

République Algérienne Démocratique et Populaire
Ministère De L'enseignement Supérieur Et De La Recherche Scientifique

Université Dr. Moulay Tahar - SAIDA –



Faculté des sciences économiques, commerciales

& sciences de gestions

Département sciences économique

Mémoire de fin d'études

pour l'obtention du diplôme de Master en science économique

Option: méthodes quantitatives de gestion

« MQG »

Thème :

**Prévision de la consommation d'électricité pour la distribution
publique par la méthode BOX & JENKINS.**

Direction De Distribution D'électricité Et Du Gaz

SAIDA

Réalisé par :

-Boubekeur Khadîdja

Encadré par docteur:

-Benhamida Mohamed

Présenté devant les jurys :

Mr.....(Président)

Mr..... Benhamida Mohamed.....(Encadreur)

Mr.....(Examineur)

Mr.....(Examineur)

Année universitaire

2012-2013

Remerciements

Cette thèse doit beaucoup aux nombreuses personnes qui m'ont encouragé, soutenu et conforté au long de toutes ces années. Qu'elles trouvent dans ce travail l'expression de mes plus sincères remerciements.

Je souhaite renouveler mes remerciements à docteur BENHAMIDA MOHAMED pour avoir accepté et encadré cette thèse. J'aimerais remercier, docteur BEN KADOUR pour la confiance qu'il m'a témoigné du début à la fin du travail, pour sa disponibilité à intuitions. Sincèrement, grâce à lui j'ai pu apprendre beaucoup de choses importantes pour mon développement personnel. J'aimerais aussi remercier docteur SOUAR YUCEF, Grâce à lui aussi j'ai pu valoriser mes travaux en me mettant en relation avec plusieurs spécialistes de mon domaine de recherche.

Je n'oublie pas enfin La justesse de ses critiques a été très constructive et utile. Je lui exprime ma très profonde gratitude. J'aimerais également remercier mes professeurs Moulay, BELLAHOUELLE, BOUMAAZA, TADJ, ABDELLAH pour avoir suivi mes travaux pendant la durée de ma thèse. Leurs conseils et remarques, toujours pertinents, Je remercie les membres du jury pour avoir accepté de participer à mon jury de thèse.

Par l'occasion, je remercie MR SOUIAH, MR MOUHAND et MME MONSEUR pour m'avoir permis de présenter mes travaux de recherches.

Une occasion qui m'a permis d'avoir des critiques constructives très utiles pour la continuation de mes travaux de recherche.

Je remercie beaucoup le directeur du sdo MR SBAA et MR TAOUCH pour leurs efforts dans la réalisation de la partie pratique de cette thèse.

Je ne peux également oublier de remercier MR FELIH pour son aide. La thèse a parfois été un moment difficile pour mes proches. Elle est très preneuse de temps ! Et j'avoue ne pas les avoir consacré le temps qu'ils le méritent. Mon fiancé, je te remercie pour tes encouragements et ton soutien. Malgré la période difficile qu'on a traversé, tu m'as été toujours disponible et compréhensive. J'ai également pu compter sur l'aide de mes parents mes sœurs et mon frère. Ils ont toujours cru

en moi et j'espère être toujours à leur hauteur. Un remerciement spécial pour mon père, qui sans lui, je n'aurai jamais eu l'opportunité de continuer mes études et effectuer cette thèse. Je remercie aussi tous mes amis, merci pour m'avoir accompagné et aidé pendant ces dernières années.

Introduction générale:

Les techniques de prévision des séries temporelles se révèlent dans la plupart des cas suffisantes pour prévoir les phénomènes économiques.

Elles sont construites à partir de l'extrapolation de composantes déterministes qui mutilent la réalité économique.

La prévision des séries temporelles a connu un développement important depuis trente ans. La diffusion de logiciels spécialisés la met à la portée de toutes les organisations. La prévision est fondamentale dans la mesure ou elle est à la base de l'action. La prise de décision doit en effet toujours reposer sur des prévisions, c'est ainsi que la Société Nationale de l'Electricité et du Gaz « SONELGAZ » (lieu de mon stage) s'intéresse aux prévisions de la consommation de l'électricité et du gaz afin de faire face à la demande de sa clientèle, mais aussi orienter sa politique commerciale (prix, marketing... etc).

C'est dans ce but que la compagnie m'a chargé de faire une étude scientifique et de préparer un modèle prévisionnel sur la consommation d'électricité pour la distribution publique.

Afin de répondre aux attentes de la SONELGAZ, j'ai pensé d'utiliser une méthode de prévision à court terme : c'est la méthode de Box & Jenkins.

les premières études sur les chroniques ont amené à considérer de façon standard quatre grandes composantes : la tendance, le cycle conjoncturel, la composante saisonnière et la composante résiduelle.

La méthode de Box & Jenkins qui est la méthode de prévision la plus élaborée, sa caractéristique essentielle est d'effectuer une modélisation stochastique où l'évolution de la série est la réalisation d'un processus aléatoire déterminé. Contrairement aux autres méthodes de prévision, ce modèle n'est pas donné a priori mais sélectionné dans une classe très étendue de modèles, les modèles *ARIMA* (Auto Regressive Integrated Moving Average).

C'est ainsi que j'ai essayé de part ce travail, de répondre à la question:

Pour quel horizon on pourra faire des prévision de la consommation d'électricité pour la distribution publique par la méthode de Box & Jenkins?

Afin de répondre à ma problématique j'ai structuré mon travail en trois chapitres :

1- Le premier chapitre est consacré à la prévision de la consommation d'électricité (définition de la prévision, objectif, classification des techniques de prévision, les types de la consommation d'électricité et clientèle électricité.

2- Le deuxième chapitre quant à lui portera sur la présentation théorique des séries temporelles à savoir la méthode de prévision Box & Jenkins.

3- Enfin le troisième chapitre est consacré à l'application de cette méthode et offrir des résultats sur cette prévision.

Pour finir ce travail, une conclusion générale est présentée.

Chapitre 01:

Prévision de la

consommation

d'électricité

La prévision de la consommation d'électricité est une des missions essentielles de la SDO. La qualité de cette prévision, élément essentiel de la préparation et de l'anticipation, contribue à garantir l'équilibre offre-demande d'électricité à tout instant, et influe donc directement sur la sûreté du système électrique.

1.1 Généralité sur la prévision:

1.1.1 Définition de la prévision :

Une prévision peut être défini comme un ensemble de probabilités associées à un ensemble d'événements futurs Cette prévision est basée sur un ensemble d'informations disponibles à l'instant t où elle a été effectuée. Cet ensemble noté Ω_t (l'indice temporel t correspond à l'instant t) représente les données disponibles, les connaissances et les théories concernant le phénomène que l'on souhaite prévoir¹.

1.1.2 Objectif de la prévision :

L'idéal pour une entreprise : produire exactement les produits que ses clients vont acheter. Mais sauf dans le cas particulier de la production à la commande cela n'est pas possible. Ainsi afin de prendre les décisions relatives à son bon fonctionnement et à sa pérennité, toute entreprise doit s'appuyer sur un système de prévisions fiables².

1.1.3 Horizon de prévision:

L'incertitude augmente avec l'horizon de planification. Selon le type de décisions à prendre, le système de prévision devra être à long, moyen ou court terme.

- ❖ **Long terme (3ans et plus)** : au niveau stratégique : diversification, lancement de produits nouveaux, investissements ou désinvestissement en équipements.
- ❖ **A moyen terme (6 mois à 2 ans)** : Planification globale et détaillé des capacités, acquérir une machine, embaucher du personnel...
- ❖ **Les prévisions à court terme** : serviront à l'activité opérationnelle de production : cycle de production, planification de la production et

¹ - CARNOT, N. et TISSOT B. (2002). *La prévision économique*, Economica, Paris, 483 p.

² - www.pearltrees.com/t/prevision-gestion

approvisionnement des matières et composants, charge des ateliers, ordonnancement des tâches.

2.1 Classification des techniques de prévision :

Les méthodes sont regroupées en catégories, de la façon suivante:

- les approches basées sur le jugement, ou informelles
- les méthodes extrapolatives ou univariées
- les méthodes explicatives ou causales
- les méthodes systémiques et économétriques¹.

2.1.1 Les méthodes informelles :

Les méthodes informelles ou de jugement sont très répandues dans le monde de l'entreprise. Plus généralement, elles sont particulièrement utiles dans toutes les applications caractérisées par une information quantitative déficiente (données non mesurables, peu fiables ou trop peu nombreuses) alors qu'un certain nombre de connaissances, d'informations qualitatives sont disponibles.

Principales méthodes informelles:

1. Réunions d'experts
2. Planification, politique de prix (évolution des ventes, du budget promotionnel), **horizon:** CT, MT, LT
3. Confrontation des forces des ventes (sales force composite). Exemple : évolution des ventes pour l'ensemble d'une firme qui vend différents produits). **horizon:** CT ou MT
4. Développements de scénarios. Ex: déchets produits par une firme, une région. **horizon:** MT ou LT
5. Approche Delphi. Variables qualitatives, planification, politique de prix. **horizon:** CT MT LT

¹ - PASTEELS, Jean-Michel (1997). Extrait de la thèse de doctorat; "L'expertise dans la prévision à court terme de variables économiques: contributions méthodologiques et empiriques". Université Libre de Bruxelles.

2.1.2 Les méthodes extrapolatives :

Les méthodes extrapolatives utilisent les observations quantitatives du passé de la variable pour prédire son futur. Autrement dit: $\omega_t = (Y_0, \dots, Y_{t-2}, Y_{t-1}, Y_t)$. Elles sont utilisées principalement pour la prévision à court terme ainsi que lorsque des variables explicatives ne sont pas disponibles ou manquent de fiabilité. Elles permettent notamment de modéliser l'inertie propre à de nombreuses variables économiques. Des informations contextuelles concernant le phénomène étudié permettent en général d'améliorer l'application des méthodes extrapolatives.

Principales méthodes extrapolatives ou univariées:

1. Prévision naïve. Phénomène incertain à prévoir (ex: prévision du cours d'une action), **horizon** : TCT
2. Lissage exponentiel (simple, double, adaptatif, amorti, etc.). Séries courtes, de nature industrielle, microéconomique, de fréquence mensuelle ou trimestrielle (ex: production du secteur textile au cours des trois prochains mois), **horizon** : TCT / CT
3. Courbes de croissance (régression linéaire, logistique, Gompertz, etc.). Séries annuelles, peu cycliques, assez régulières (ex: cycle de vie d'un produit). **horizon**: MT / LT
4. Décomposition saisonnière (moyennes mobiles, X-11, X-11 ARIMA, X-12, etc.). Séries trimestrielles et surtout mensuelles qui présentent d'importantes fluctuations saisonnières (ex: consommation de biens alimentaires). **horizon** : CT
5. Modélisation ARIMA (méthodologie de Box et Jenkins, etc.). Séries longues (ex: production industrielle trimestrielle). **horizon**:CT
6. Décomposition spectrale (composantes non observées). Séries longues (ex: production industrielle trimestrielle). **horizon**: CT
7. Modèles ARFIMA, mémoire longue. Séries très longues, séries annuelles (ex: évolution des prix pétroliers). **horizon**: MT / LT
8. Modèles stochastiques non linéaires (à seuil, bilinéaire, à erreurs GARCH, à coefficients dépendant du temps, etc.). Séries longues, de nature financière (données journalières, en temps réel), séries économiques volatiles (variables monétaires) (ex: taux de change \$/DM, taux d'intérêt). **horizon**: TCT / CT

2.1.3 Les méthodes explicatives :

Les méthodes explicatives utilisent les valeurs passées d'une ou de k variables, y compris, éventuellement, la variable étudiée Y (appelée variable dépendante). Ceci implique que la qualité des prévisions obtenues à l'aide de ces méthodes dépend entre autres de la qualité de prévision ou d'estimation des variables explicatives sur l'horizon de prévision. Cette catégorie de méthodes est particulièrement recommandée lorsqu'il existe des variables explicatives dont les observations sont disponibles plus rapidement que la variable dépendante. Pour la prévision conjoncturelle, on se base notamment sur ce qu'on appelle des indicateurs avancés ("leading indicators") de l'économie.

Principales méthodes explicatives:

1. Régression multiple. Planification (ex: prévision des ventes en fonction de la politique de prix et du budget consacré à la promotion). **horizon:** CT / MT / LT
2. Modèles de fonction de transfert. Séries longues (ex: prévision de l'activité en fonction d'indicateurs avancés). **horizon :** CT MT
3. Modèles à correction d'erreurs, modèles VAR, VARMA. Séries longues et variables économiques qui interagissent toutes entre elles (ex: variables macro-économiques, consommation privée et produit intérieur brut). **horizon:** CT / MT
4. Modèles dynamiques linéaires bayésiens. Séries longues pour lesquelles on dispose des informations à priori (ex: ventes de la société). **horizon:** CT

2.1.4 Les méthodes systémiques :

Les méthodes systémiques tendent à considérer le système économique dans sa totalité. Elles utilisent les relations et les interactions possibles entre de nombreuses variables. Ces méthodes sont utilisées pour la modélisation macro-économique et méso-économique principalement. Elles servent d'instrument de simulation (planification, politique budgétaire, politique monétaire) et de prévision à moyen et long terme.

Principales méthodes systémiques:

1. Equations simultanées (double moindres carrés, triple moindres carrés, etc.). Quand la distinction entre variables exogènes et endogènes est possible (ex: offre et demande en fonction du prix d'un bien). **horizon:** CT / MT / LT

2. Modèles macroéconomiques. Variables macroéconomiques pour lesquelles il existe un cadre théorique bien défini (équilibre général).
horizon:LT¹

1.2 Consommation d'électricité :

1.2.1 Définition de la consommation :

La **consommation**, au sens économique du terme, c'est l'action **d'utiliser** ou de **détruire**, immédiatement ou progressivement, des **biens** et des **services**, dans le but de satisfaire un besoin².

En ce qui concerne la consommation d'électricité, Contrairement à d'autres formes d'énergie, l'énergie électrique ne peut généralement pas être stockée telle quelle à grande échelle, chaque kilowatt-heure devant donc être produit au moment de sa consommation. Ce délicat équilibre entre l'offre et la demande doit être maintenu en tout temps pour assurer la fiabilité du service électrique.

1.2.2 Les types de la consommation d'électricité:

Plusieurs secteurs consomment de l'électricité et de la chaleur :

- ☞ le secteur de la transformation, et l'industrie de l'énergie au sein du secteur de l'énergie ;
- ☞ le transport et la distribution d'électricité et de chaleur ;
- ☞ les différents secteurs et branches de la consommation finale (industrie, transport, résidentiel, services, etc.).

1.2.3 la consommation finale d'électricité :

La consommation finale d'électricité et de chaleur est la somme d'électricité et de chaleur consommée dans les secteurs de l'industrie, du transport, de l'agriculture, des services commerciaux/publics et dans le secteur résidentiel. Le relevé de données relatives à l'usage non énergétique d'électricité et de chaleur n'est pas prévu.³

¹ - [fr.wikipedia.org/wiki/Prévision économique](https://fr.wikipedia.org/wiki/Prévision_économique).

² - *Société de consommation : société d'un pays industriel avancé qui crée sans cesse des besoins artificiels*
1992 P 14

³ Agence international de l'énergie manuel sur les statistiques de l'énergie page 48-53

1.3 Présentation de la SONELGAZ :

PRESENTATION DE L'ENTREPRISE MERE

Groupe SONELGAZ Spa - Secteur d'activité, Energie.

PRESENTATION DE L'ENTREPRISE MERE

Groupe SONELGAZ Spa - Secteur d'activité, Energie.

1.2.1 Historique:

Plus de soixante ans d'existence, l'organisme est passé par les étapes suivantes :

- Electricité et Gaz d'Algérie, EGA par abréviation (loi de la nationalisation de 1947):Un établissement public qui détient le monopole de la production, du transport et de la distribution de l'électricité et du gaz.
-
- Société Nationale de l'Electricité et du Gaz, Sonelgaz par abréviation (ordonnance de 1969) EGA devient Sonelgaz. Elle a les capacités organisationnelles et gestionnaires pour accompagner et soutenir le développement économique du pays.
- Société Nationale de l'Electricité et du Gaz : En 1983 les unités Sonelgaz travaux et de fabrication matériel, en l'occurrence : KAHRIF, KANAGHAZ, INERGA, ETTERKIB, KAHRAKIB, et AMC sont transformées en entreprises autonomes.
- Sonelgaz EPIC (décret 1991) : La Sonelgaz change de nature juridique et devient Etablissement Public à caractère Industriel et Commercial. Elle est placée sous tutelle du Ministre de l'Energie et des Mines.
- Création de la Sonelgaz Spa (loi sur l'électricité et la distribution du gaz par canalisations de Février 2002) : Cet engagement confère à la Sonelgaz la possibilité d'élargir ses activités à d'autres domaines relevant du secteur de l'énergie et aussi d'intervenir au niveau international.

