostiquer un modèle ?

On distingue principalement deux types de modèles :

-les *modèles déterministes*, ces modèles relèvent de la statistique descriptive. Ils ne font intervenir que de manière sous-jacente le calcul des probabilités et consistent à supposer que l'observation de la série à la date t est une fonction du temps t et d'une variable ϵ_t centrée faisant office d'erreur au modèle, représentant la différence entre la réalité et le modèle proposé :

$$X_t = f(t, \epsilon_t)$$

On suppose de plus que les ϵ_t sont non corrélés.

Les deux modèles de ce type les plus utilisés sont :

1. le *modèle additif*. c'est le "modèle classique de décomposition" dans le traitement des modèles d'ajustement. La variable X_t s'écrit comme la somme de trois termes :

$$X_t = Z_t + S_t + \epsilon_t$$

, où Z_t représente la tendance (déterministe), S_t la saisonnalité (déterministe aussi) et ϵ_t les composantes erreurs au modèle aléatoires i.i.d¹.

1. indépendants et identiquement distribués

2. le **modèle multiplicatif** : la variable X_t s'écrit au terme d'erreur près comme le produit de la tendance et de la saisonnalité :

$$X_t = Z_t \times S_t \times \epsilon_t.$$

L'ajustement est ici multiplicatif et intervient dans les modèle **GARCH**

-les **modèle stochastiques** : Ils sont du même type que les modèles déterministes à ceci près que les variables de bruit (composante résiduelle) ϵ_t ne sont pas i.i.d mais possèdent une structure de corrélation non nulle : ϵ_t est une fonction des valeurs passées et d'un terme d'erreur η_t

Les deux types de modèles ci-dessus induisent des techniques de prévision bien particulières. Schématiquement, on s'intéresse tout d'abord à la saisonnalité éventuelle que l'on isole d'abord, ensuite on cherche à les modéliser, les estimer. Enfin on les élimine de la série : s'appellent la **détendancialisat**ion et la **désaisonnalisat**ion de la série. Une fois ces composantes éliminées on obtient la série aléatoire ϵ_t [6].

2.5.5 Diagnostic du modèle

Une fois le modèle construit et ses paramètres estimés ; on vérifie que le modèle proposé est bon c'est-à-dire l'ajustement au modèle :

- en étudiant les résidus
- en faisant des tests

2.5.6 Prévision

Enfin, une fois ces différentes étapes réalisées, nous sommes en mesure de faire de la prévision.

2.6 Processus stationnaire de second ordre

Une des grandes questions dans l'étude des série temporelles est de savoir si celles-ci suivant un processus stationnaire. On entend par là le fait que la structure du processus sous-jacent supposé évolue ou non avec le temps. Si la structure reste la même, le processus est dit alors stationnaire.

Définition 2.6.1. (**Stationnarité faible**) $(X_t)_{t \in Z}$ est un processus **stationnaire de second ordre** (ou faiblement stationnaire) s'il vérifie :

- (i) $\forall t \in Z, E(X_t) = m$
 - (ii) $\forall t \in Z, V(X_t) = \sigma^2 = \gamma(0)$
 - (iii) $\forall t \in Z, \forall h \in T, COV(X_t, X_{t+h})$ (ne dépend que de h)
- $\gamma(h)$ est **l'auto-covariance** d'ordre h de X_t [1]

Remarque 2.6.1. Dans la suite, les processus stationnaires désignent les processus de la définition précédente.

2.6.1 Bruit blanc

-**Bruit blanc faible** : $(\epsilon_t)_{t \in Z}$, si et seulement si :

$$\begin{aligned} E(\epsilon_t) &= 0, \forall t \in Z \\ V(\epsilon_t) &= \sigma^2, \forall t \in Z \\ COV(\epsilon_t, \epsilon_s) &= 0, \text{ si } t \neq s \end{aligned}$$

On notera $\epsilon \sim BB(0, \sigma^2)$

$BB()$ signifie Bruit blanc.

$-\epsilon_t$ est un **Bruit blanc fort** si et seulement si les ϵ_t sont i. i. d, $E(\epsilon_t) = 0$ et $V(\epsilon_t) = \sigma^2$. [11]

Définition 2.6.2. L'auto-covariance d'un processus stationnaire $(X_t)_{t \in Z}$ est définie par : [11]

$$\gamma : T \longrightarrow R$$

$$h \longrightarrow \gamma(h) = Cov(X_t, X_{t-h})$$

Définition 2.6.3. (Fonction d'auto-corrélation)

La fonction d'**auto-corrélation** d'un processus stationnaire $(X_t)_{t \in Z}$ est définie par :

$$\forall h \in Z, \rho(h) = \gamma(h) \div \gamma(0) = Corr(X_t, X_{t+h})$$

Définition 2.6.4 (Fonction d'auto-corrélation partielle)

On appelle auto-corrélation partielle d'ordre p

$$r(p) = Corr(X_t - EL(X_t|X_{t-1}, \dots, X_{t-p+1}), X_{t-p} - EL(X_{t-p}|X_{t-1}, \dots, X_{t-p+1}))$$

$$= \frac{Cov(X_t - EL(X_t|X_{t-1}, \dots, X_{t-p+1}), X_{t-p} - EL(X_{t-p}|X_{t-1}, \dots, X_{t-p+1}))}{[Var(X_t - EL(X_t|X_{t-1}, \dots, X_{t-p+1}))Var(X_{t-p} - EL(X_{t-p}|X_{t-1}, \dots, X_{t-p+1}))]^{\frac{1}{2}}} [11]$$