Dans la mise en œuvre de cette loi, le Groupe Sonelgaz s'est engagé en groupe industriel avec la création de plusieurs sociétés filiales à savoir:

- 2004 : Création des filiales production SPE (Société de production d'électricité) et des filiales transport GRTE/GRTG (gestionnaire du réseau travaux électricité/Gaz),
- 2005 : Mise en place de la CREG (Commission de Régularisation d'Electricité et du Gaz),
- 2006 : Création des filiales Distribution Ouest, Centre, Alger, Est,
- 2006 : Mise en place de l'opérateur Système (OS),
- 2006 : Réintégration des cinq sociétés de travaux.
- 2008 : Mise en place de l'opérateur marché (OM).
- Travaux en cours : Maison mère en cours de réorganisation des fonctions juridique, communication et HSE (Environnement, Santé et sécurité). ¹

¹ WWW.KHRIF.DZ

1.3.2 Macro structure de la DR-SAÏDA:

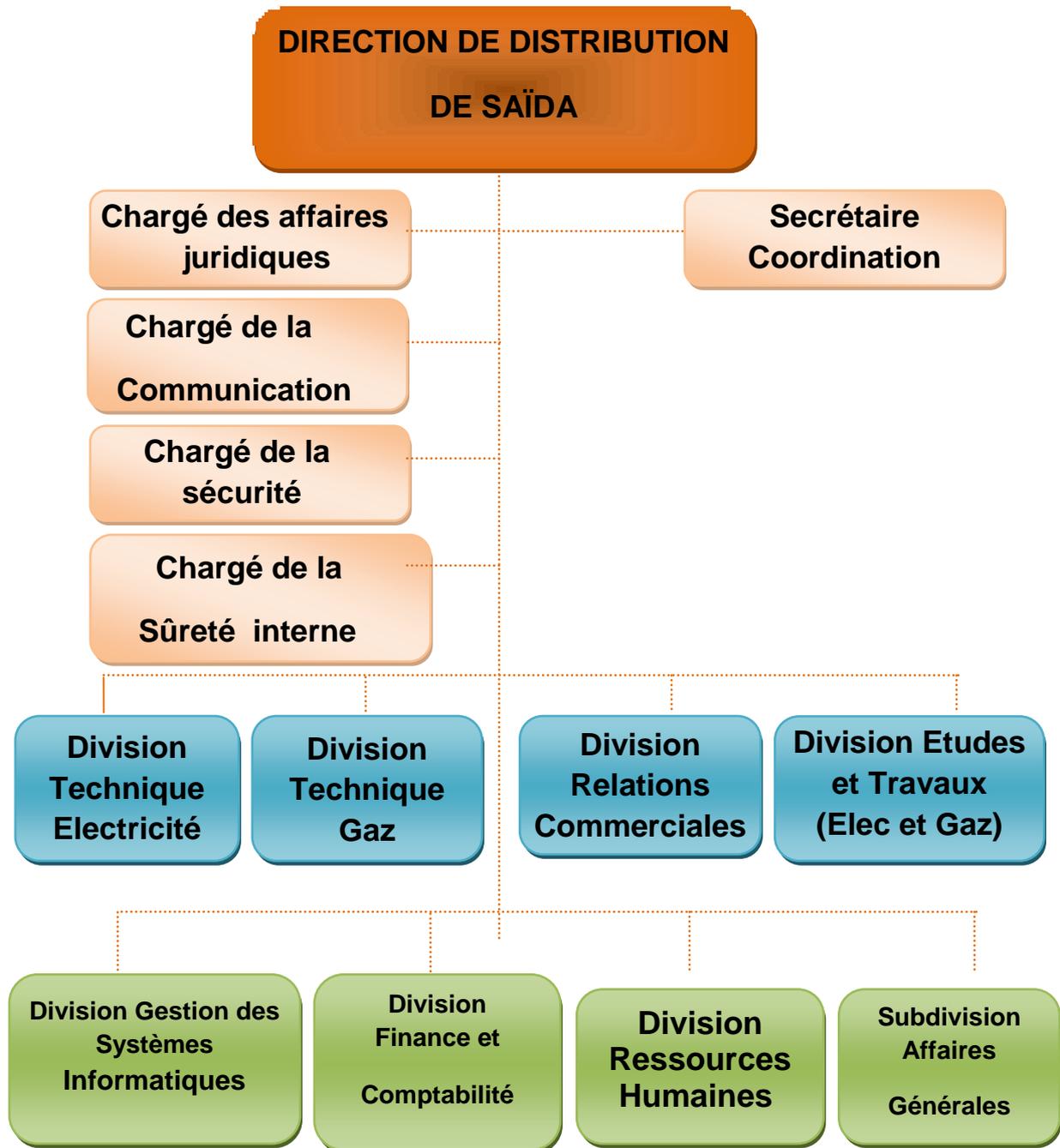


Schéma (1.1) : Macro structure de la DR-SAÏDA

SOURCE: Société de distribution de l'électricité et du gaz/unité
SAIDA

1.4 Présentation de la direction Analyse et Prévision (DAP) :

La DAP est une sous-direction de la Direction Générale Développement et Stratégie (DGDS)

1.4.1 Principales activités de la DAP :

Statistiques :

- Collecte des données traitement des données Publication (bulletin Statistique, bilan énergétique, rapports mensuels,...)

Analyse de la demande :

- Classification de la clientèle selon la forme de la courbe de charge :
- Panels Moyenne Tension (MT) / Basse Tension (ÂÔ)
- Clientèle Haute Tension (HT),
- Modélisation macro-économique de la demande énergétique. - Analyse de la clientèle par commune / wilaya.

Prévision de la demande :

Prévision de la demande énergétique : une stratégie en trois piliers

1er pilier : Fiabilité des données

- Segmentation et connaissance de la clientèle,
- Optimisation de la segmentation des données selon les directions régionales de distribution,
- Meilleure organisation des données par type, nature,... à l'échelle régionale et nationale.
- Plus de coordination entre les structures régionales des filiales de distribution et de la DAP/DGDS (suite à la restructuration de l'entreprise)

2ème pilier : Outils de calculs performants

- Le choix de la méthode de prévision est décisif.
- Il dépend essentiellement de la qualité des données.
- Outil de calcul convivial (simple d'utilisation, dynamique, extensible,...).

3ème pilier : Efficacité de la ressource humaine

- Formation continue de la ressource humaine sur les outils de prévision de la demande énergétique.
- Préparation d'une ressource humaine compétitive par rapport à celle de concurrents potentiels.
- Encouragement et promotion de la ressource humaine¹.

¹ - www.sonelgaz.dz

Conclusion :

Nous avons vu dans un premier temps une généralité sur la prévision (définition, les différentes techniques de prévision ...)

Dans un second temps nous avons offert une explication sur la consommation d'électricité. Puis nous avons donné une présentation de la SONELGAZ et enfin nous avons présenté la DAP et ses principales activités.

Ce chapitre est consacré à une présentation sommaire des techniques d'analyse des séries chronologiques. Nous étudions les caractéristiques statistiques-en terme de stationnarité- des séries temporelles en présentant les différents tests. Puis nous présentons différentes classes de modèles (MA, AR, ARMA) en étudiant leurs propriétés. Enfin la méthode de BOX et Jenkins qui systématise une démarche d'analyse des séries temporelles.

2.1 concepts de base sur les séries chronologiques :

2.1.1 Définition d'une série chronologique :

- On appelle série chronologique ou chronique une suite (Y_t)

d'observations chiffrées d'un même phénomène, ordonnées dans le temps.

Notation : Une série chronologique est aussi appelée série temporelle ou chronique.

Exemples :

A. Nombre mensuel de vente de voitures neuves en France.

B. Nombre annuel de naissance en Algérie.¹

- On appelle série chronologique (série temporelle, chronique) une suite d'observations numériques d'une grandeur effectuées a intervalles réguliers au cours du temps.

L'échelle de mesure et la variabilité de la grandeur sont telles que celle-ci sera toujours représentée par une variable continue à valeurs réelles.

En général l'unité est choisie de sorte que la variable s'exprime avec 2 ou 3 chiffres significatifs.

La fréquence des observations peut être journalière, hebdomadaire, mensuelle, trimestrielle, annuelle ou autre.²

¹ Alain Pirotte, L'économétrie : Des origines aux développements récents, Paris, CNRS, 22 juin 2004, p242

² University of Bath, UK Department of Mathematical Sciences University of Bath, UK © 2000 by Chapman & C Hall/CRC p 20 consulter au site:

[http://ccoms-imsuerj.org.br/capfts/wp-content/uploads/2011/08/Chatfield_Ch_Tim_Series_Forecasting\(2000\)\(1st%20edition\)\(280\).pdf](http://ccoms-imsuerj.org.br/capfts/wp-content/uploads/2011/08/Chatfield_Ch_Tim_Series_Forecasting(2000)(1st%20edition)(280).pdf)

2.1.2 Les composantes d'une série chronologique :

Le but de la décomposition d'une série chronologique est de distinguer dans l'évolution de la série, une tendance « générale », des variations saisonnières qui se répètent chaque année, et des variations accidentelles imprévisibles.

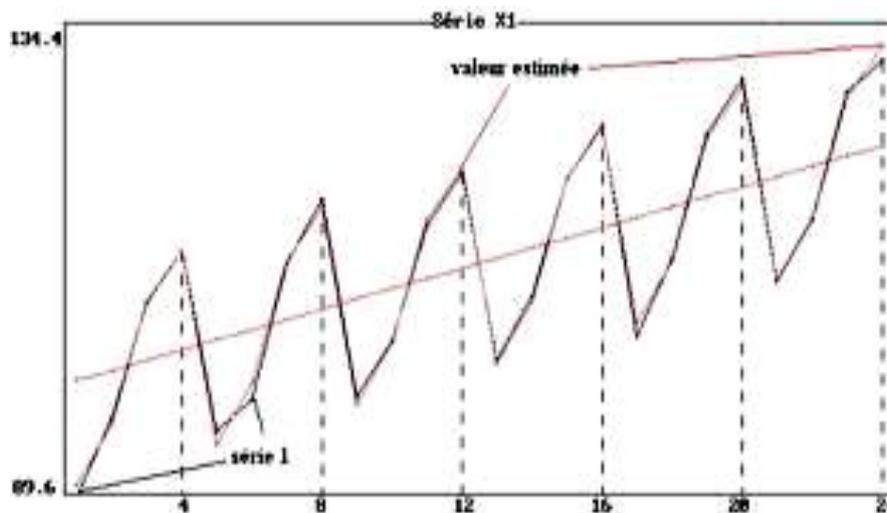
L'intérêt de ceci est d'une part de mieux comprendre, de mieux décrire l'évolution de la série, et d'autre part de prévoir son évolution (à partir de la tendance et des variations saisonnières).¹

1/La tendance : Tt

Ou « trend » notée Tt, censée décrire le mouvement de long terme, de fond ou encore structurel du phénomène. Ce mouvement est traditionnellement représenté par des formes analytiques simples: polynomiales, logarithmiques, exponentielles, cycliques, logistiques.²

Exemple : augmentation du chiffre d'affaire de 1978 à 1982.

Figure (2-1) :représente la tendance



¹ - نصيب رجم-الإحصاء التطبيقي، 2004 دار العلوم للنشر و التوزيع-عنايه ص 139

² - analyse des séries temporelles applications à l'économie et à la gestion , Régis bourbonnais, michel terraza page 6

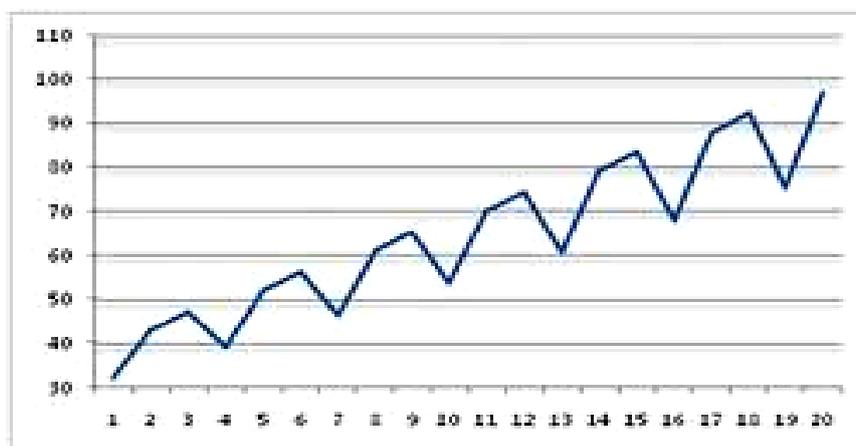
2/La composante cyclique : Ct

Notée Ct qui regroupe des variations à période moins précise autour de la tendance. Ces phases durent généralement plusieurs années, mais n'ont pas de durée fixe. Sans informations spécifiques, il est généralement très difficile de dissocier la tendance du cycle. Dans la plupart des travaux sur les séries temporelles la tendance regroupe aussi la composante cyclique.¹

3/Les variations saisonnières : St.

Ou variations saisonnières notées St sont des variations se reproduisant périodiquement à des moments bien déterminés et qui sont liées au rythme imposé par les variations météorologiques des saisons (production agricole, consommation de gaz, . . .), ou encore par des activités économiques et sociales (fêtes, vacances, solde, le ramadhan, etc.).²

Figure (2-2) : représente la composante saisonnalité



¹ - الأساليب الكمية التطبيقية في إدارة الأعمال المؤلف العلمي الثلاثي، د-عبد الحميد عبد المجيد البلداوي، د-نجم عبد الله الحميدي دار وائل للنشر 2008 ص 228

² - Régis Bourbonnais - Prévision des ventes P 22, Polycopié du produit multimédia, Université de Paris-Dauphine, OCTOBRE 2001

4/ Les variations accidentelles ou résiduelles : Rt.

Notée Rt. Elle rassemble tout ce que les autres composantes n'ont pu expliquer du phénomène observé. Elle contient donc de nombreuses fluctuations, en particuliers accidentelles, dont le caractère est exceptionnel et imprévisible, (catastrophes naturelles, grèves, guerres...). Comme par hypothèse ce type d'événement est censé être corrigé, le résidu présente en général une allure aléatoire plus ou moins stable autour de sa moyenne.¹

Exemple : Les tempêtes de décembre 1999 seront à l'origine de variations accidentelles dans la série chronologique mensuelle de consommation d'électricité.

Figure (2-3):représente la composane résiduelle



Source :

1 مولود حشمان السلاسل الزمنية و تقنيات التنبؤ القصير المدى

ديوان المطبوعات الجامعية الطبعة الثالثة 2010 ص 29-30

¹ - Régis Bourbonnais opc - p 33.

2.1.3 Schémas de décomposition d'une série chronologique :

La technique de décomposition d'une série chronologique, repose sur un modèle qui l'autorise. Ce modèle porte le nom de schéma de décomposition.

Il en existe essentiellement trois grands types :

1. Le modèle additif :

Dans un modèle additif, on suppose que les 3 composantes : tendance, variations saisonnières et variations accidentelles sont indépendantes les unes des autres.

On considère que la série X_t s'écrit comme la somme de ces 3 composantes :

$$X_t = E_t + S_t + R_t [1]$$

2. Le modèle multiplicatif :

1° forme de modèle multiplicatif:

On suppose que les variations saisonnières dépendent de la tendance.

Et on considère que X_t s'écrit de la manière suivante :

$$Y_t = E_t \times S_t + R_t [1]$$

2° forme de modèle multiplicatif : complet

On suppose que les variations saisonnières et les variations accidentelles dépendent de la tendance.