2.6.2 Polynomes retard, avance et différence

Définition 2.6.5. -l'opérateur retard L (lag) ou B (backward) est défini sur la classe des processus stationnaire comme étant :

$$L : (X_t)_{t \in Z} \longmapsto (Y_t)_{t \in Z} \text{ tel que } Y_t = X_{t-1}$$

$$\text{On note : } LX_t = X_{t-1}$$

-De la même façon, **l'opérateur avance** F (forward) correspond à

$$F : (X_t)_{t \in Z} \longmapsto (Y_t)_{t \in Z} \text{ tel que } Y_t = X_{t+1}$$

$$\text{On note : } FX_t = X_{t+1}$$

-**l'opérateur de différence** $\Delta = (1 - L)$

-l'opérateur de différence d'ordre k $\Delta_k = (1 - L^k)[10]$

2.7 Processus ARMA et ARIMA

2.7.1 Processus auto-régressifs d'ordre p (AR(p))

$(X_t)_{t \in \mathbb{Z}}$ est un processus AR(p) si :

(X_t) est stationnaire.

(X_t) vérifie une équation $(X_t) = \mu + \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \epsilon_t$ avec φ_i sont des réels, $\varphi_p \neq 0$ et ϵ_t est un bruit blanc de variance σ^2 , $\epsilon_t \sim BB(0, \sigma^2)$.

On note $\phi(L)X_t = \mu + \epsilon_t$ où $\phi(L) = 1 - \varphi_1 L + \dots + \varphi_p L^p [2]$

Propriétés des processus AR(p)

On suppose que $\phi(L)X_t = \mu + \epsilon_t$ où
 -les racines de ϕ sont de module strictement supérieur à 1
 - ϵ_t suit un bruit blanc.

On peut se ramener ensuite à $\mu = 0$ par centrage car

$\phi(L)(X_t - m) = \epsilon_t$ où $m = \frac{\mu}{\phi(1)} = E(X_t)$
 On considère donc le cas où $\phi(L)(X_t) = \epsilon_t (E(X_t) = 0)$

Auto-covariance, Auto-corrélation et équations de Yule-Walker

-L'auto-covariance :

$\gamma(h) = Cov(X_t, X_{t-h}) = E(X_t X_{t-h})$ pour $h \geq 0$ ($carm = 0$)

On a alors

$$\gamma(0) = \varphi_1 \gamma(1) + \dots + \varphi_p \gamma(p) + \sigma_\epsilon^2$$

Si $h > 0$ on a :

$$\gamma(h) = \varphi_1 \gamma(h-1) + \dots + \varphi_p \gamma(h-p) + E(\epsilon_t X_{t-h})$$

-Les auto-corrélations :

A partir de la relation de récurrence de $\gamma(h)$ on déduit celle de $\rho(h) = \frac{\gamma(h)}{\gamma(0)}$.

$$\rho(h) = \varphi_1 \rho(h-1) + \dots + \varphi_p \rho(h-p), \forall h \geq 0$$

Pour $h > 0$, les $\gamma(h)$ et les $\rho(h)$ vérifient une relation de récurrence d'ordre p et

$$1 = \varphi_1 \rho(1) + \dots + \varphi_p \rho(p) + \frac{\sigma_\epsilon^2}{\gamma(0)}$$

$$\implies \gamma(0) = \sigma_\epsilon^2 \frac{1}{1 - \varphi_1 \rho(1) + \dots + \varphi_p \rho(p)}$$

Les équation de Yule-Walker pour $h = 1, \dots, p$ peuvent s'écrire :

$$\begin{pmatrix} 1 & \rho(1) & \dots & \rho(p-1) \\ \rho(1) & 1 & \rho(1) & \rho(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(p-1) & \dots & \dots & 1 \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_p \end{pmatrix} = \begin{pmatrix} \rho(1) \\ \vdots \\ \rho(p) \end{pmatrix}$$

Les solutions de l'équation de récurrence sont déterminées par la donnée de conditions initiales $\rho(1), \dots, \rho(p)$: elle permettent d'obtenir $\varphi_1, \dots, \varphi_p$.

$$\begin{aligned} \rho(1) &= \varphi_1 + \varphi_2 \rho(1) + \dots + \varphi_p \rho(p-1) \\ \rho(p) &= \varphi_1 \rho(p-1) + \dots + \varphi_{p-1} \rho(1) + \varphi_p \\ \varphi_1 &= (1 - \varphi_2) \rho(1) - \dots - \varphi_p \rho(p-1) \end{aligned}$$

...

$$\varphi_1 = \rho(p) - \varphi_2 \rho(p-1) - \dots - \varphi_{p-1} \rho(1)$$

On peut donc aussi abtenir $\rho(1), \dots, \rho(p)$ en fonction de $\varphi_1, \dots, \varphi_p$.

Proposition 2.7.1. Si $X_t \sim AR(p)$ alors les $|\rho(h)|$ et les $\gamma(h)$ décroissent vers 0 exponentiellement avec h.

Proposition 2.7.2. Si $(X_t)_{t \in \mathbb{Z}} \sim AR(p)$ et si $\phi(L)X_t = \mu + \epsilon_t$ est sa représentation canonique, alors :

$$r(h) = \begin{cases} 0 & \text{si } h > 0 \\ \neq 0 & \text{sinon} \end{cases}$$

2.7.2 Processus moyenne mobile d'ordre q (MA(q))

$(X_t)_{t \in \mathbb{Z}} \sim MA(q)$ s'il existe $\epsilon_t \sim BB(0, \sigma^2)$ et $\theta_1, \dots, \theta_q$ tel que

$$X_t = m + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q}$$

On note $X_t = m + \theta(L)\epsilon_t$ où $\theta(L) = 1 - \theta_1 L - \dots - \theta_q L^q$ [2]

Propriétés des processus MA(q)

On suppose que la représentation étudiée est la représentation canonique.

$$X_t = m + \theta(L)\epsilon_t$$

toute les racines de θ sont de module > 1 .

$$\theta(L) = 1 - \theta_1 L - \dots - \theta_q L^q$$

$$\epsilon_t \sim BB(0, \sigma^2)$$

Proposition 2.7.3. Auto-corrélation

$\rho(h) = 0$ si $|h| > q$ et $\rho(q) \neq 0$

Proposition 2.7.4. On a pas de résultat particulier pour les auto-corrélations partielles.