Et on considère que X_t s'écrit de la manière suivante :

$$X_t = E_t \times S_t \times R_t [1]$$

Remarque :

- Dans le cas d'une série (X_t) à valeurs positives, ce 2e modèle multiplicatif se ramène à un modèle additif en considérant la série ($\ln(X_t)$) :

$$\ln(X_t) = \ln(E_t) + \ln(S_t) + \ln(R_t).$$

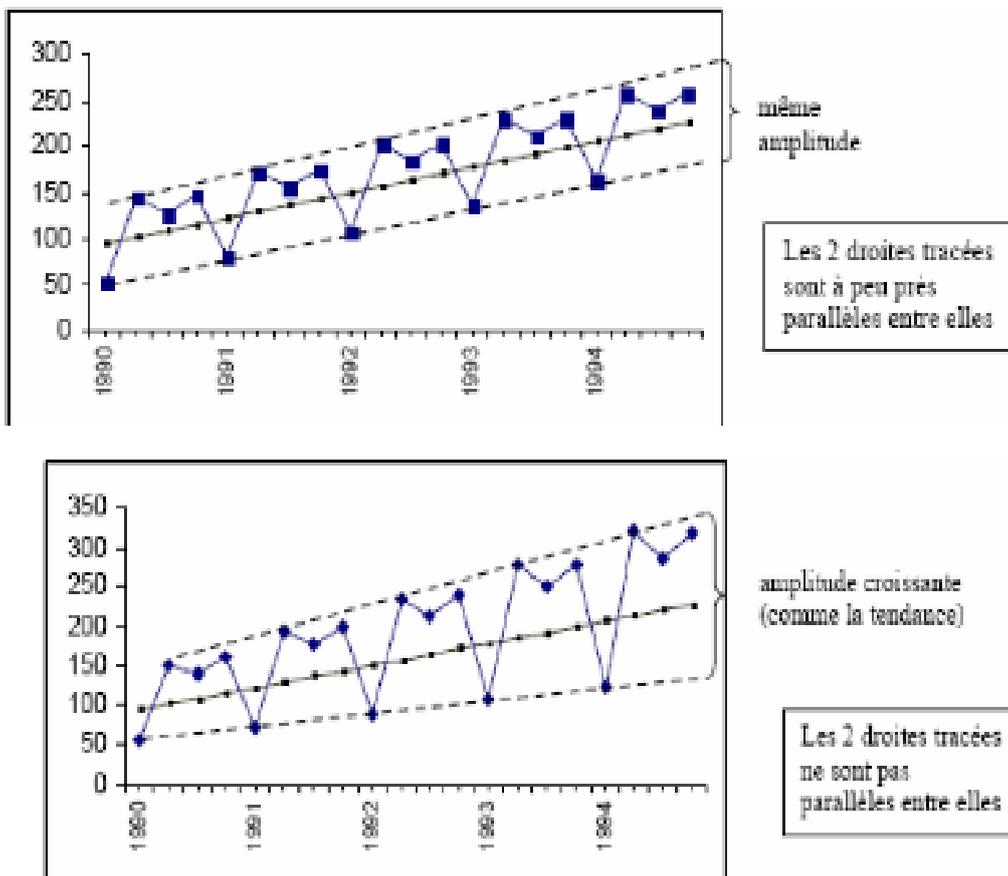
➤ La seule différence entre les 2 modèles multiplicatifs est dans l'estimation des R_t , qui n'a pas une grande importance.¹

2.1.3.1 Procédure de choix d'un schéma de décomposition :

*** la procédure de la bande :**

La procédure de la bande consiste a partir de l'examen visuel du graphique de l'évolution de la série brute à relier, par une ligne brisée, toutes les valeurs « hautes » et toutes les valeurs « basses » de la chronique. Si les deux lignes sont parallèles, la décomposition de la chronique peut se faire selon un schéma additif ; dans le cas contraire le schéma multiplicatif semble plus adapté.

Figure 2.4 : Exemple de schéma additif



¹ - analyse des séries temporelles applications à l'économie et à la gestion, Régis bourbonnais, Michel Terraza page 6-7

SOURCE :ANALYSE DES SERIE TEMPORELLES Régis bourbonnais, Michel Terraza P 20-21

2.1.4 Détections des composantes des séries temporelles :

Avant de faire l'analyse des séries temporelles on va tous d'abord détecter s'il ya ces composantes dans la série ou non et parmi les méthodes de détection :

2.1.4.1 Analyse des graphes :

Les composantes de la série chronologique peuvent être révélés par l'analyse de données graphiques, dans le cas de la composante de tendance elle pousse avec une courbe d'évolution de la série à travers le temps, en hausse si la tendance est positive, ou en baisse si elle est négative. Quant à la variable saisonnière elle se révèle par la régularité dans l'enregistrement du pic de la dernière saison de chaque année ou par une baisse sensible au début de la nouvelle année, par rapport à la composante cyclique se représente dans le graphe avec des augmentations ou réductions régulières permettant de déterminer la période de l'apparition de ce phénomène , que ce soit dans une saison ou un mois donné, dans le cas de la composante résiduelle est représenté dans ces fluctuations au niveau de la série et qui perturbe les composants réguliers leur donnant une forme aléatoire.¹

2.1.4.2 Tests statistiques :

Dans la plupart des cas, le test graphique est insuffisant pour bien examiner les composants de la série ce qui nécessite l'utilisation d'autres instruments statistiques (tests) pour détecté ces composantes ; nous proposons deux tests dont le premier déterminera l'existence de la composante de tendance et le deuxième déterminera l'existence de la composante saisonnier.²

2.1.4.2.1 Détection de la tendance :

*** Les tests non paramétrique :**

Les **tests non paramétriques** ne font aucune hypothèse sur la distribution sous-jacente des données. On les qualifie souvent de tests *distribution free*.

¹مولود حشمان السلاسل الزمنية وتقنيات التنبؤ قصير المدى ديوان المطبوعات الجامعية الطبعة الثالثة 2010 ص 29

² المرجع نفسه ص 30

***Le test de Daniel:**

Considérons ce test le plus fort parmi les autres tests non paramétriques il conviens par le coefficient de corrélation de Spirman pour mesurer la corrélation linéaire entre deux ordres : ordre R_t et temporelle t ou :

$$t=1.2 \rightarrow T, R_t = 1.T \text{ Et } R_t = f(t)$$

Connaissant le coefficient de corrélation définir par :

$$r_s = \frac{COV(R_t, t)}{\sqrt{Var(R_t)Var(t)}} = \frac{\sum_{t=1}^T (R_t - \bar{R})(\bar{t} - t)}{\sum_{t=1}^T (t - \bar{t})^2} = 1 - \frac{6 \sum_{t=1}^T d_t^2}{t(t^2-1)}$$

Ou :

d_t : La différence entre l'ordre ascensionnel et temporelle ce que veut dire :

$$d_t = R_t - t \text{ et } r_s = \{-1, 1\}$$

Hypothèse de test :

$$\begin{cases} H_0: \text{série aleatoire (il ya pas de tendance)} \\ H_1: \text{série non aleatoire (il existe une tendance)} \end{cases}$$

Décision:

Après trouver le coefficient de Spirman r_s on rejette l'hypothèse nulle avec prendre en valeur la proportion d'échantillon :

1) Proportion d'échantillon inférieur :

$$|r_s| > r_{\alpha/2}, T \ll 30$$

2) Proportion d'échantillon inférieur :

$$|z| > z_{\alpha/2}, T > 30$$

$$\text{Ou : } r_s = \frac{1}{\sqrt{T-1}} \leq z = \frac{r_s - \mu_{r_s}}{\delta_{r_s}} \quad \mu_{r_s} = 0; \delta_{r_s} = \frac{1}{\sqrt{T-1}}$$

$$Z = \frac{r_s}{\delta_{r_s}} = r_s \sqrt{T-1}$$

Tests paramétrique:

Cette méthode suppose une existence de la composante de tendance dans la série ainsi que la composante aléatoire et elle suppose également la connaissance de la distribution de probabilité des erreurs se que signifie : $\mu_t \sim (0; \delta^2)$ ou

$$y_t = f(t, \mu_t)$$

Et après déterminer la forme de la fonction $f(t, \mu_t)$ en estime ses paramètres par utiliser la statistique de Student ou la variance standard.¹

2.1.4.2.2 Detection de la saisonnalité :

***Tests non paramétrique :**

****test de Kruskall Wallis:**

On applique ce test après retirer la tendance

Hypothèse de test :

$$\begin{cases} H_0: \text{il ya pas de saisonnalité} \\ H_1: \text{il ya une saisonnalité} \end{cases}$$

On va tester ces hypothèses par la relation suivante :

$$KW = \frac{12}{T(T+1)} \sum_{i=1}^P \frac{R_i^2}{n_i} - 3(T+1) \rightarrow X^2(p-1)$$

R_i : la somme des ordre d'observation conforme a i

n : nombre d'observation conforme a la saison i

p : la circulation égale a quatre dans le cas ou le nombre d'observation trimestrielle, douze dans le cas mensuelle

Si N_i supérieur de 5 et l'hypothèse nulle n'est juste alors KW peut suivre la distribution χ^2 avec un degré de $(p-1)$

Décision :

En rejette H_0 si $KW > \chi_{p-1}^2$

¹ ص i مولود حشمان مرجع سبق ذكره 64

Test paramétrique:***le test de Buys-Ballot :**

On calcule, pour chacune des années, la moyenne et l'écart type, puis on estime par MCO les paramètres a_1 et a_0 de l'équation $\sigma_i = a_1 \bar{x}_i + a_0 + \varepsilon_i$.

Dans le cas, où le paramètre a_1 n'est pas significativement différent de 0 (test de Student) alors on accepte l'hypothèse d'un schéma additif ; dans le cas contraire, nous retenons un schéma multiplicatif.¹

2.2 Elément d'analyse des séries temporelles:**2.2.1 Processus stochastique :**

Un processus aléatoire est une application X qui associe au couple (ω, t) la quantité $x_t(\omega)$. Elle est telle que $\forall t \in T$ fixé, x_t est une variable aléatoire définie sur un espace probabilisé. Un processus stochastique est donc une famille de variables aléatoires indicées par t noté $(x_t, t \in T)$ ou encore x_t .

Dans la suite de l'exposé, l'espace des indices T est le temps, t est alors l'instant d'observation de la variable aléatoire x sur l'individu. Si T est l'ensemble des réels, le processus est dit continu. Si $T = Z$ ou N ou N^* ou N^* borné, le processus est dit discret.

On suppose, par la suite que la série temporelle notée x_t $t \in N^*$ borné (soit une succession d'observations régulièrement espacées dans le temps d'une valeur économique) est une réalisation d'un processus stochastique discret univarié².

2.2.2 Stationnarité :**2.2.2.1 Définition et propriétés :**

Avant le traitement d'une série chronologique, il convient d'en étudier les caractéristiques. Si ces caractéristiques c'est-à-dire son espérance et sa variance se trouvent modifiées dans le temps, la série chronologique est considérée comme non stationnaire ; dans le cas d'un processus stochastique invariant, la série temporelle est alors stationnaire. de manière formalisée, le processus stochastique y_t est stationnaire si :

¹ Régis Bourbonnais. Michel T OPC p 21

² Régis Bourbonnais. Michel T OPC p73

$E(y_t) = E(y_{t+m}) = \mu \forall t \text{ et } \forall m$, la moyenne est constante et indépendante de temps ;

$var(y_t) < \infty \forall t$, la variance est finie et indépendante de temps ;

$cov(y_t, y_{t+k}) = E[(y_t - \mu)(y_{t+k} - \mu)] = \gamma_k$, la covariance est indépendante de temps.

Il apparaît, à partir de ces propriétés, qu'un processus de bruit blanc ε_t dans lequel les ε_t sont indépendants et de même loi $N(0, \sigma_\varepsilon^2)$ est stationnaire.

Une série chronologique est donc stationnaire si elle est la réalisation d'un processus stationnaire. Ceci implique que la série ne comporte ni tendance, ni saisonnalité et plus généralement aucun facteur n'évoque avec le temps.¹

2.2.2.2 Fonction d'autocorrelation simple et partielle :

La fonction d'autocorrelation FAC est la fonction notée p_k qui mesure la corrélation de la série avec elle-même décalée de k période, sa formulation est la suivante :

$$p_k = \frac{cov(y_t, y_{t-k})}{\sigma_{y_t} \sigma_{y_{t-k}}} = \frac{\sum_{t=k+1}^n (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sqrt{\sum_{t=k+1}^n (y_t - \bar{y})^2} \sqrt{\sum_{t=k+1}^n (y_{t-k} - \bar{y})^2}} \quad [1]$$

Avec \bar{y} moyenne de la série calculée sur n-k périodes, n=nombre d'observations.

Nous pouvons en déduire que :

$$p_0 = 1 \text{ et } p_k = p_{-k}$$

Cette formule [1] est malaisée à manier puisqu'elle exige de recalculer pour chaque terme p_k les moyennes et les variances, c'est pour quoi on lui préfère la fonction d'autocorrelation d'échantionnage :

$$\widehat{p}_k = \frac{\sum_{t=k+1}^n (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2} \quad [2]$$

Avec \bar{y} moyenne de la série calculée sur n périodes.

2.2.2.3 Tests de « bruit blanc » et de stationnarité :

¹ Regie Bourbonnais-Econométrie-Manuel et exercices corrigés. 6édition 2007- p 224.

Nous nous pouvons identifier clairement les caractéristiques stochastiques d'une série chronologique que si elle est stationnaire. Cette étude de stationnarité s'effectue essentiellement à partir de l'étude des fonctions d'autocorrelation (ou de leur représentation graphique appelée « corrélogramme »). Une série chronologique est stationnaire si elle ne comporte ni tendance ni saisonnalité.

Nous pouvons distinguer différents types de séries stationnaires :

- ☞ A mémoire, c'est-à-dire dont on peut modéliser, par une loi de reproduction, le processus ;
- ☞ Identiquement et indépendamment distribuée notée *i.i.d* ou appelée Bruit Blanc (« White Noise ») ;
- ☞ Normalement (selon une loi normale) et indépendamment distribuée notée *n.i.d* ou appelée Bruit Blanc gaussien.¹

2.2.2.3.1 Analyse des fonctions d'autocorrelation :

Lorsque nous étudions la fonction d'autocorrélation d'une série chronologique, la question qui se pose est de savoir quels sont les termes p_k qui sont significativement différents de 0.

Le test d'hypothèses pour un terme p_k est le suivant :

$$H_0: p_k = 0$$

$$H_1: p_k \neq 0$$

Si la taille d'échantillon $n > 30$, le coefficient p_k tend de manière asymptotique vers une loi normale de moyenne 0 et d'écart type $\frac{1}{\sqrt{n}}$.

L'intervalle de confiance du coefficient p_k est alors donné par :

$$p_k = 0 \pm t^{\alpha/2} \frac{1}{\sqrt{n}}$$

N=nombre d'observation

la valeur calculer $\frac{\widehat{p}_k}{\sqrt{\frac{1}{n}}} \rightarrow n(0,1)$

¹ Regie Bourbonnais-Econométrie-Manuel et exercices corrigés. 6^e édition 2007- p 240

On compare avec la valeur tab de la loi normale au seuil de coefficient 5%

Si la $V_{cal} > V_{tab}$ on accepte H_0

2.2.2.3.2 Statistique de Box-Pierce et Ljung-Box :

Il est établi à partir de la statistique $Q = n \sum_{k=1}^k r_k^2(a_t)$ qui est fonction de la somme des carrés des autocorrélations $r_k^2(a_t)$ de la FAC et de nombre d'observations n. il permet de vérifier l'hypothèse :

$$H_0: p_1 = p_2 = \dots = p_k = 0$$

H_1 : il existe au moins un p_i significativement différent de 0.

Pour effectuer ce test, il est conseillé de choisir K proche du tiers du nombre d'observations. L'hypothèse H_0 est rejetée au seuil de 5% si Q est supérieur au quantile 0,95 de loi du x^2 .

Nous pouvons utiliser aussi une autre statistique, dont les propriétés asymptotiques sont meilleures, dérivée de la première qui est le \hat{Q}

de Ljung et Box : $\hat{Q} = n(n+2) \sum_{n-k}^h \frac{\widehat{p_k^2}}{n-k}$ qui est aussi distribuée selon un x^2 à h degrés de liberté et dont les règles de décisions sont identiques au précédent.¹

2.2.2.3.3 Les tests de normalité:

a) test de SKEWNESS et test de KURTOSIS :

Soit $\mu_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k$ le moment centré d'ordre k, le coefficient de Skewness ($\beta_1^{1/2}$) est égale à : $\beta_1^{1/2} = \frac{\mu_3}{\mu_2^{3/2}}$ et le coefficient de Kurtosis $\beta_2 = \frac{\mu_4}{\mu_2^2}$

Si la distribution est normale et le nombre d'observations grand :

$$\beta_1^{1/2} \rightarrow N\left(0; \sqrt{\frac{6}{n}}\right)$$

$$\beta_2 \rightarrow N\left(3; \sqrt{\frac{24}{n}}\right)$$

¹ - Regie Bourbonnais-Econometrie- OPC P227

On construit alors les statistiques :

$$v_1 = \frac{\beta_1^{1/2} - 0}{\sqrt{\frac{6}{n}}} \text{ et } v_2 = \frac{\beta_2 - 3}{\sqrt{\frac{24}{n}}} \text{ que l'on compare à } 1.96 \text{ au seuil de } 5\%.$$

Si les hypothèses $H_0: v_1 = 0$ (symétrie) et $v_2 = 0$ (aplatissement normal)

sont vérifiées alors $|v_1| < 1.96$ et $|v_2| < 1.96$, dans le cas contraire, l'hypothèse de normalité est rejetée.¹

b) Test de Jarque-Bera :

Il s'agit d'un test qui regroupe les résultats précédents, si $\beta_1^{1/2}$ et β_2 obéissent à des lois normales alors la qualité s :

$$s = \frac{n}{6}\beta_1 + (\beta_2 - 3)^2 \text{ suit un } \chi_{1-\alpha}^2(2) \text{ à } 2 \text{ degrés de liberté.}$$

Donc si $s \geq \chi_{1-\alpha}^2(2)$, on rejette l'hypothèse H_0 de normalité des résidus au seuil α^2

2.2.3 Les processus aléatoires non stationnaires:

2.2.3.1 Description des processus TS et DS :

Processus TS :

Définition :

Un processus TS s'écrit : $x_t = f_t + \varepsilon_t$ ou f_t est une fonction polynomiale du temps, linéaire ou non linéaire et ε_t un processus stationnaire de type ARMA.

Le processus TS le plus simple est représenté par une fonction polynomiale de degré 1. Ce processus s'écrit : $x_t = a_0 + a_1 t + \varepsilon_t$

¹ - Régis Bourbonnais et Michel Terraza, Analyse des séries temporelles, 2^e édition, Dunod, 2008, p. 237

² - Jarque, Carlos M. & Anil K. Bera (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. Economics Letters 6 (3): 255–259.

Si ε_t est un bruit blanc (gaussien ou non), les caractéristiques de ce processus sont alors :

$$E[x_t] = a_0 + a_1 t + E[\varepsilon_t] = a_0 + a_1 t$$

$$V[x_t] = 0 + V[\varepsilon_t] = \sigma_\varepsilon^2$$

$$COV[x_t, x_{t'}] = 0 \text{ pour } t \neq t'$$

La non stationnarité de ce processus est dû au fait que son espérance dépend du temps.

La méthode pour stationnariser un processus *TS* est d'estimer les coefficients a_0 , a_1 par MCO (Moindre Carrés Ordinaires) et de retrancher de la valeur de X_t en t la valeur estimée de sa moyenne $\widehat{a}_0 + \widehat{a}_1 t$.

Cours des séries temporelles théorie et applications volume 1¹.

Processus DS :

DS sans dérive :

Soit le processus DS sans dérive (AR(1)) :

$$x_t = x_{t-1} + \varepsilon_t$$

$$(1 - B)x_t = \beta + \varepsilon_t$$

La racine du polynôme caractéristique $(1 - B)$ est égale à 1. On dit que le processus x_t a une racine unité, il est donc non stationnaire.

Ce processus DS sans dérive peut se réécrire sous la forme :

$$x_t = x_{t-1} + \varepsilon_t$$

$$x_{t-1} = x_{t-2} + \varepsilon_{t-1} \rightarrow x_t = x_{t-2} + \varepsilon_{t-1} + \varepsilon_t$$

$$x_{t-2} = x_{t-3} + \varepsilon_{t-2} \rightarrow x_{t-3} + \varepsilon_{t-2} + \varepsilon_{t-1} + \varepsilon_t$$

Etc.

¹ Introduction à la théorie des processus en temps discret modèle ARIMA et méthode de box & Jenkins. ARTHURE CHARPENTIER UNIVERCIT2 parie d'AUPHINE P21

Si le premier terme de la chronique est x_0 le modèle s'écrit alors :

$$x_t = x_0 + \sum_{i=1}^t \varepsilon_i$$

Les caractéristiques de ce processus sont (on supposant x_0 certain) :

$$E[x_t] = x_0$$

$$V[x_t] = t\sigma_\varepsilon^2$$

$$\text{cov}[x_t, x_{t'}] = \sigma_\varepsilon^2 \times \text{MIN}(t, t') \text{ si } t \neq t'$$

Un processus DS sans dérive est un processus stationnaire en moyenne et non stationnaire en variance.

DS avec dérive :

Considérons un processus DS avec dérive

$\beta = 0$: le processus porte alors le nom de processus DS avec dérive[1]

S'écrit :

$$x_t = x_{t-1} + \beta + \varepsilon_t$$

Comme précédemment, on peut rechercher sa forme équivalente développée :

$$x_t = x_{t-1} + \beta + \varepsilon_t$$

$$x_{t-1} = x_{t-2} + \beta + \varepsilon_{t-1} \rightarrow x_t = x_{t-2} + 2\beta + \varepsilon_{t-1} + \varepsilon_t$$

$$x_{t-2} = x_{t-3} + \beta + \varepsilon_{t-2} \rightarrow x_{t-3} + 3\beta + \varepsilon_{t-2} + \varepsilon_{t-1} + \varepsilon_t$$

Etc.

Si on suppose la valeur d'origine x_0 connue et déterministe, on a alors :

$$x_t = x_0 + \beta t + \sum_{i=1}^t \varepsilon_i$$

On peut analyser les caractéristiques de ce processus :

$$E[x_t] = x_0 + \beta t$$

$$V[x_t] = t\sigma_\varepsilon^2$$

$$\text{cov}[x_t, x_{t'}] = \sigma_\varepsilon^2 \times \text{MIN}(t, t') \text{ si } t \neq t'$$

Un processus DS avec dérive est un processus **non stationnaire en moyenne et en variance**. Ces moments évoluent en fonction du temps t .

Un processus DS est un processus que l'on peut stationnariser par l'application du filtre aux différences :

$$x_t = x_{t-1} + \beta + \varepsilon_t \leftrightarrow (1 - \beta)x_t = \beta + \varepsilon_t^1$$

2.2.3.2 Les tests de racines unitaires et la stratégie séquentielle de test :

Les structures DS et TS jouent un rôle très important dans le traitement statistique d'une chronique. Comment choisir entre l'une ou l'autre des structures? Les tests de racine unitaire tentent de répondre à cette question.

2.2.3.2.1 Les tests de Dickey-Fuller simple (DF) :

Les tests proposés par DICKEY & FULLER (1979) ont pour but de vérifier la stationnarité de la série étudiée. Ils permettent de déceler le type de non stationnarité de la série.²

L'application du test se fait en estimant par la méthode des moindres carrés ordinaires MCO trois modèles suivant que le processus qui représente la série x_t contient ou non une constante et une tendance. Les modèles de bases sont :

Modèle 1 : $(1 - \phi_1 B)x_t = a_t$ modèle autorégressif d'ordre 1 : AR(1)

Modèle 2 : $(1 - \phi_1 B)(x_t - \mu) = a_t$ modèle AR(1) avec constante, ou $E[x_t] = \mu$

Modèle 3 : $(1 - \phi_1 B)(x_t - \alpha - \beta t) = a_t$ modèle AR(1) avec tendance et $a_t \rightarrow i. i. d. (0; \sigma_a^2)$

¹ - Régis Bourbonnais et Michel Terraza, Analyse des séries temporelles, 2^e édition, Dunod, 2008, p. 136

²- "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," Journal of the American Statistical Association, 74, p. 427-431.

Caractéristiques des modèles :

Modèle 1 : $(1 - \phi_1 B)x_t = a_t \leftrightarrow x_t = \phi_1 x_{t-1} + a_t$

Soit l'hypothèse $H_0: \phi_1 = 1$

Le modèle s'écrit alors sous cette hypothèse :

$$(1 - B)x_t = a_t \leftrightarrow x_t = x_0 + \sum_{i=1}^t a_i$$

Il s'agit d'un modèle DS « promenade aléatoire » dont les caractéristiques non stationnaires ont été décrites précédemment :

Si H_0 est acceptée alors le processus stationnarisé est : $\Delta x_t = a_t$

Si l'hypothèse alternative $H_1: |\phi_1| < 1$ est acceptée alors x_t est un AR(1) stationnaire.

Modèle 2 :

$$(1 - \phi_1 B)(x_t - \mu) = a_t$$

$$x_t = \phi_1 x_{t-1} + c + a_t \text{ ou } c = \mu(1 - \phi_1)$$

Sous $H_0: \phi_1 = 1$.

Le modèle s'écrit (avec $c = 0$ par conséquent) : $x_t = x_{t-1} + a_t$

C'est un modèle DS de promenade aléatoire (soit le modèle 1)

On le rend stationnaire par la transformation : $\Delta x_t = a_t$

Sous l'hypothèse alternative $H_1: |\phi_1| < 1$ on a un AR(1) stationnaire avec constante.

Modèle 3 :

$$(1 - \phi_1 B)(x_t - \alpha - \beta t) = a_t$$

$$x_t = \phi_1 x_{t-1} + bt + c + a_t$$

Avec $c = \alpha(1 - \phi_1) + \phi_1 B$ et $b = \beta(1 - \phi_1)$

Sous $H_0: \phi_1 = 1$

Alors $b = 0$ et $c = \beta$ et le modèle s'écrit: $x_t = x_{t-1} + \beta + a_t$

Il s'agit d'une promenade aléatoire avec dérive (processus DS avec dérive)

Avec l'hypothèse alternative $H_1: |\phi_1| < 1$, le modèle peut s'écrire en posant : $\varepsilon_t = (x_t - \alpha - \beta t) : (1 - \phi_1 B)\varepsilon_t = a_t$.¹

En résumé, le tableau ci-dessus synthétise la typologie des modèles

Tableau (2-1) typologie des modèles

	Modèle 1	Modèle 2	Modèle 3
H0	DS sans dérive	DS sans dérive	DS avec dérive
	Stationnaire à l'ordre 1	Stationnaire à l'ordre 1	Non stationnaire
	Non stationnaire à l'ordre 2	Non stationnaire à l'ordre 2	
H1	AR(1)	AR(1) avec constante	TS avec erreurs ARMA

Source : ANALYSE DES SERIES TEMPORELLES REGIS BOURBONNAIS ; MICHEL T

2.2.3.2.2 Les tests de Dickey-Fuller Augmentés (ADF) :

1/ Transformations des modèles de base :

Dans les modèles précédents, utilisés pour les tests de Dickey-Fuller simples, le processus est, par hypothèse, un bruit blanc. Or il n'y a aucune raison pour que, a priori, l'erreur soit non corrélée ; on appelle tests de Dickey-Fuller Augmentés (DFA, 1981) la prise en compte de cette hypothèse.

Les tests ADF sont fondés, sous l'hypothèse alternative $|\phi_1| < 1$, sur l'estimation par les MCO des modèles :

¹ - Régis Bourbonnais et Michel Terraza, Analyse des séries temporelles, 2^e édition, Dunod, 2008, p. 144.

$$\Delta x_t = \rho x_{t-1} - \sum_{j=2}^p \phi_j \Delta x_{t-j+1} + \varepsilon_t \dots (4)$$

$$\Delta x_t = \rho x_{t-1} - \sum_{j=2}^p \phi_j \Delta x_{t-j+1} + c + \varepsilon_t \dots (5)$$

$$\Delta x_t = \rho x_{t-1} - \sum_{j=2}^p \phi_j \Delta x_{t-j+1} + c + b_t + \varepsilon_t \dots (6)$$

La valeur p peut être déterminée selon les critères d'information de Akaike ou Schwarz, ou encore, en partant d'une valeur suffisamment importante de p , on estime un modèle à $p-1$ retards puis à $p-2$ retards

Jusqu'à ce que le coefficient du p^{eme} retard soit significatif¹.

2.2.3.2.3 Le test de Philips et Perron :

Le test de Phillips et Perron (1988) est construit sur une correction non paramétrique des

statistiques de Dickey-Fuller pour prendre en compte des erreurs

hétéroscédastiques et/ou autocorrélées. Il se déroule en quatre étapes :

- estimation par les moindres carrés ordinaires des trois modèles de base des tests de Dickey-Fuller et calcul des statistiques associées, soit et le résidu estimé ;
- estimation de la variance dite de court terme des résidus $\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^n e_t^2$;
- estimation d'un facteur correctif s_t^2 (appelé variance de long terme) établi à partir de la structure des covariances des résidus des modèles précédemment estimés de telle sorte que les transformations réalisées conduisent à des distributions identiques à celles du Dickey-Fuller standard :

$$s_t^2 = \frac{1}{n} \sum_{t=1}^n e_t^2 + 2 \sum_{i=1}^l \left(1 - \frac{i}{l+1}\right) \frac{1}{n} \sum_{t=i+1}^n e_t e_{t-i}$$

pour estimer cette variance de long terme, il est nécessaire de définir un nombre de retard l (troncature de Newey-West) estimé en fonction du nombre d'observations n , $l \approx 4(n/100)^{2/9}$;

- calcul de la statistique de $pp: t_{\phi_1}^* = \sqrt{k} \times \frac{\hat{\phi}_1 - 1}{\hat{\sigma}_{\phi_1}} + \frac{n(k-1)\hat{\sigma}_{\phi_1}}{\sqrt{k}}$ avec

$$k = \frac{\hat{\sigma}^2}{s_t^2} \text{ (qui est égal à 1- de manière asymptotique - si } e_t \text{ est un$$

¹- Régis Bourbonnais-ECONOMETRIE OPC P 148

bruit blanc). Cette statistique est à comparer aux valeurs critiques de la table de MacKinnon.

Il est à noter que les logiciels RATS et Eviews permettent directement l'utilisation de ces tests.¹

2.2.3.2.4 Stratégie de tests :

Nous constatons que pour réaliser un test de racine unitaire, le résultat n'est identique selon l'utilisation de l'un des trois modèles comme processus générateur de la chronique de départ. Les conclusions auxquelles on parvient sont donc différentes et peuvent entraîner des transformations erronées. C'est la raison pour laquelle Dickey-Fuller, et à leur suite d'autres auteurs, ont élaboré des stratégies de tests. Nous présentons un exemple simplifié (schéma 1) d'une stratégie de tests.