2.7.3 Processus ARMA(p, q)

Un processus stationnaire $(X_t)_{t \in \mathbb{Z}}$ admet une représentation ARMA(p, q) canonique minimale s'il vérifie une équation :

$$\phi(L)X_t = \mu + \theta(L)\epsilon_t$$

où

(i) $\epsilon_t \sim BB(0, \sigma^2)$

(ii) $\phi(L) = 1 - \varphi_1 L \dots \varphi_p L^p$ avec $\varphi_p \neq 0$

(iii) $\theta(L) = 1 - \theta_1 L - \dots - \theta_q L^q$ avec $\theta_q \neq 0$

(iv) $\phi(L)$ et $\theta(L)$ ont toutes leurs racines de module strictement supérieur à 1 (représentation canonique).

(v) $\phi(L)$ et $\theta(L)$ n'ont pas de racines communes (représentation minimale). [4]

(Auto-covariance et auto-corrélation)

(i) pour $h > q$ les $\gamma(h)$ et les $\rho(h)$ vérifient les équation de récurrence d'ordre p :

$$\gamma(h) - \varphi_1 \gamma(h-1) - \dots - \varphi_p \gamma(h-p) = 0$$

$$\rho(h) - \varphi \rho(h-1) - \dots - \varphi_p \rho(h-p) = 0$$

(ii) Elles décroissent donc vers 0 exponentiellement avec h , pour $h > q$.

Un processus ARMA(p, q) vérifie alors (équations de Yule-walker) :

$$\begin{pmatrix} \rho(p) & \dots & \rho(q+p-1) \\ \vdots & \ddots & \vdots \\ \rho(q+p-1) & \dots & \rho(q) \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_p \end{pmatrix} = \begin{pmatrix} \rho(p+1) \\ \vdots \\ \rho(p+q) \end{pmatrix}$$

Quand ρ est connu ou estimé, on peut alors calculer les ϕ_j ou inversement. quand les ϕ_j sont connus, on calcule $\rho(q+1) \dots \rho(q+p)$.

2.7.4 Processus ARIMA(p, d, q)

Ces processus sont non stationnaires dès que $d \leq 1$. Les séries économiques sont souvent non stationnaires.

Définition 7. 1. $(X_t)_{t \in \mathbb{Z}}$ est un processus ARIMA(p, d, q) en représentation canonique minimale s'il vérifie une équation de type :

$$\forall t \geq 0, (1 - L)^d \phi(L) X_t = \mu + \theta(L) \epsilon_t$$

Et ceci avec :

(i) $\epsilon_t \sim BB(0, \sigma^2)$

(ii) $\phi(L) = 1 - \varphi_1 L - \dots - \varphi_p L^p$ où $\varphi_p \neq 0$

$\theta(L) = 1 - \theta_1 L - \dots - \theta_q L^q$ où $\theta_q \neq 0$

(iii) $\theta(L)$ et $\phi(L)$ ont leurs racines de module > 1 et n'ont pas de racines communes.

(iv) conditions initiales

$$Z = (X_{-1}, \dots, X_{-p-d}, \epsilon_{-1}, \dots, \epsilon_{-q})$$

telles que $COV(\epsilon_t, T) = 0$. [5]

2.7.5 Processus SARIMA

Définition 2.7.2. [5]

le processus stationnaire, $(X_t)_{t \in \mathbb{Z}}$ satisfait une représentation SARIMA saisonnier (Seasonal Integrated AutoRegressive Moving Average) d'ordre p et q , intégré d'ordre d , de période s et s' , noté $SARIMA_{s,s'}(p, d, q)$ si :

$$\phi(L^s)(1 - L^s)^d X_t = \mu + \theta(L^{s'}) \epsilon_t$$

avec :

(i) $\epsilon_t \sim BB(0, \sigma^2)$

(ii) $\phi(L) = 1 - \varphi_1 L - \dots - \varphi_p L^p$ où $\varphi_p \neq 0$

$\theta(L) = 1 - \theta_1 L - \dots - \theta_q L^q$ où $\theta_q \neq 0$

(iii) $\theta(L)$ et $\phi(L)$ ont leurs racines de module > 1 et n'ont pas de racines communes.

(iv) $\mu \in \mathbb{R}, \phi_0 = \theta_0 = 1$.

Chapitre 3

Application à la prévision des ventes de Florex équipements

introduction

Le travail qui suit reposera sur le journal des ventes mensuelles des années 2017 et 2018 récoltées au niveau de l'entreprise Florex équipements. l'étude concernera les trois produits :

- Casque de sablage (Cds)
- Grenaille acier angulaire(Gaa)
- Unité de sablage 200 litres(Us 200)

Les données seront traitées via le logiciel **R** ; Ce dernier est langage de programmation et un logiciel libre destiné aux statistique et à la science des données soutenu par la R Foundation for Statistical Computing. Il est largement utilisé par les statisticiens, les data miners, data scientists pour le développement de logiciels statistiques et l'analyse des données[7].