¹ - Davidson, Russell and James G. MacKinnon (2004), *Econometric Theory and Methods*, p.623, ISBN 978-0-19-512372-2

¹-Phillips, P.C.B and P. Perron (1988), "Testing for a Unit Root in Time Series Regression", *Biometrika*, 75, 335–346

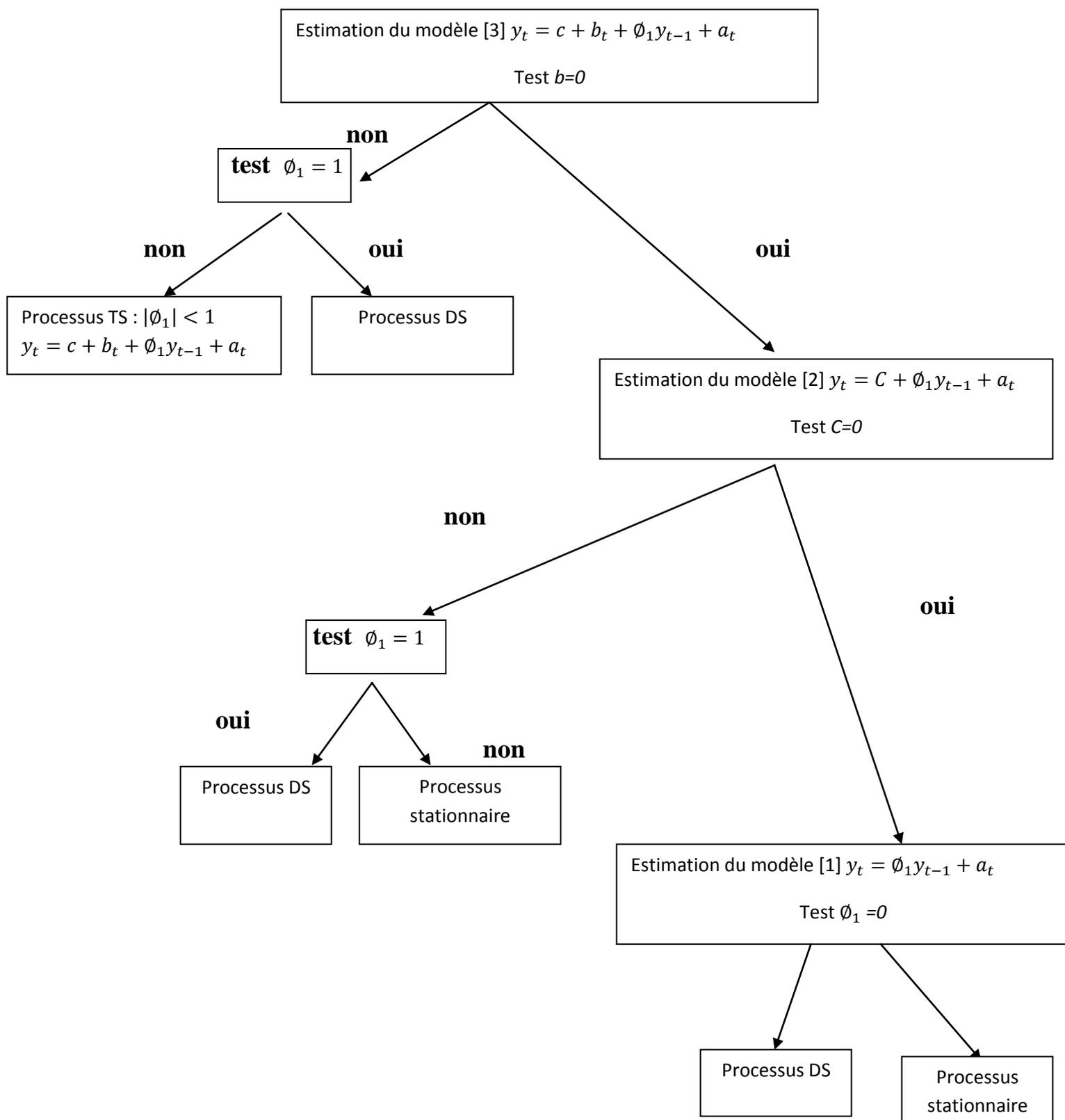


Schéma 2-1 : Stratégie simplifiée des tests de racine unitaire

Source : R.Bourbonnais ,op-cit, p 249

2.3 Processus linéaires des séries temporelles :

Dans cette partie, nous entamons l'étude détaillée de la partie aléatoire d'une série chronologique. Les processus linéaires constituent le modèle le Plus simple pour décrire cette composante.

Parmi ces processus on a les processus Moyenne Mobile, processus Autorégressive et les processus mixte .

2.3.1 Processus Moyenne Mobile (Moving Average) q : MA(q) ;

formulation :

dans le processus de moyenne mobile d'ordre q, chaque observation y_t est générée par une moyenne pondérée d'aléas jusqu'à la q-ième période.

$$MA(1) : y_t = \varepsilon_t - \alpha_1 \varepsilon_{t-1}$$

$$MA(2) : y_t = \varepsilon_t - \alpha_1 \varepsilon_{t-1} - \alpha_2 \varepsilon_{t-2}$$

$$\dots MA(q) : y_t = \varepsilon_t - \alpha_1 \varepsilon_{t-1} - \alpha_2 \varepsilon_{t-2} - \dots - \alpha_q \varepsilon_{t-q}$$

Où $\alpha_1, \alpha_2, \dots, \alpha_q$ sont des paramètres pouvant être positifs ou négatifs et ε_t est un aléa gaussien.

Dans ce processus, tout comme dans le modèle autorégressif AR, les aléas sont supposés être enregistrés par un processus de type **bruit blanc**. Nous pouvons interpréter le modèle MA comme étant représentatif d'une série chronologique fluctuant autour de sa moyenne de manière aléatoire, d'où le terme de moyenne mobile car celle-ci, en lissant la série, gomme le bruit créé par l'aléa.¹

Il est à noter qu'il ya équivalence entre un processus MA(1) et un processus AR d'ordre p infini :

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

2.3.1.1 Caractéristiques des corrélogrammes :

le corrélogramme simple d'un processus MA(q) est de la forme générale :

$$p_k = \frac{\sum_{i=0}^{i=q-k} \alpha_i \alpha_{i+k}}{\sum_{i=0}^{i=q} \alpha_i^2} \text{ pour } k=0,1, \dots, q \text{ et } p_k = 0 \text{ pour } k > q$$

¹ Regie Bourbonnais-Michel Terrazza OPC P197

C'est-à-dire que seuls les q premiers termes du corrélogramme simple sont significativement différents de 0.

Le corrélogramme partiel est caractérisé par une décroissance géométrique des retards.¹

2.3.2 Processus Auto Régressif d'ordre p : AR (p) :

a)formulation :

dans le processus autorégressif d'ordre p , l'observation présente y_t est générée par une moyenne pondérée des observations passées jusqu'à la p -ième période sous la forme suivante :

$$\text{AR}(1) : y_t = \theta_1 y_{t-1} + \varepsilon_t$$

$$\text{AR}(2) : y_t = \theta_1 y_{t-1} + \theta_2 y_{t-2} + \varepsilon_t$$

.....

$$\text{AR}(p) : y_t = \theta_1 y_{t-1} + \theta_2 y_{t-2} + \dots + \theta_p y_{t-p} + \varepsilon_t \quad [2]$$

Où $\theta_1, \theta_2, \dots, \theta_p$ sont des paramètres à estimer positifs ou négatifs, ε_t est un aléa gaussien.

Nous pouvons ajouter à ce processus une constante qui ne modifie en rien les propriétés stochastiques. L'équation [2] peut aussi s'écrire à l'aide de l'opérateur décalage D :

$$(1 - \theta_1 D - \theta_2 D^2 - \dots - \theta_p D^p) y_t = \varepsilon_t$$

2.3.2.1 Caractéristiques des processus AR(p) :

Conditions de stationnarité

Le paramètre φ détermine si le processus AR(1) est stationnaire ou non:

$$|\varphi| = \begin{cases} < 1 & \text{le processus est stationnaire} \\ 1 & \text{marche aleatoire: le processus est donc non stationnaire} \\ > 1 & \text{le processus est explosif} \end{cases}$$

¹ - Regie Bourbonnais-Econometrie- OPC P232 - 1

Théorème : Un processus AR(p) est stationnaire si le module des solutions (les racines) de son équation caractéristique est à chaque fois strictement supérieure à 1 en valeur absolue¹.

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

$$p_k = p^k$$

Le corrélogramme partiel a ses seuls p premiers termes différent de 0.²

* **Caractéristiques de la FAC d'un AR(p)** :

Pour $K > 0$, calculons la fonction d'autocovariance : $\gamma_k = E[x_t x_{t-k}]$

$$\gamma_k = \phi_1 E[x_{t-1} x_{t-k}] + \dots + \phi_p E[x_{t-p} x_{t-k}] + E[a_{t-k}]$$

En divisant les deux membres de l'équation par :

$$p_k = \frac{\gamma_k}{\gamma_0}$$

$$p_k - \phi_1 p_{k-1} - \dots - \phi_p p_{k-p} = 0 \text{ (Équation de YULE-WALKER)}$$

$$p_k (1 - \phi_1 B - \dots - \phi_p B^p) = 0$$

En conclusion, la FAC théorique d'un AR(p) se présente selon une exponentielle amortie et/ ou une sinusoïdale amortie.

* **Caractéristiques de la FAP d'un AR(p)** :

On obtient le système d'équations linéaires de Yule-Walker :

¹ James Douglas Hamilton, Time Series Analysis, Princeton N.J, Princeton University Press, 1994
p. 799

² - Regie Bourbonnais-Econometrie OPC P238 .

$$p_1 = \phi_1 + \phi_2 p_1 + \dots + \phi_p p_{p-1}$$

$$p_p = \phi_1 p_{p-1} + \phi_2 p_{p-2} + \dots + \phi_p$$

Utilisons le système d'équation de Y-W pour calculer les premiers paramètres des processus AR(1) et AR(2).

*la FAP d'un AR(p) a ses P paramètres valeurs différentes de 0 les autres sont nulle.¹

2.3.3 Processus mixte Auto Régressif Moyenne Mobile d'ordre p et q : ARMA (p, q) :

*formulation :

les modèles ARMA sont donc représentatifs d'un processus généré par une combinaison des valeurs passées et des erreurs passées. Ils sont définis par l'équation :

$$ARMA(p, q) : (1 - \theta_1 D - \theta_2 D^2 - \dots - \theta_p D^p) x_t = (1 - \alpha_1 D - \alpha_2 D^2 - \dots - \alpha_q D^q) \varepsilon_t$$

Nous avons :

$$ARMA(1,0) = AR(1) ; ARMA(0,1) = MA(1)$$

Dans le cas d'un processus ARMA (p,q) avec constante :

$$y_t = \mu + \theta_1 x_{t-1} + \theta_2 x_{t-2} + \dots + \theta_p x_{t-p} + \varepsilon_t - \alpha_1 \varepsilon_{t-1} - \alpha_2 \varepsilon_{t-2} - \dots - \alpha_q \varepsilon_{t-q}$$

$$E[x_t] = \frac{\mu}{(1 - \theta_1 - \theta_2 - \dots - \theta_p)}$$

Donc connaissant l'espérance du processus (EViews calcul directement l'espérance du processus) la constante du processus ARMA est déterminée par :

$$\mu = E(x_t) \times (1 - \theta_1 - \theta_2 - \dots - \theta_p)$$

¹ - Regie Bourbonnais-Econometrie OPC P240

2.3.3.1 Caractéristiques des processus ARMA (p, q) :*** Caractéristiques des correlogrammes:**

les correlogrammes simples et partiels sont, par voie conséquence, un mélange des deux correlogrammes des processus AR et MA pure. Il s'avère ainsi plus délicat d'identifier ces processus à partir de l'étude des fonctions d'autocorrélation empiriques.[1]

*** Caractéristiques de la FAC d'un ARMA(p,q) :**

Pour $k > 0$. calculons la fonction d'autocovariance

$$\gamma_k = E[x_t x_{t-k}] = \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p} - \theta_1 E[a_{t-1} a_{t-k}] \dots \theta_p [a_{t-q} x_{t-k}]$$

En divisant les deux membres de l'équation par :

$$p_k = \frac{\gamma_k}{\gamma_0} = \phi_1 p_{k-1} + \dots + \phi_p p_{k-p}$$

Soit l'équation de Yule-W : $\phi_p(B)p_k = 0$

Ce résultat montre que pour établir une règle générale de comportement de la FAC d'un ARMA(p,q), il est nécessaire d'envisager deux cas :

$-q - p < 0$ c'est à dire l'ordre de la parti MA est inférieur à celui de la partie AR, dans ce cas la FAC se comporte comme un mélange de fonctions exponentielle-sinusoidale amorties dont l'allure dépend de la résolution de l'équation de Y-W : $\phi_p(B)p_k = 0$, mais aussi des valeurs initiales des coefficients $p_q, p_{q-1}, \dots, p_{q-p+1}$;

$-q - p \geq 0$ c'est à dire l'ordre de la parti MA est alors supérieur à l'ordre de AR. Dans ce cas, on obtient $-q - p + 1$ valeur initiales $(p_0, p_1, \dots, p_{q-p})$ qui n'obéiront pas au comportement général précédent,

On aura une FAC qui tendra vers 0 quand k augmente, mais son allure est inconnue pour $k \leq q - p$.¹

¹ - تومي صالح مدخل لنظرية القياس الاقتصادي دراسة نظرية ج 2 الطبعة الثانية 2010 ص 179

Tableau 2-2 : résumé des propriétés des fonctions d'autocorrélation simple et partielle.

<i>Processus</i>	<i>Fonction autocorrélation simple</i>	<i>Fonction autocorrélation partielle</i>
<i>AR(1)</i>	<i>Décroissance exponentielle ($\theta_1 > 0$) ou sinusoidale amortie ($\theta_1 < 0$)</i>	<i>Pic significatif pour le premier retard : positif si ($\theta_1 > 0$) et négatif si ($\theta_1 < 0$) les autres coefficients nuls pour des retards > 1</i>
<i>AR(2)</i>	<i>Décroissance exponentielle ou sinusoidale selon les signes de θ_1 et θ_2</i>	<i>Pic significatif pour le premier et second retard, les autres coefficients sont nuls pour des retards > 2</i>
<i>AR(p)</i>	<i>Décroissance exponentielle ou-et sinusoidale</i>	<i>Pic significatif pour les p premiers retards, les autres coefficients sont nuls pour des retards $> p$</i>
<i>MA(1)</i>	<i>Pic significatif pour le premier retard : positif si ($\alpha_1 < 0$) et négatif si ($\alpha_1 > 0$) les autres coefficients sont nuls pour des retards > 1</i>	<i>Décroissance exponentielle ($\alpha_1 > 0$) ou sinusoidale amortie ($\alpha_1 < 0$)</i>
<i>MA(2)</i>	<i>Pic significatif pour le premier et second</i>	<i>Décroissance exponentielle ou sinusoidale selon les</i>

	<i>retard, les autres coefficients sont nuls pour des retards > 2</i>	<i>signes de α_1 et α_2</i>
<i>MA(q)</i>	<i>Pic significatif pour les q premiers retards, les autres coefficients sont nuls pour des retards > q</i>	<i>Décroissance exponentielle ou-et sinusoidale</i>
<i>ARMA(1,1)</i>	<i>Décroissance géométrique à partir du premier retard le signe est déterminé par $\theta_1 - \alpha_1$</i>	<i>Décroissance exponentielle ($\alpha_1 > 0$) ou sinusoidale amortie ($\alpha_1 < 0$)</i>
<i>ARMA(p,q)</i>	<i>Décroissance exponentielle ou sinusoidale amortie tronquée après (q-p) retard</i>	<i>Décroissance exponentielle ou sinusoidale amortie tronquée après (p-q) retard</i>

SOURCE : Régie Bourbonnais-Econométrie 2007 OPC P : 240

Conditions d'utilisation :

Les modèles AR, MA, ARMA ne sont représentatifs que de chroniques :

- Stationnaires en tendance
- Corrigées des variations saisonnières

2.3.4 Processus mixte Auto Régressif Moyenne Mobile Intégré: ARIMA :

Supposons une chronique x_t échantillon d'un processus DS d'ordre d . ce processus est stationnarisé par l'utilisation du filtre aux différences d'ordre d

$$x_t \xrightarrow{(1-B)^d} y_t = (1-B)^d x_t$$

Le processus x_t contient d racines réelles unité. Il est dit intégré d'ordre d et on l'écrit :

$x_t \rightarrow I(d)$ avec d l'ordre d'intégration.

Si le processus sortant y_t est un ARMA (p,q) , alors le processus entrant x_t est un ARIMA (p,d,q) . un ARIMA (p,d,q) non saisonnier s'écrit :

$$\phi_p(B)(1-B)^d x_t = \theta_q(B) a_t$$

{ARIMA (p,d,q) = ARIMA $(p+d,q)$ non stationnaire}.

2.3.5 Processus Saisonnier Auto Régressif Moyenne Mobile Intégré d'ordre: SARIMA :

Considérons une chronique x_t échantillon d'un processus DS d'ordre D de saisonnalité s . ce processus est stationnarisé après passage dans un filtre saisonnier aux différences D -ièmes.

$$x_t \xrightarrow{(1-B^s)^D} y_t (1-B^s)^D x_t$$

Le processus x_t contient D racines réelles ou complexes unitaires. Il est dit intégré saisonnier d'ordre D et on l'écrit $x_t \rightarrow I_s(D)$ avec D l'ordre d'intégration et s la période de la saisonnalité.

Supposons que le processus sortant du filtre y_t est un processus $ARMA_{s,\hat{s}}(P, Q)$ alors le processus entrant x_t est non stationnaire :

$ARI MA_{s,\hat{s}}(P, D, Q)$, il s'écrit :

$$\phi_p(B^s)(1 - B^s)^D x_t = \theta_Q(B^s)a_t$$

On constate qu'un $ARI MA_{s,\hat{s}}(P, D, Q)$ est un $ARMA_{s,\hat{s}}(P + D, Q)$ non stationnaire et qu'un $ARI MA_{s,\hat{s}}(P, 0, Q)$ est un $ARMA_{s,\hat{s}}(P, Q)$ stationnaire.

2.4 La méthodologie de Box & Jenkins:

La partie autorégressive d'un processus, notée AR, est constituée par une combinaison linéaire finie des valeurs passées du processus. La partie moyenne mobile, notée MA, est constituée d'une combinaison linéaire finie en t des valeurs passées d'un bruit blanc. Wold(1954) montre que les modèles ARMA permettent de représenter la plupart des processus stationnaires. l'approche de BOX & JENKINS(1976) consiste en une méthodologie d'étude systématique des séries chronologiques à partir de leurs caractéristiques afin de déterminer, dans la famille des modèles ARIMA le plus adapté à représenter le phénomène étudié. Quatre étapes principales sont définies.

2.4.1 Recherche de la représentation adéquate :

2.4.1.1L'Identification :

La phase d'identification est la plus importante et la plus difficile : elle consiste à déterminer le modèle adéquat dans la famille des modèles ARIMA. Elle est fondée sur l'étude des corrélogramme simple et partiel.

A) Désaisonnalisation:

Dans le cas d'une série affectée d'un mouvement saisonnier, il convient de la retirer préalablement à tout traitement statistique. Cette saisonnalité est ajoutée à la série prévue à la fin du traitement afin d'obtenir une prévision en terme brut.

B) Recherche de la stationnarité en terme de tendance :

Si l'étude du corrélogramme simple et les tests statistiques s'y rapportant (statistique Q) présentent d'une série affectée d'une tendance, il convient d'en étudier les caractéristiques selon les tests de Dickey-Fuller. La méthode d'élimination de la tendance est fonction du processus DS ou TS sous-jacent à la chronique étudiée.

Après la stationnarisation, nous pouvons identifier les valeurs des paramètres p, q du modèle ARMA.

- Si le corrélogramme simple n'a que ses q premiers ($q=3$ maximum) différents de 0 et que les termes du corrélogramme partiel diminuent lentement, nous pouvons pronostiquer un MA(q).
- Si le corrélogramme partiel n'a que ses p premiers termes ($p=3$ maximum) différents de 0 et que les termes du corrélogramme simple diminuent lentement, cela caractérise un AR(p).
- Si les fonctions d'autocorrélation simple et partiel ne paraissent pas tronquées, il s'agit alors d'un processus de type ARMA, dont les paramètres dépendent de la forme particulière des corrélogramme

2.4.2 Estimation des paramètres :

Les méthodes d'estimation diffèrent selon le type des processus diagnostiqué.

A) Estimation du modèle AR(p):

Nous pouvons appliquer une méthode des moindres carrés ou bien nous pouvons utiliser les relations existantes entre les autocorrélations et les coefficients des modèles (équation de Yule-Walker)¹

- Méthode des moindres carrés :

$$y_t = \theta_1 y_{t-1} + \theta_2 y_{t-2} + \theta_3 y_{t-3} + \dots + \theta_p y_{t-p} + \varepsilon_t$$

A cause des obstacles de départ on va débiter l'opération d'estimation par : $t(p+1)$

$$y_{p+1} = \theta_1 y_p + \theta_2 y_{p-1} + \dots + \theta_p y_1 + \varepsilon_{p+1}$$

$$y_{p+2} = \theta_1 y_{p+1} + \theta_2 y_p + \theta_3 y_{p-1} \dots + \theta_p y_2 + \varepsilon_{p+2}$$

$$y_t = \theta_1 y_{t-1} + \theta_2 y_{t-2} + \dots + \theta_p y_{t-p} + \varepsilon_t$$

¹ مولود حشمان مرجع سبق ذكره ص 177

Sous les hypothèses des MCO on trouve :

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix}$$

$$\hat{\theta} = (\hat{y}y)^{-1} \hat{y}X$$

$$X = \begin{pmatrix} y_{p+1} \\ y_{p+2} \\ y_{p+3} \\ \dots \\ M \\ \dots \\ y_t \end{pmatrix}$$

B) Méthode de Yull-Walker:

Nous pouvons utiliser les relations existantes entre les autocorrélations et les coefficients des modèles pour estimer les modèles AR(p)

$$p_1 = \theta_1 + \theta_2 p_1 + \theta_3 p_2 + \dots + \theta_p p_{p-1}$$

$$p_2 = \theta_1 p_1 + \theta_2 p_2 + \dots + \theta_p p_{p-2}$$

$$p_p = \theta_1 p_{p-1} + \theta_2 p_{p-2} \dots + \theta_p$$

B) Estimation du modèle MA(q) et ARMA(p,q) :

L'estimation des paramètres de ces modèles s'avèrent plus complexes. Box et Jenkins suggèrent d'utiliser une procédure itérative de type balayage que nous pouvons illustrer de la manière suivante :

$$y_t = \theta_1 y_{t-1} + \theta_2 y_{t-2} + \theta_p y_{t-p} \dots + \varepsilon_t - \alpha_1 \varepsilon_{t-1} - \alpha_2 \varepsilon_{t-2} - \dots - \alpha_q \varepsilon_{t-q}$$

Action de déférence ;

$$(1 - \theta_1 D - \theta_2 D^2 - \dots - \theta_p D^p) y_t = (1 - \alpha_1 D - \alpha_2 D^2 - \dots - \alpha_q D^q) \varepsilon_t$$

Que nous pouvons écrire :

$$y_t = \frac{1}{(1 - \theta_1 D - \theta_2 D^2 - \dots - \theta_p D^p)} (1 - \alpha_1 D - \alpha_2 D^2 - \dots - \alpha_q D^q) \varepsilon_t$$

Nous posons : $v_t = \frac{1}{(1 - \theta_1 D - \theta_2 D^2 - \dots - \theta_p D^p)} \varepsilon_t$

Ce qui nous donne :

$$y_t = v_t - \alpha_1 v_{t-1} - \dots - \alpha_q v_{t-q} \text{ ou } v_t = y_t + \alpha_1 v_{t-1} + \dots + \alpha_q v_{t-q} \quad [1]$$

Nous posons : $\widehat{v}_0 = 0$ et $\widehat{v}_t = 0$ et nous calculons les valeurs estimées de \widehat{v}_t à partir de la relation [1] :

$$\widehat{v}_2 = y_2$$

$$\widehat{v}_3 = y_3 + \widehat{\alpha}_1 \widehat{v}_2$$

$$\widehat{v}_4 = y_4 + \widehat{\alpha}_1 \widehat{v}_3 + \widehat{\alpha}_2 \widehat{v}_2 \dots \text{etc.}$$

Après calcul de toute les valeurs de \widehat{v}_t , nous estimons les paramètres $\alpha_1 \dots \alpha_q$ par la méthode des moindres carrés.

2.4.3 Tests d'adéquation du modèle:

Les paramètres du modèle étant estimé (on vérifie la convergence de la procédure itérative d'estimation), nous examinons les résultats d'estimation.

- Les coefficients du modèle doivent être significativement différents de 0 (le test du t student s'applique de manière classique). Si un coefficient n'est pas significativement différent de 0 , il convient d'envisager une nouvelle spécification éliminant l'ordre du modèle AR ou MA non valide.
- L'analyse des résidus : si les obéissent à un bruit blanc, il ne doit pas exister d'autocorrélation dans la série et les résidus doivent être homoscédastique. Les tests suivants peuvent être utilisés.
 - ❖ Le test de DURBIN WATSON ,bien qu'il ne permette de détecter que des autocorrélation d'ordre 1 ;
 - ❖ Les tests de BOX et PIERCE et de LJUNG et BOX qui permettent de tester l'ensemble des termes de la fonction d'autocorrélation.3 si le résidu est à mémoire, cela signifie que la spécification du modèle est incomplète et qu'il convient d'ajouter au moins un ordre au processus ;2

- ❖ Le test ARCH d'hétéroscédasticité effectué à partir de la fonction d'autocorrelation du résidu au carré.
- ❖ Le test ARCH d'hétéroscédasticité effectué à partir de la fonction d'autocorrelation du résidu au carré.
- ❖ Les critères d'information de Akaik (AIK) a montré que le meilleur des modèles ARMA non filtré enlogst celui qui minimise la statistique :

$$AIK(p, q) = n \log \sigma_a^2 + 2(p + q)$$

- ❖ Le critère de Schwarz(1978)

$$SC(p, q) = \sigma_{at}^2 + (p + q) \log n$$

La phase de validation du modèle est très importante et nécessite le plus souvent un retour à la phase d'identification.¹

2.4.4Prévision:

A)les transformations de départ :

A l'étape de la transformation, plusieurs techniques ont pu être employées afin de stationnariser le processus générateur pour qu'il soit identifiable dans la classe des processus ARMA ;

Il est nécessaire lors de l'étape de prévision, de prendre en compte la transformation retenu (« recoloration de la prévision ») :

- Si on a utilisé la régression afin de retirer une ou plusieurs composantes déterministes estimées par les MCO, ces dernières sont extrapolées jusqu'à l'horizon prévisionnel choisi puis combinées aux valeurs prévisionnelles de l'ARMA.
- Si la transformation résulte de l'utilisation d'un filtre linéaire, alors le processus sous-jacent complet est un processus de type ARIMA que l'on peut considérer –par extension –pour la prévision comme un processus ARMA ;
- Si la transformation utilisée est celle de Box-Cox, il faudra recolorer la prévision issue du processus ARMA (prendre l'exponentielle des données si la transformation logarithmique a été appliquée sur la série brute).

¹ Econométrie des séries temporelles- cour- Hellène Hamisultarme consulté au site: helene-hamisultane.voila.net/travaux/SERIES_TEMPORELLES.pdf p14- 2

La théorie Des processus ARMA suppose que le processus de départ est un processus centré. Le fait de centrer les données peut introduire dans le modèle ARMA un terme constant. Deux cas peuvent se présenter :

- La chronique n'a pas subi de transformation, elle est centrée $(x_t - \bar{x})$ et le processus générateur obéit donc à un ARMA(p,q) :

$$\phi_p(B)(x_t - E(x_t)) = \theta_q(B)a_t$$

Soit :

$$\begin{aligned} \phi_p(B)x_t &= \phi_p(B)(x_t - E(x_t)) + \theta_q(B)a_t \\ &= (1 - \phi_1 B - \dots - \phi_p B^p)E(x_t) + \theta_q(B)a_t \\ &= \left(1 - \sum_{j=1}^p \phi_j B^j\right)E(x_t) + \theta_q(B)a_t \end{aligned}$$

D'ou $\phi_p(B)x_t = \mu + \theta_q(B)a_t$; $\mu = (1 - \sum_{j=1}^p \phi_j)E(x_t)$, du fait de la neutralité de l'opérateur B sur une constante. D'où :

$$E(x_t) = \frac{\mu}{\left(1 - \sum_{j=1}^p \phi_j\right)}$$

- La chronique a été transformée par un filtre. Considérons, par exemple, une chronique centrée $x_t - \bar{x} = y_t$ du processus générateur $y_t = x_t - E(x_t)$ filtré : $z_t = (1 - B)y_t$ et supposons que z_t obéisse à un ARMA(p,q) (ou que y_t obéisse à (un ARIMA(p,1,q)). On a :

$$\phi_p(B)(1 - B)y_t = \theta_q(B)a_t$$

$$\phi_p(B)(1 - B)(x_t - E(x_t)) + \theta_q(B)a_t$$

$$\phi_p(B)(1 - B)x_t = \phi_p(B)(1 - B)E(x_t) + \theta_q(B)a_t$$

$$= \phi_p(B)(1 - B)[E(x_t) - E(x_t)] + \theta_q(B)a_t$$

$$= \theta_q(B)a_t$$

On peut vérifier un résultat similaire à partir d'une chronique centrée après transformation par le filtre : $y_t = (1 - B)x_t = x_t - x_{t-1}$. D'où $\bar{y} \approx 0$

Si n suffisamment grand. Si y_t centrée obéit à un ARMA(p,q) alors x_t suit le processus ARIMA(p,1,q) précédent.

B) Calcul du prédicteur :

Supposons que l'on veuille prévoir un processus ARMA(p,q), (ou p,q sont les ordres maximums saisonniers, non saisonniers des parties AR et MA. A un instant t quelconque, de l'axe des temps, le processus possède une valeur x_t que l'on désire prévoir (calculer) à l'instant $t+l$ au moyen du processus ARMA (p,q) dont on connaît à présent les caractéristiques. On note cette prévision $\hat{x}_t(l)$.

l est dit date d'induction, et $\hat{x}_t(l)$ représente la prévision faite en t pour la date $t+l$. comme le processus ARMA possède différentes écritures théoriques, la prévision théorique peut s'écrire :

$$x_{t+l} = \phi_1 x_{t+l-1} + \dots + \phi_p x_{t+l-p} + a_{t+l} - \theta_1 a_{t+l-1} - \dots - \theta_q a_{t+l-q}$$

$$x_{t+l} = \sum_{j=0}^{\infty} \varphi_j a_{t+l-j}; \varphi_0 = 1$$

$$x_{t+l} = \sum_{j=1}^{\infty} \pi_j x_{t+l-j} + a_{t+l} \quad (a_t \text{ est un bruit blanc gaussien})$$

Considérons, par exemple, la seconde forme et le cas simple ou $l=1$.
La vraie valeur la prévision est donc :

$$x_{t+1} = a_{t+1} + \varphi_1 a_t + \varphi_2 a_{t-1} + \varphi_3 a_{t-2} + \dots$$

Dans cette écriture, à l'instant t, a_{t+1} est un inconnue et donc la valeur que l'on peut prévoir à l'aide de l'ARMA estimé est :

$$\hat{x}_t(1) = \varphi_1 a_t + \varphi_2 a_{t-1} + \dots$$

On commet donc une erreur à l'instant t pour la date d'induction 1 :

$$e_t(1) = x_{t+1} - \hat{x}_t(1) = a_{t+1}$$

Ainsi, on constate que les a_t utilisés dans les modèles ARMA comme étant les valeurs d'un bruit blanc sont aussi les erreurs de prévision ayant pour origine le pas de temps immédiatement antérieur. C'est la raison pour laquelle on appelle aussi les a_t les erreurs prévisionnelle ou les innovations.

Considérons à présent une date d'induction l quelconque et supposons que la meilleure prévision soit donnée par l'estimation :

$$\hat{x}_t(l) = \tilde{\varphi}_l a_t + \tilde{\varphi}_{l+1} a_{l-1} + \tilde{\varphi}_{l+2} a_{l-2} + \dots$$

Dans cette expression, les coefficients $\tilde{\varphi}_l$ sont à estimer. L'erreur de prévision est alors :

$$e_t(l) = x_{t+l} - \hat{x}_t(l).$$

Pour trouver les $\tilde{\varphi}_l$ on va utiliser le critère des MCO, c'est-à-dire calculer les qui minimisent l'erreur quadratique moyenne : $Min E[e_t(l)]^2$ avec

$$e_t(l) = x_{t+l} - \hat{x}_t(l) \\ = a_{t+l} + \varphi_1 a_{t+l-1} + \dots + \varphi_{l-1} a_{t+1} + (\varphi_l - \hat{\varphi}_l) a_t + (\varphi_{l+1} - \hat{\varphi}_{l+1}) a_{t-1} + \dots$$

Comme a_t est un bruit blanc gaussien, $e_t(l)$ obéit donc à une loi $N(0; V[e_t(l)])$

$$V[e_t(l)] = \sigma_a^2(1 + \varphi_1^2 + \dots + \varphi_{l-1}^2) + \sum_{j=1}^{\infty} (\varphi_j - \tilde{\varphi}_j)^2 \sigma_a^2$$

Comme $E[e_t(l)] = 0$ alors $V[e_t(l)] = E[e_t(l)]^2$ et cette contité est minimale si $\varphi_j = \tilde{\varphi}_j$

On peut de ce fait écrire que :

$$\hat{x}_t(l) = \varphi_l a_t + \varphi_{l+1} a_{t-1} + \dots$$

et l'erreur $e_t(l)$ s'écrit :

$$e_t(l) = a_{t+l} + \varphi_1 a_{t+l-1} + \dots + \varphi_{l-1} a_{t+1}$$

On réalise ainsi une prévision non biaisée et la variance de l'erreur de prévision est :

$$V[e_t(l)] = (1 + \varphi_1^2 + \dots + \varphi_{l-1}^2) \sigma_a^2$$

On peut noter que la fonction de prédiction $\hat{x}_t(l)$ est la nature adaptative. En effet :

$$\hat{x}_{t+l}(l) = \psi_l a_{t+1} + \psi_{l+1} a_t + \psi_{l+2} a_{t-1} + \dots$$

$$\hat{x}_{t+l}(l) = \psi_l a_{t+1} + \hat{x}_t(l+1)$$

$$\hat{x}_{t+l}(l) = \hat{x}_t(l+1) + \psi_l e_t(l)$$

Les prévisions issues des processus ARMA permettent donc de calculer le nouveau prédicteur connaissant l'ancien prédicteur $\hat{x}_{t+l}(l)$ de x_{t+l+1} d'origine t dès que la valeur du processus en $t+1$: x_{t+1} est connue. On peut ainsi adapter les prévisions faites sur la valeur x_{t+l} à l'époque t lorsque des transformations se manifestent entre t et l . Cette adaptativité montre la force des modèles ARMA par rapport au modèle de type déterministe. Cela explique aussi pourquoi les méthodes de lissage exponentiel (processus ARMA particulier) ont été abondamment utilisées pour réaliser des prévisions de court terme.