3.1 Modélisation des séries de données de ventes avec R

3.1.1 importation des données à partir d'un fichier Excel

Les données collectées ont été organisées dans un fichier excel nommé "Donnée.xlsx". Les commandes R avec lesquels on charge ces dernières dans l'environnement de travail R sont :

library(readxl) : permet de charger la bibliothèque readxl de R avec laquelle on peut importer des données directement à partir des feuilles de calcul d'excel. Il existe plusieurs autres commandes mais nécessitent un détour énorme et donc moins efficace dans notre cas.

`d=read_excel("c ://Donnée.xlsx")` : la commande de la bibliothèque précédente qui permet de lire les données à partir du fichier Excel "donnée.xlsx").

`d<-as.data.frame(d)` : sauvegarde ces données comme étant un tableau R comme suit :

Le résultat qu'affiche R en exécutant cette commande est le suivant :

	mois	Cds	Gaa	Us2001
Jan 2017	5	7	7050	11
Feb 2017	4	10	4000	0
Mar 2017	9	6	4000	7
Apr 2017	2	1	4025	5
May 2017	8	5	4000	7
Jun 2017	7	18	6000	0
Jul 2017	6	11	0	1
Aug 2017	1	4	5000	2
Sep 2017	12	7	500	6
Oct 2017	11	0	3425	4
Nov 2017	10	27	2000	4
Dec 2017	3	8	1000	6
Jan 2018	5	0	6000	2
Feb 2018	4	3	1000	4
Mar 2018	9	2	2000	1
Apr 2018	2	2	0	3
May 2018	8	0	2000	10
Jun 2018	7	1	5000	0
Jul 2018	6	1	12000	0
Aug 2018	1	25	2000	8
Sep 2018	12	4	1000	1
Oct 2018	11	7	8450	2
Nov 2018	10	7	0	3
Dec 2018	3	6	0	1

FIGURE 3.1 – Tableau des données relatives aux ventes mensuelles 2017–2018 des produits cds, gaa et Us 200

3.1.2 Modélisation de la série Cds : Casque de sablage

Recherche du modèle ARIMA adéquat à cette série

a) structuration des données comme étant une série temporelle avec la commande `ts` comme suite :

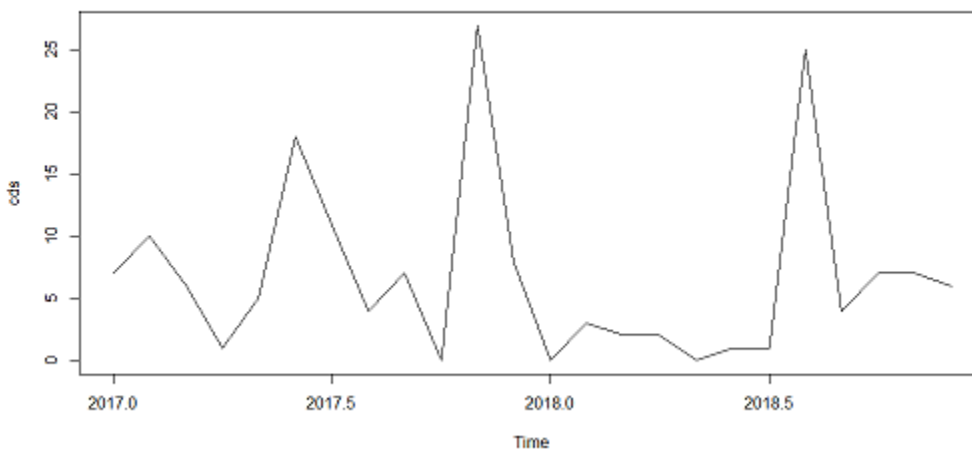
```
cds=ts(d$Cds,start=c(2017,1),end=c(2018,12),frequency = 12)
```

```
cds
```

b) On obtient alors un tableau (série temporelle (`cds`)) qui représente l'évolution mensuelle de la quantité de Casque de sablage vendu.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2017	7	10	6	1	5	18	11	4	7	0	27	8
2018	0	3	2	2	0	1	1	25	4	7	7	6

d) Examen du graphe de la série : Avec la commande `plot(cds)` on obtient le graphe de la série 'cds' :



Le graphe montre que la série évolue horizontalement à l'axe des abscisses, ce qui nous laisse penser à une absence de la tendance dans cette série.

d'autre part on remarque des piques qui se répètent à des périodes précises de la chronique ce qui implique une composante saisonnière.

En examinant le corrélogramme et corrélogramme partiel, on pourra voir plus clair l'existence ou non des ces deux composantes.

e) ACF et PACF de la série `cds` :