La prévision $\hat{x}_t(l)$ de variance minimale est aussi l'espérance mathématique de x_{t+1} conditionnelle en x_t, x_{t-1}, \dots . En effet :

$$E[x_{t+l}/x_t, x_{t-1}, \dots] = \psi_l a_t + \psi_{l+1} a_{t-1} + \dots = \hat{x}_t(l)$$

On constate qu'il est possible de calculer la prévision de variance minimale en utilisant la première écriture des processus ARMA. C'est ainsi que procèdent les logiciels informatiques. Compte tenu de la remarque précédente, on peut écrire :

$$\hat{x}_t(l) = \sum_{i=1}^p \phi_i E[x_{t+l-j}/x_t, \dots] - \sum_{j=0}^q \phi_j E[a_{t+l-j}/x_{1, \dots}] + E[a_{t+1-j}/x_{t \dots}]$$

Or on sait que :

$$E[x_{t-j}/x_1] = x_{t-j} \quad j=0,1,2,\dots$$

$$E[x_{t+j}/x_1] = \hat{x}_t(j) \quad j=1,2,\dots$$

$$E[a_{t-j}/x_1] = a_{t-j} - \hat{x}_{t-j-l}(1) \quad j=0,1,2$$

$$E[a_{t+j}/x_1] = 0$$

Considérons, par exemple, un processus ARMA (1.1) :

$$x_t = \phi_l x_{t-l} + a_t - \phi_l a_{t-l}$$

Calculons les prévisions en t pour des dates d'induction égales à 1,2,...,l-1 :

$$\hat{x}_t(1) = \phi_l x_t - \phi_l a_t$$

$$\hat{x}_t(2) = \phi_l \hat{x}_t(1)$$

$$\hat{x}_t(3) = \phi_l \hat{x}_t(2)$$

.
.
.

$$\hat{x}_t(n) = \phi_l \hat{x}_t(l-1)$$

Cet exemple montre que l'on peut prévoir la partie AR de l'ARMA (1.1) à une date d'induction quelconque / puisque les valeurs observées pourront être remplacées par les valeurs prévues. En revanche. La partie MA, égal ici à 1.

UN processus AR peut donc être prévu jusqu'à un horizon quelconque, tandis qu'un processus MA (q) ne pourra être prévu que jusqu'à l'horizon 1+q. Pour réaliser les calculs de la prévision lorsqu'il y a une partie MA, il faut tenir compte soit des q premières valeurs de a, inconnues soit du résultat de la prévision arrière de Box et Jenkins.

On peut enfin construire des intervalles de confiance autour des prévisions réalisées. En effet, par hypothèse les a, sont un bruit blanc gaussien : la distribution conditionnelle de la valeur à prévoir est donc une variable aléatoire d'espérance

La valeur observée :

$$\frac{x_{t+l} - \hat{x}_t(l)}{(1 + \sum_{j=1}^{l-t} \psi_j^2)^{1/2} \sigma_a} \quad \text{suit donc une loi normale } N(0 ; 1).$$

L'intervalle de confiance de la prévision est donné par :

$$x_{t+l} = \hat{x}_t(l) \pm t^{\alpha/2} (1 + \sum_{j=1}^{l-1} \psi_j^2)^{1/2} \sigma_{a_t}$$

Ou s_{a_t} est l'estimation de l'écart type du bruit blanc et $t^{\alpha/2}$ la valeur de la loi normale centrée réduite ($t^{\alpha/2} = 1.96$ pour un intervalle à 95%).

Comme nous avons les relations :

Déduisons que $x_t = \frac{\theta_B}{\phi_B} a_t$ et $x_t = \psi(B)a_t$, nous en déduisons que

$\phi_p(B)\psi(B) = \theta_q(B)$. Les coefficients du polynôme d'opérateur $\psi(B)$ peuvent être calculés connaissant ceux de $\phi_p(B)$ et $\theta_q(B)$ par identification :

$$\psi_0 = 1$$

$$\psi_1 = \phi_1 - \theta_1$$

$$\psi_2 = \phi_1\psi_1 + \phi_2 - \theta_2$$

.

.

.

$$\psi_j = \phi_1\psi_{j-1} + \dots + \phi_p\psi_{j-p} - \theta_j^1$$

Nous pouvons résumer les différentes étapes de la méthodologie de BOX & JENKINS à partir du schéma :

¹ - c'est ainsi que l'on procède pour un LES (ARIMA(0,1,1)) ou un LED (ARIMA(0,2,2)).

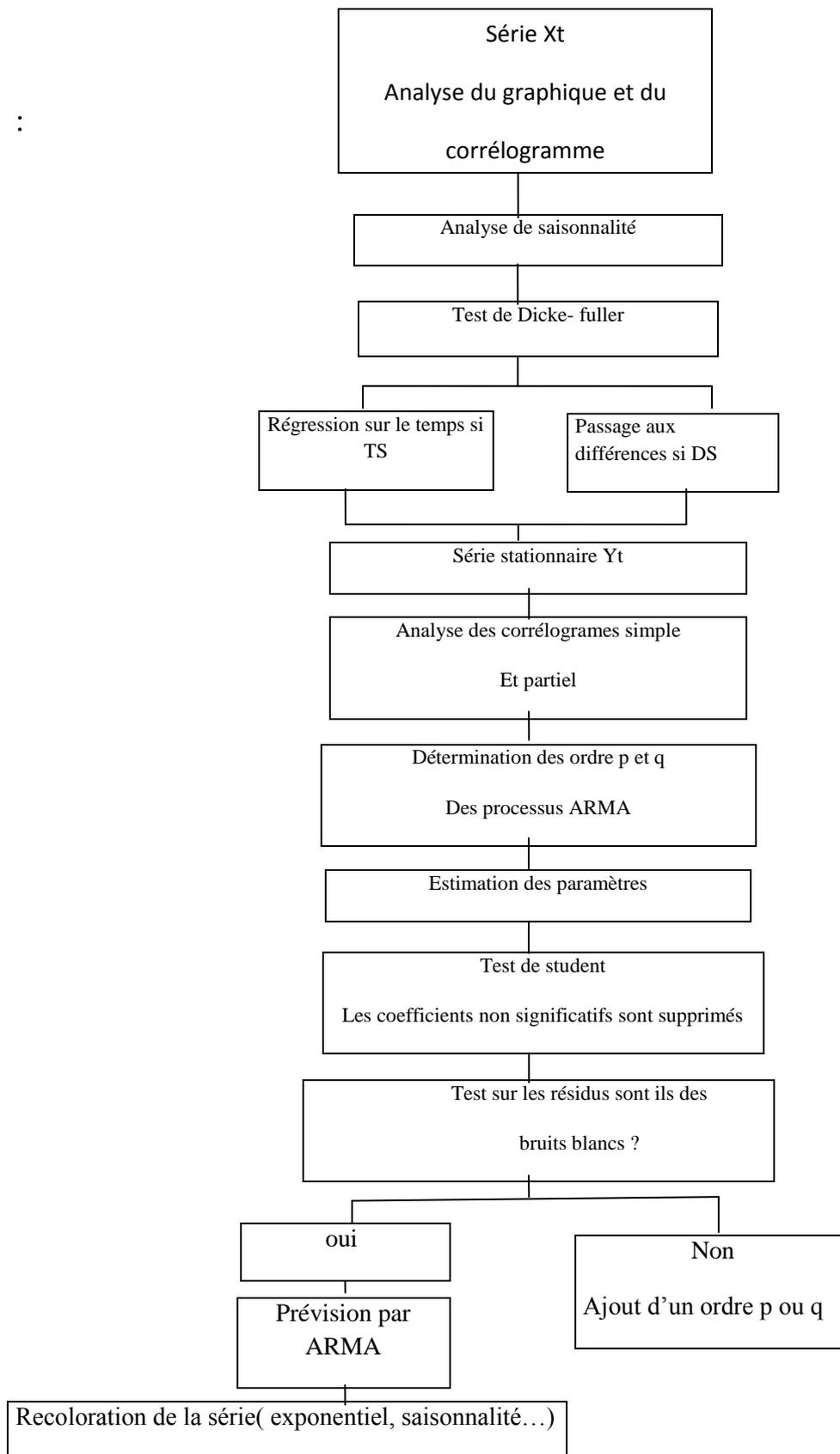


Schéma 2-2 : les étapes de la méthodologie de Box et Jenkins

Conclusion :

Les modèles des séries temporelles est très importantes pour les prévisions à court terme, ou on trouve une diversité des méthodes d'analyse statistique l'importance de ces modèles ressortir dans l'analyse des phénomènes économiques, considérons la méthode de Box et Jenkins parmi les méthodes adopté dans les prévisions économiques a cause de sa facilité ; l'application de cette méthode constate à donner des résultats à-peu-près similaire avec la réalité.

Dans ce chapitre, on va tout d'abord faire une analyse descriptive de la série consommation d'électricité pour la distribution publique, ensuite faire une présentation des données afin de se familiariser avec celle-ci, enfin nous allons appliquer la méthode de prévision vue dans le chapitre 2 afin de dégager des arguments à propos de cette méthode.

3.1 Analyse descriptive :

3.1.1 Présentation des données :

Toutes les données utilisées dans notre étude proviennent de la Direction Analyse et Prévision (DAP) de la SDO. Il s'agit de données mensuelles d'octobre 2003 à avril 2013 et cela pour la ville de SAIDA.

Pour la distribution publique qui représente les ménages et les artisans, l'électricité est essentielle dans notre vie quotidienne, pour nous éclairer, pour nous simplifier les activités ménagères et pour nous divertir. Elle est donc un produit noble et précieux.

Certaines industries utilisent de l'électricité pour l'électrolyse, mais les statistiques faisant la distinction entre cet usage et les autres au sein de ces entreprises sont rarement disponibles et, par conséquent, toute la consommation doit être déclarée comme usage énergétique. Donc les données sur lesquelles nous allons travailler sont des données mensuelles allant d'octobre 2003 à avril 2013 ce qui nous fait 115 observations.

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
01		3926910	8786880	4224490	4695911	4126426	4518314	5072196	5107660	5768487	6220438
02		3417709	3512780	3500875	4029317	3006165	3151627	4391396	4509261	4836573	5318887
03		1373369	1503336	1452986	1710744	2909750	2870592	2409652	3332125	3565577	3767979
04		3536249	4242345	3719279	4495672	4153218	4366273	4424184	5147441	5323426	5869535
05		3696139	3729230	3204310	3537441	2775740	3548163	3865962	4091901	4447594	
06		1372962	1728439	1553828	1717197	2828533	2093497	2727764	2882544	3145849	
07		4221176	4148989	1236768	4992409	5144023	5705691	5180011	5872699	6659117	
08		2826847	3756076	4091126	4270929	3380633	4738194	4961750	5867015	7216418	
09		2074222	2133120	1697606	1807244	3455509	2869253	3813130	4871599	5757852	
10	400128	4516373	4902255	4881027	5796227	5716961	6212417	6386308	7792518	9068457	
11	3427479	3975481	4145648	4095819	2970758	3301909	4288198	4995476	5651996	5874156	
12	1275277	1776640	1528824	1716490	2289148	2681615	2083082	3348488	3576552	3678690	

tableau3.1 Consommation d'électricité en GWH DP pour la série de la ville de SAIDA

3.2 Application de la méthode de Box & Jenkins:

Nous allons appliquer la méthodologie de Box & Jenkins pour notre série et cela pour effectuer des prévisions allant de Jan 2014 à Dec 2014.

La méthode de prévision de Box & Jenkins à l'instar des méthodes les plus anciennes (traditionnelle, lissage exponentiel) demande un nombre élevé de données, comme nous disposons de 115 observations nous prévoyons obtenir des prévisions de qualité supérieure.

Pour le traitement économétrique de notre série on utilise le logiciel Eviews 7.0

3.2.1 Identification et étude de la stationnarité:

A) Etude de la stationnarité de la série temporelle ELC:

Avant d'arriver à l'étape d'identification il faut d'abord analyser la stationnarité de la série temporelle au phénomène sous l'étude et ca sera par l'application des tests de stationnarités suivantes :

1) Représentation graphique de l'évolution de la série temporelle ELC:

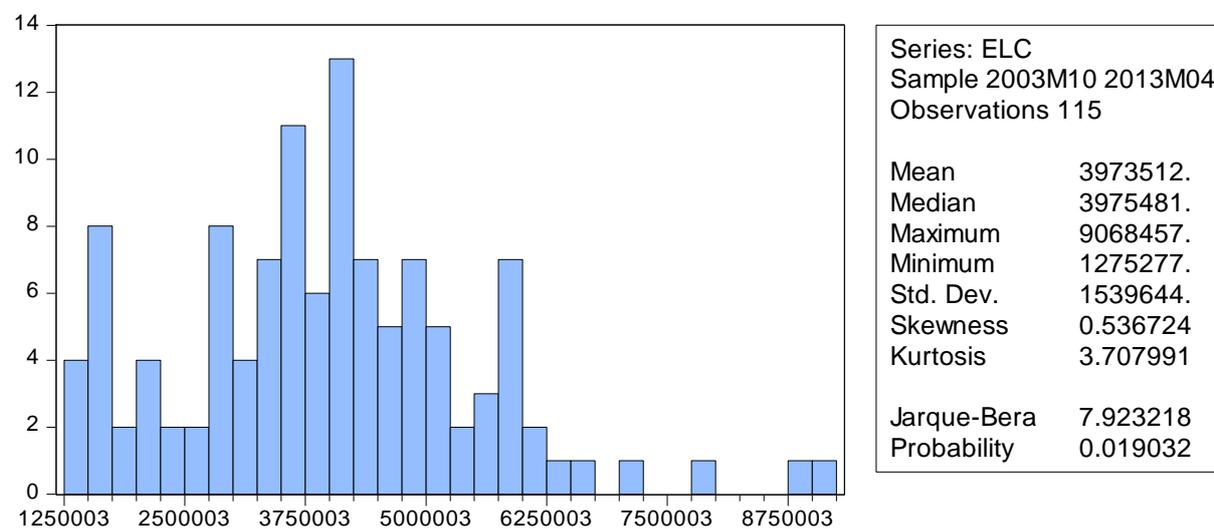
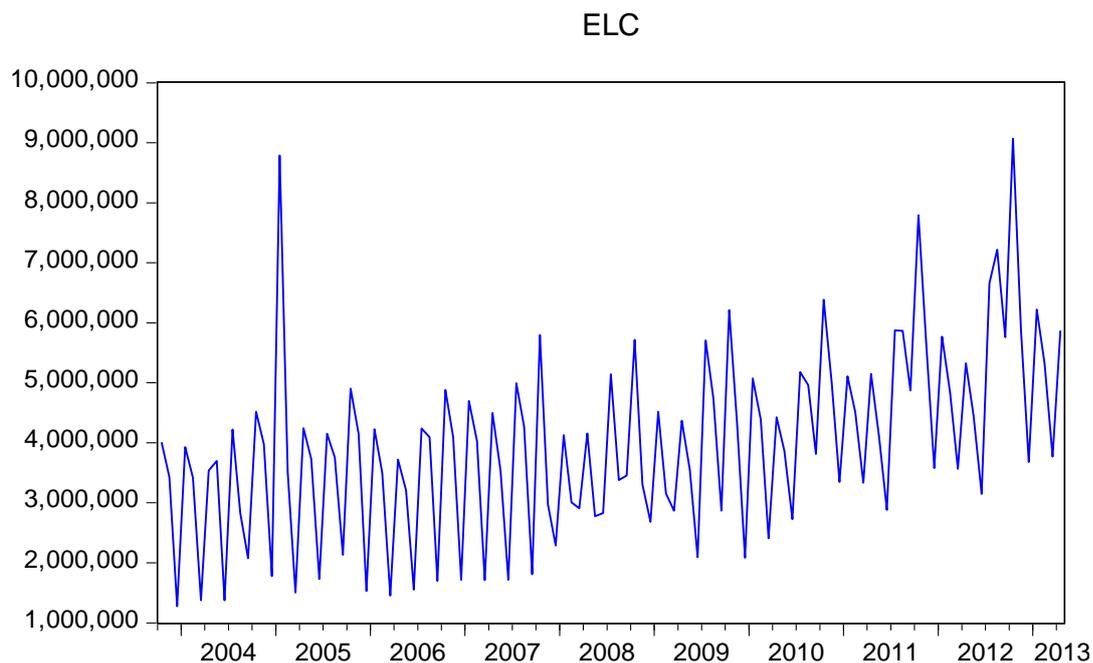


Figure (3-1) : série consommation d'électricité ville de Saida

Figure 3-2 : Evolution de la consommation d'électricité depuis 2003

**Resource : élaboré par l'étudiante a l'aide du programme
EVIIEWS 07**

D'après l'analyse graphique de la série chronologique on remarque une présence de plusieurs routeurs ce que signifie la non stationnarité de la série temporelle, et pour bien éclairer la non stationnarité on va étudier la fonction d'autocorrélation simple et partiel.