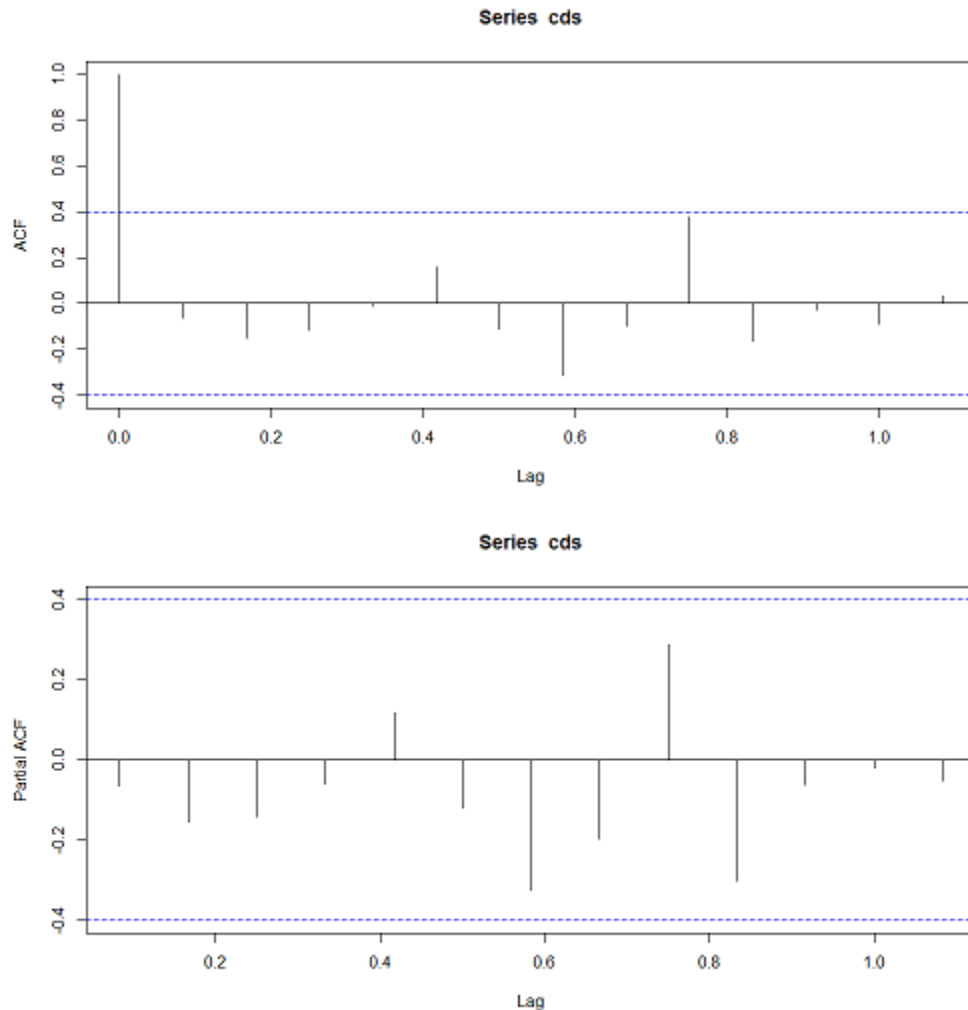


FIGURE 3.2 – Corrélogramme et corrélogramme partiel de la série Cds

L'examen de ces deux graphes montre que la série ne contient aucune saisonnalité (d'après l'ACF), par contre le PACF montre une variation sinusoïdale des auto-corrélations partielles, qui confirme l'existence de la saisonnalité.

L'examen des ACF et PACF montrent aussi que les auto-corrélations et les auto-corrélations partielles ont comprise dans l'intervalle de confiance d'un bruit blanc.

f) Identification du modèle :

D'après les résultats précédents, on constate que le modèle adéquat pour la série Cds est un bruit blanc blanc saisonnier.

Ce dernier est soit avec où sans dérive (constante); on l'estime dans R avec :

```
out = arima(cds, order = c(0, 1, 0), seasonal = list(order = c(0, 1, 0), period = 12))
```

toute fois R comprend une commande qui examine plusieurs modèles ARIMA candidats pour modéliser cette série pour en sélectionner le meilleur selon un critère de choix, comme suit :

```
auto.arima(cds, seasonal = TRUE)
```

Le résultat de cette commande nous donne : Alors la série "cds" est un $ARIMA(0, 0, 0)$ écrit sous la forme : $X_t = \mu + \epsilon_t$, où $\epsilon_t \sim BB(0, \sigma^2)$, $\mu = 6,75$

```

> auto.arima(cds)
Series: cds
ARIMA(0,0,0) with non-zero mean

Coefficients:
  mean
 6.7500
s.e. 1.4521

sigma^2 estimated as 52.8: log likelihood=-81.14
AIC=166.29 AICc=166.86 BIC=168.64

```

Le modèle retenu est donc : $X_t = 6,75 + \epsilon_t$.

g) Analyse des résidus

La dernière étape d'une modélisation par modèle ARIMA avant de faire des prévisions, est de vérifier que les résidus forment un bruit blanc.

Dans R on procède de la manière suivante :

1. On calcule les résidus avec la commande `res = out$residuals`
2. On examine le corrélogramme et le corrélogramme partiel des résidus (figure 3.3) :

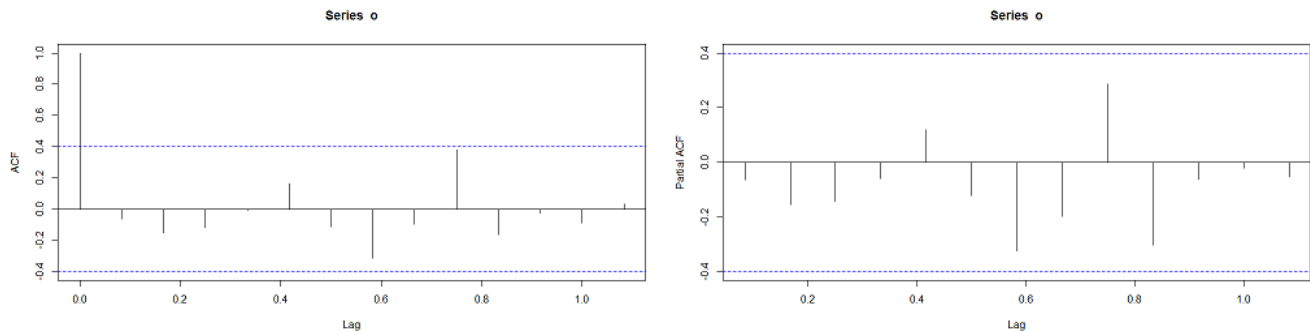


FIGURE 3.3 – Corrélogramme et corrélogramme partiel des résidus de la série Cds

D'après le graphique les résidus sont tous dans l'intervalle de confiance d'un bruit blanc. Ce résultat peut être confirmé avec le test de Box et Pierce.

3. Le test de Box et Pierce : `box.test(o)` :

```
> Box.test(o)
```

```
Box-Pierce test
```

```
data: o
```

```
X-squared = 0.10235, df = 1, p-value = 0.749
```

FIGURE 3.4 –

La p -value = 0.749 > 0,05, donc les résidus vérifient la propriété d'un bruit blanc. Une fois le modèle est validé on passe à la prévision.

3.1.3 Prévision de la série Cds

Pour calculer des prévisions d'une série chronologique, on utilise la fonction `predict()[3]`
La commande R pour prévoir la série Cds est donc :

```
prevision = predict(out, n.ahead = 6)(on fais une prévision sur un horizon de six mois)
```

Le résultat de cette commade nou donne :

```
> prevision = predict(out,n.ahead=6)
> prevision
$pred
      Jan Feb Mar Apr May Jun
2019 6.75 6.75 6.75 6.75 6.75 6.75

$se
      Jan Feb Mar Apr May Jun
2019 7.113661 7.113661 7.113661 7.113661 7.113661 7.113661
```

3.1.4 Modèles des séries Gaa e Us 200

Un travail similaire est réalisé sur R pour les deux séries Gaa et Us 200. Les modèles validés pour ces deux produits sont comme suit :

Série Gaa

(Gaa) $\sim ARMA(0,0,0)$ alors l'équation de ce modèle s'écrit sous la forme :

$$Gaa_t = 3352,083 + \epsilon_t$$

Série Us 200

La série temporelle (us) est un $ARIMA(0,0,1)$ (Figure 3.10)

Donc elle s'écrit sous la forme :

$$U_{st} = 3,5962 + \epsilon_t + 0,3585\epsilon_{t-1}$$

Les différents graphes et figures des corrélogrammes et corrélogrammes partiels de ces deux séries, les graphes et corrélogrammes des résidus ainsi que les résultats des différents tests ont résumés dans l'annexe A

3.2 Conclusion

Dans ce chapitre, on a montré la valeur de l'analyse des séries chronologiques dans la modélisation et la prévision des ventes des produits au niveau de l'entreprise Florex equipments. les modèles SARIMA sont très efficaces pour de telles série où l'historique des ventes est assez grand.

Conclusion Générale

Parmi toutes les prévisions qu'une organisation peut faire l'estimation de la demande future est une donnée clé, car c'est l'une des plus utiles pour les entreprises. En effet bien prévoir la demande client permet par exemple d'établir quelle capacité de production est requise afin d'ajuster l'offre à la demande, de déterminer les meilleures stratégies de production, de planifier l'utilisation des équipements et les besoins en équipements, de planifier la main-d'oeuvre requise, d'orienter la politique et les stratégies de gestion des stocks.

Dans ce travail, on a pu élaborer une modélisation des ventes de trois produits fabriqués au niveau de l'entreprise FLOREX EQUIPEMENTS avec les modèles SARIMA offert par l'analyse des séries chronologiques.

Comme suite à ce travail, on peut penser à :

- Généraliser les résultats sur l'ensemble de tout les produits de cette entreprise,
- essayer d'autres modèles de prévisions et comparer les résultats des différents modèles,
- Etudier la relation entre ces prévisions et les autres fonctions de l'entreprise telle que la gestion des stocks, pour optimiser le bénéfice de l'entreprise à son maximum.

Annexe A

Graphes et figures des Résultats De Modélisation sous R

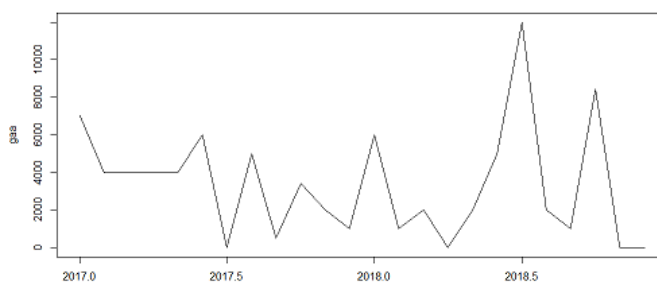


FIGURE A.1 – Graphe de la série Gaa

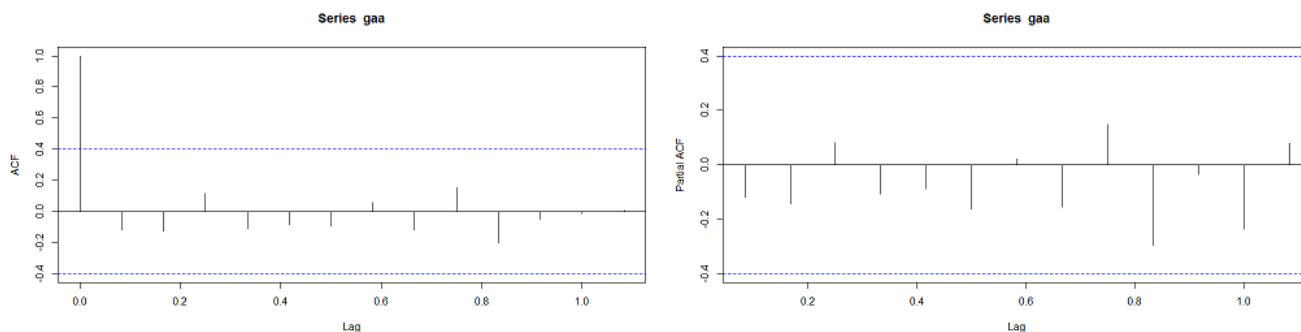


FIGURE A.2 – Corrélogramme et corrélogramme partiel de la série Gaa

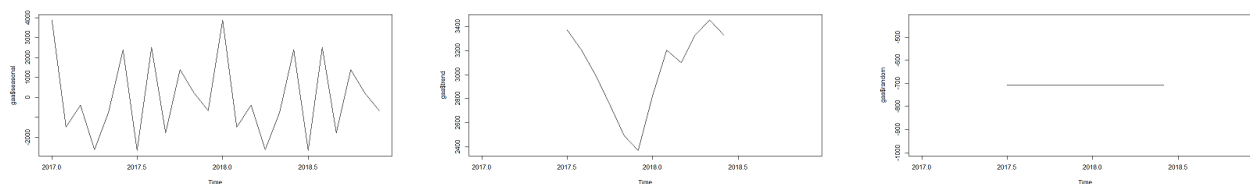


FIGURE A.3 – Les trois composantes de la série Gaa(saisonnalité, tendance e résiduelle)

```
> auto.arima(gaa)
Series: gaa
ARIMA(0,0,0) with non-zero mean
```

```
Coefficients:
  mean
3352.083
s.e. 607.551
```

```
sigma^2 estimated as 9243963: log likelihood=-
226.02
AIC=456.04 AICc=456.61 BIC=458.39
```

```
> out = arima(gaa,order = c(0,0,0))
> out
```

```
Call:
arima(x = gaa, order = c(0, 0, 0))
```

```
Coefficients:
  intercept
3352.083
s.e. 607.551
```

```
sigma^2 estimated as 8858798: log likelihood = -226.02, aic = 456.04
```

FIGURE A.4 – Résultats de la commande auto.arima pour la série Gaa

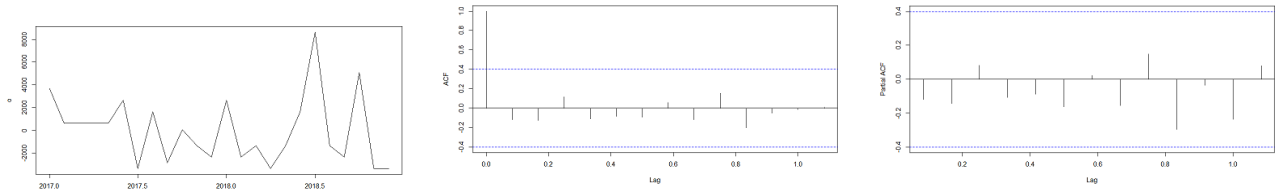


FIGURE A.5 – Graphe, ACF et PACF des Résidus Gaa

```
> Box.test(o)
```

Box-Pierce test

```
data: o
```

```
X-squared = 0.34164, df = 1, p-value = 0.5589
```

FIGURE A.6 – Résultat du Test de Box et Pierce sur les Résidus Gaa

```
> predict(out,n.ahead=6)
$pred
      Jan      Feb      Mar      Apr      May      Jun
2019 3352.083 3352.083 3352.083 3352.083 3352.083 3352.083

$se
      Jan      Feb      Mar      Apr      May      Jun
2019 2976.373 2976.373 2976.373 2976.373 2976.373 2976.373
```

FIGURE A.7 – Prévission des ventes Gaa sur l’horison Janvier 2019-Juin 2019

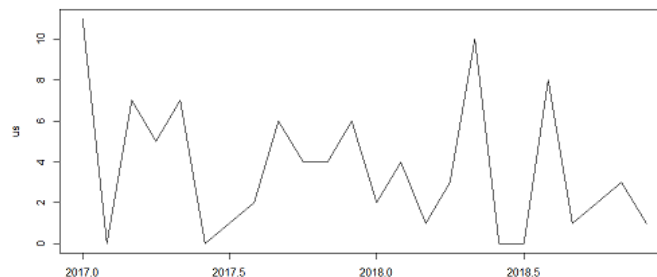


FIGURE A.8 – Graphe de la série Us 200

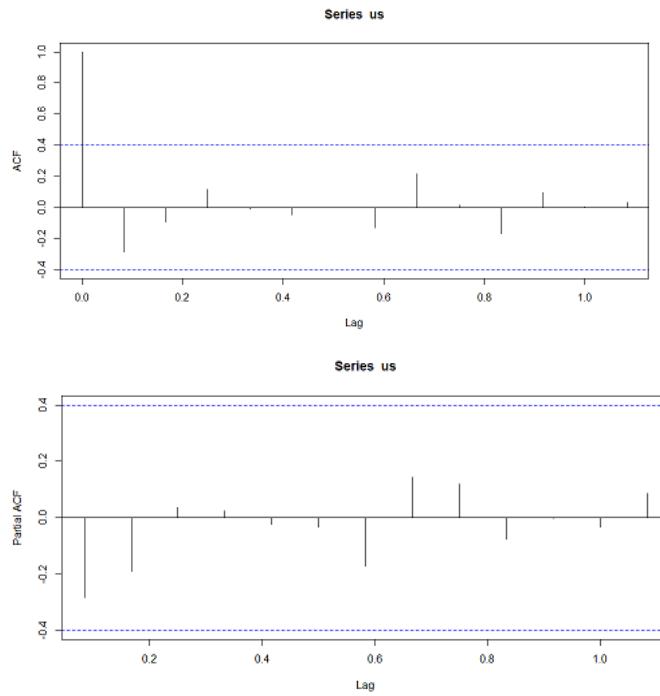


FIGURE A.9 – ACF et PACF de la série Us 200

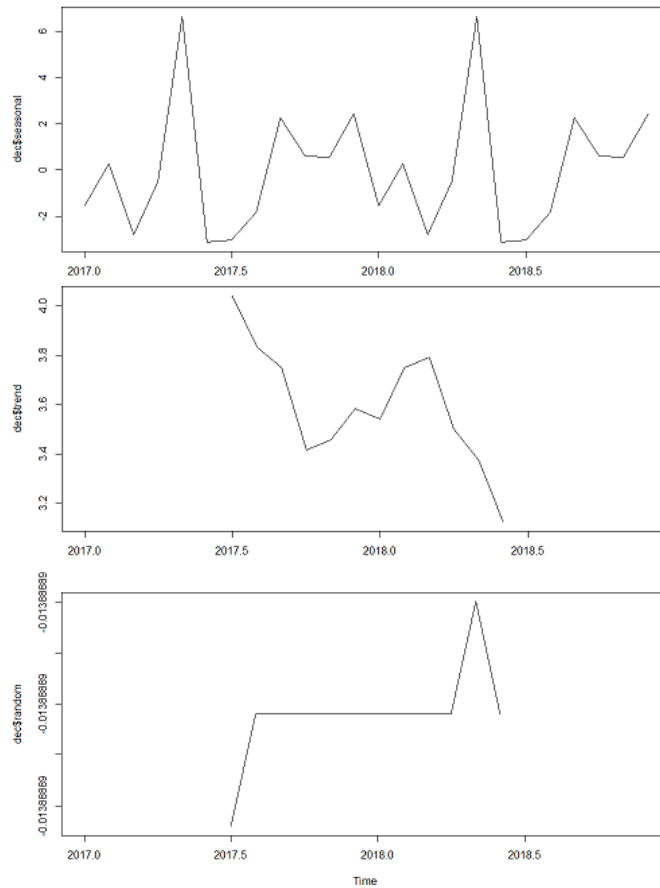


FIGURE A.10 – Les trois composantes de la série Us 200(saisonnalité, tendance e résiduelle)

```

> auto.arima(us)
Series: us
ARIMA(0,0,1) with non-zero mean

Coefficients:
    ma1 mean
   -0.5385 3.5962
s.e. 0.2884 0.2884

sigma^2 estimated as 9.153: log likelihood=-59.75
AIC=125.5 AICc=126.7 BIC=129.04

> outus = arima(us,order = c(0,0,1))
> outus

Call:
arima(x = us, order = c(0, 0, 1))

Coefficients:
    ma1 intercept
   -0.5385  3.5962
s.e. 0.2884  0.2884

sigma^2 estimated as 8.39: log likelihood = -59.75, aic = 125.5

```

FIGURE A.11 – Résultats de la commande auto.arima pour la série Us 200

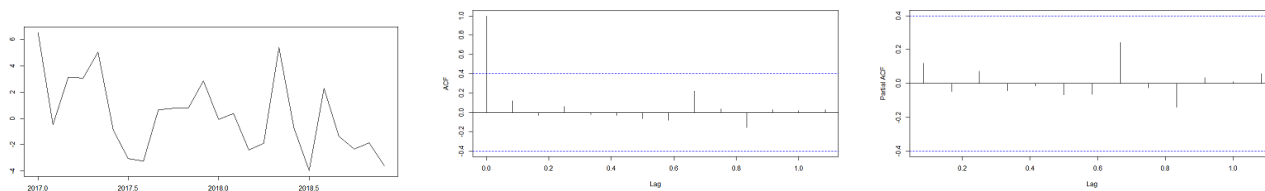


FIGURE A.12 – Graphe, ACF et PACF des résidus Us 200

```

> Box.test(o)

Box-Pierce test

data: o
X-squared = 0.33674, df = 1, p-value = 0.5617

```

FIGURE A.13 – Résultat du Test de Box et Pierce sur les Résidus Us 200

```

> prevision = predict(outus,n.ahead=6)
> prevision
$pred
      Jan      Feb      Mar      Apr      May      Jun
2019 5.532486 3.596248 3.596248 3.596248 3.596248 3.596248

$se
      Jan      Feb      Mar      Apr      May      Jun
2019 2.896585 3.289933 3.289933 3.289933 3.289933 3.289933

```

FIGURE A.14 – Préviation des ventes Us 200 sur l’horison Janvier 2019-Juin 2019

Bibliographie

- [1] Hélène Hamisultane. Econométrie des séries temporelles. Lecture note Licence. CNRS, Paris. 2002.
- [2] Agnès Lagnoux. Séries chronologique. Cours Master, Université de Toulouse. 2008.
- [3] Aragon Yives. Séries temporelles avec R. EDP Sciences. 2016.
- [4] Olivier Cappé, Maurice Charbit and Eric Moulines. Introduction aux Séries temporelles. Springer. 2017.
- [5] Régis Bourbonnais. Économétrie - 9e édition : Cours et exercices corrigés. Dunod. 2015
- [6] Régis Bourbonnais, Michel Terraza. Analyse des séries temporelles : Applications à l'économie et à la gestion - Cours et exercices corrigés. 4eme édition. Dunod. 2016.
- [7] Pierre Lafaye de Micheaux, Rémy Drouilhet and Benoit Liquet. Le logiciel R - Maîtriser le langage - Effectuer des analyses statistiques. Springer. 2010.
- [8] René DARMON. Pilotage dynamique de la force de vente. Edition Village Mondial, Paris. 2001.
- [9] Seghier Abdellatif and all. Prediction d'un processus stationnaire du second ordre de covariance connue sur un intervalle fini. Illinois journal of Mathematics. 1978.
- [10] Patricia. Charpentier. Organisation et gestion de l'entreprise. Edition Nathan, Paris. 1997.
- [11] George E. P. Box, Gwilym M. Jenkins and Gregory C. Reinsel. Time Series Analysis : Forecasting and Control. Wiley ; 4th edition (June 30, 2008).