2) Analyse du corrélogramme:

Tableau(3-2) :fonction d'autocorrélation simple et partiel

Date: 06/08/13 Time: 21:11
 Sample: 2003M10 2013M04
 Included observations: 115

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.066	0.066	0.5073	0.476
		2	0.031	0.027	0.6203	0.733
		3	0.740	0.740	66.357	0.000
		4	-0.062	-0.317	66.827	0.000
		5	-0.081	-0.114	67.631	0.000
		6	0.645	0.322	118.91	0.000
		7	-0.106	-0.090	120.31	0.000
		8	-0.067	0.084	120.88	0.000
		9	0.621	0.142	169.89	0.000
		10	-0.055	0.092	170.28	0.000
		11	-0.007	0.097	170.28	0.000
		12	0.679	0.222	230.45	0.000
		13	-0.041	-0.071	230.66	0.000
		14	-0.046	-0.136	230.94	0.000
		15	0.542	-0.128	270.49	0.000
		16	-0.097	0.088	271.78	0.000
		17	-0.103	-0.003	273.22	0.000
		18	0.479	-0.039	305.09	0.000
		19	-0.107	0.026	306.68	0.000
		20	-0.111	-0.059	308.42	0.000
		21	0.481	0.104	341.53	0.000
		22	-0.090	-0.079	342.71	0.000
		23	-0.060	0.073	343.24	0.000
		24	0.523	0.090	383.61	0.000
		25	-0.069	-0.031	384.31	0.000
		26	-0.086	-0.028	385.43	0.000
		27	0.430	-0.098	413.65	0.000
		28	-0.156	-0.099	417.42	0.000
		29	-0.147	-0.002	420.79	0.000
		30	0.378	0.037	443.35	0.000
		31	-0.174	-0.027	448.18	0.000
		32	-0.162	-0.070	452.43	0.000
		33	0.375	0.014	475.49	0.000
		34	-0.181	-0.054	480.95	0.000
		35	-0.113	0.051	483.08	0.000
		36	0.382	-0.066	507.88	0.000

D'après l'analyse statistique du graphe on remarque que la plupart des coefficients de la fonction d'autocorrélation ne sont pas significatifs comme on allusion qu'elles sortent de l'intervalle **[0.166,-0.166]**

Donc la série est non stationnaire.

Test de Ljung Box :

On utilise ce test pour étudier la signification globale des coefficients d'autocorrélation inférieur de 30.

La statistique de L b est donné par :

$$L b = n(n + 2) \sum_{k=1}^{30} \frac{\hat{p}_k^2}{n - k}$$

$$Lb = 155(155 + 2) \sum_{k=1}^{30} \frac{\hat{p}_k^2}{155 - k}$$

$$Lb = 443,35 > x_{0,05,30}^2 = 43,373$$

Analyse statistique:

On a Lb supérieur de Ttab $x_{0,05,30}^2$

Donc on rejette l'hypothèse nulle qui propose que tous les coefficients d'autocorrélation égal à 0.

On peut également tester la non stationnarité par la méthode statistique comme on a vu dans le deuxième chapitre ;

Cette méthode basée sur le test de la racine unitaire de Dickey-Fuller :

3) Test de Dickey-Fuller de la série ELC :

Choix du nombre de retards optimal :

Avant de pouvoir appliquer le test de Dickey-Fuller, nous devons déterminer le nombre de retards p qui minimise les critères d'Akaike et Schwartz pour les trois modèles (avec tendance et constante (trend and intercept), avec constante (intercept), sans tendance ni constante (none)).

Les valeurs des critères d'Akaike et Schwartz sont fournies par le logiciel Eviews et sont résumées dans le tableau suivant :

Tableau 3.3- Critères d'Akaike et Schwartz pour la série

Modèles / Retards		0	1	2	3	4
MODELE 3	AIK	41.16626	41.16441	41.25871	41.27472	41.26160
	SC	41.30093	41.34580	41.48773	41.55226	41.58854
MODELE 2	AIK	41.12688	41.10688	41.20017	41.21089	41.20407
	SC	41.21666	41.24292	41.38339	41.44218	41.48431
MODELE 1	AIK	41.06859	41.05299	41.14436	41.16120	41.14990
	SC	41.11348	41.14368	41.28177	41.34623	41.38343

D'après le tableau (3.3) nous constatons que le critère d'Akaike est minimisé pour les trois modèles pour un nombre de retard $p = 1$ tandis que le critère de Schwartz est minimisé pour $p = 0$. En suivant le principe de parcimonie nous retiendrons le nombre de retards qui permet d'estimer le minimum de paramètres c'est-à-dire $p = 0$. Dans ce cas on utilise le test de Dickey-Fuller simple (DF), donc il n'y a pas d'autocorrélation des erreurs.

3-1) Le test de Dickey-Fuller Augmenté :

On commence par tester la présence d'une racine unitaire à partir du modèle le plus général à savoir le modèle (3) incluant une constante et une tendance :

$$ELC_t = \phi_1 ELC_{t-1} + b_t + c + \varepsilon_t$$

$$\text{Forme du test : } \begin{cases} \phi_j = 1 \\ \phi_j \neq 1 \end{cases}$$

On teste alors l'hypothèse nulle $H_0 : b = 0$ de présence de racine unitaire (la série n'est pas stationnaire) contre l'hypothèse alternative de stationnarité $H_1 : b \neq 0$.

Eviews nous fournit les résultats suivants :

Tableau 3.4 : test de Dickey-Fuller estimé par le modèle[3]

Null Hypothesis: ELC has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 11 (Automatic - based on Modified AIC, maxlag=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-0.572458	0.9784
Test critical values:		
1% level	-4.049586	
5% level	-3.454032	
10% level	-3.152652	

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

D'après les résultats précédents on accepte hypothèse H_0 de-ce-fait il existe un racine unitaire donc la série ELC est non stationnaire.

$$t\phi_{jcal} = -0.572458 > t\phi_{jtab} \text{ au seuil de coefficient (1\%, 5\%, 10\%)(-4.049586, -3.454032, -3.152652)}$$

Pour rendre la série stationnaire on utilise la première différence ce que nous permis les résultats ci-dessus :

Tableau 3.5: test de Dickey-Fuller augmenté

Null Hypothesis: D(ELC) has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 0 (Automatic - based on Modified AIC, maxlag=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-17.62755	0.0000
Test critical values:		
1% level	-4.041280	
5% level	-3.450073	
10% level	-3.150336	

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

Analyse statistique :

$$\text{Test d'hypothèse : } \begin{cases} H_0: \phi_j = 1 \\ H_1: \phi_j \neq 1 \end{cases}$$

Dans ce cas on accepte l'hypothèse alternative $t\phi_{jcal} = -17.62755 < t\phi_{jtab}$ au seuil de coefficient (1%, 5%, 10%)(-4.041280, -3.450073, -3.150336)

Donc il n'existe pas un racine unitaire

Alors la série ELC est stationnaire à l'ordre 1 $D(ELC) \rightarrow coin(1)$

3-2) Test de philips-Pérron:

$$s_t^2 = \frac{1}{115} \sum_{t=1}^{115} e_t^2 + 2 \sum_{i=1}^1 \left(1 - \frac{1}{l+1}\right) \frac{1}{115} \sum_{t=i+1}^{115} e_t e_{t-i}$$

Avant estimer s_t^2 il faut d'abord calculer le nombre de retard :

(Troncature de Newey-West)

$$L = 4 \left(\frac{n}{100}\right)^{2/9} = 4 \left(\frac{115}{100}\right)^{2/9} = 1.38$$

Tableau 3-6 : test de Philips-perron

Null Hypothesis: ELC has a unit root
 Exogenous: Constant, Linear Trend
 Bandwidth: 2 (Newey-West automatic) using Bartlett kernel

	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	-13.20047	0.0000
Test critical values:		
1% level	-4.040532	
5% level	-3.449716	
10% level	-3.150127	

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

Analyse statistique :

A partir de ce tableau et après faire la correction non paramétrique (PP) avec les retards on rejette l'hypothèse H_0 dans la série ELC

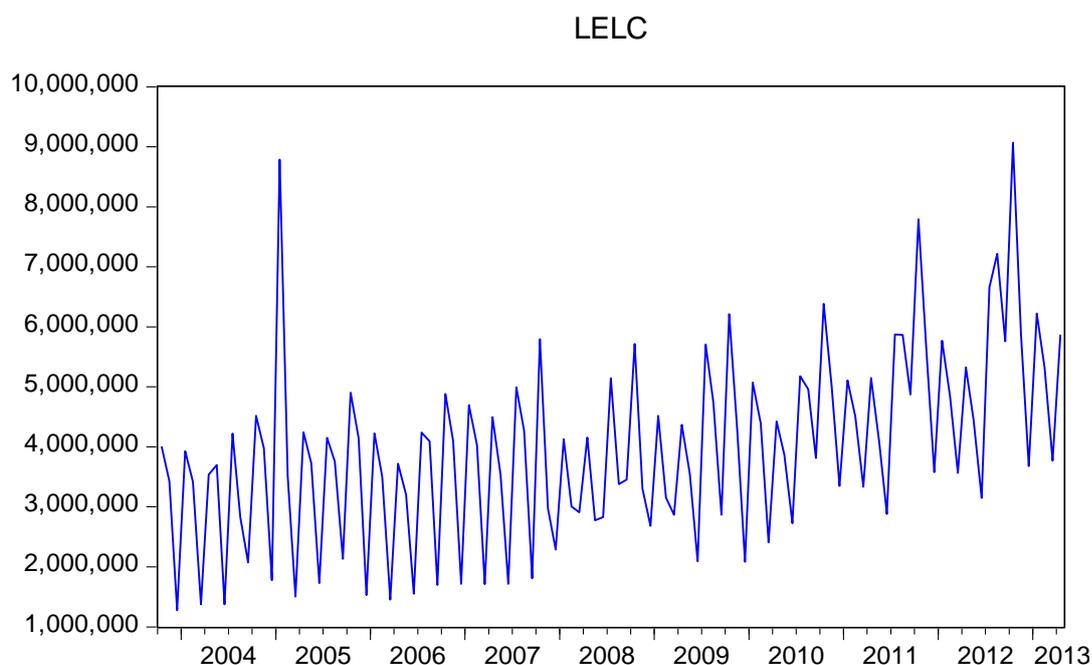
La statistique (pp) inférieur des valeurs critiques

Conclusion:

Après les tests ADF-PP on accepte l'hypothèse de stationnarité de la série ELC.

- **Dessaisonalisation des composantes de la série ELC:**

Féfigure 3-2 : Evolution de la consommation d'électricité avec LOG a partir d'octobre 2003.



On remarque à partir du graphe de la variable sous l'étude ELC qu'elle existe une composante de la tendance et une autre composante saisonnière.

Pour retirer la composante saisonnière il existe plusieurs méthodes :

**** Entrer les coefficients saisonniers :**

Après utiliser le programme EVIEWS on trouve des valeurs saisonnières adéquates pour retirer la composante saisonnière de la série chronologique.

On va retirer la composante saisonnière de la série IELC :

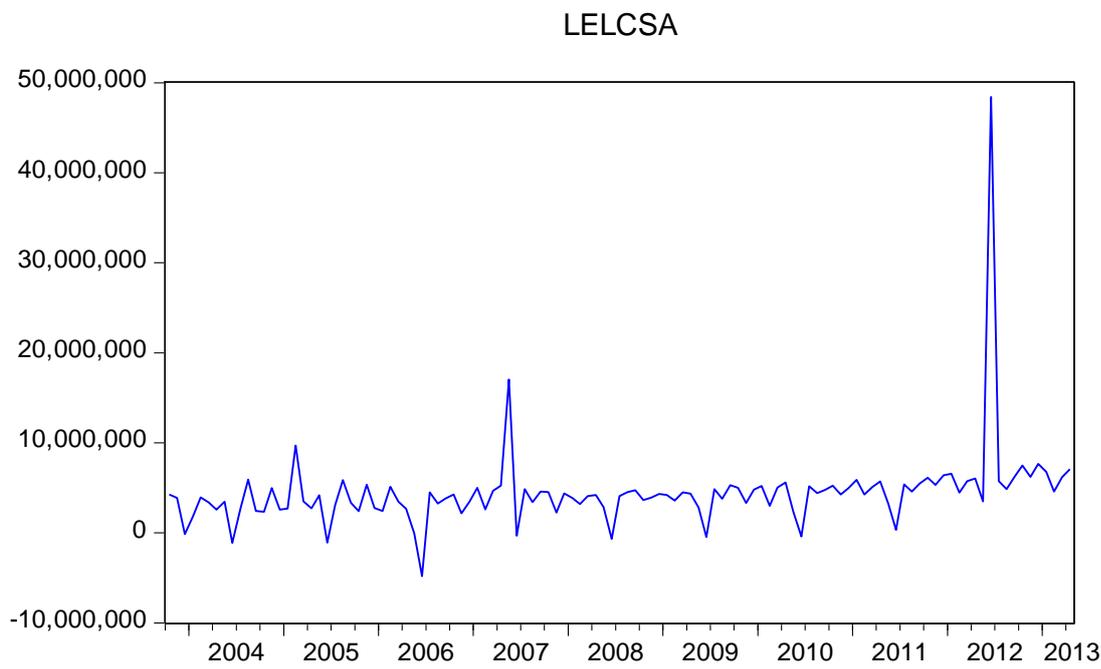
Tableau 3.7 :représente la série LELC après retirer la saisonnalité

Date: 06/11/13 Time: 14:37
 Sample: 2003M10 2013M04
 Included observations: 115
 DifferencefromMovingAverage
 Original Series: LELC
 AdjustedSeries: LELCSA

ScalingFactors:

1	-905475.3
2	-903747.7
3	31381.19
4	-1209857.
5	67098.89
6	4827832.
7	-1306064.
8	-1693570.
9	407548.8
10	-262935.1
11	-457876.8
12	1405666.

Figure3.3 :la série désaisonnalisé (LELCSAt)



Analyse du graphe:

On trouve la nouvelle série désaisonnalisé IELCSA après la soustraction des coefficients. Après l’analyse graphique on peut exprimer que la série est stationnaire.

Tableau 3.8 : Fonction d'autocorrélation de la série désaisonnalisé

Date: 06/11/13 Time: 14:50
 Sample: 2003M10 2013M04
 Included observations: 115

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.041	0.041	0.1993	0.655
		2	0.058	0.056	0.6006	0.741
		3	0.098	0.094	1.7600	0.624
		4	0.074	0.065	2.4230	0.658
		5	0.093	0.080	3.4917	0.625
		6	0.135	0.117	5.7437	0.453
		7	0.071	0.048	6.3795	0.496
		8	0.042	0.011	6.6037	0.580
		9	0.076	0.040	7.3391	0.602
		10	0.049	0.014	7.6479	0.663
		11	0.006	-0.031	7.6531	0.744
		12	0.023	-0.016	7.7240	0.806
		13	-0.026	-0.057	7.8122	0.856
		14	0.028	0.008	7.9166	0.894
		15	0.035	0.017	8.0798	0.921
		16	-0.012	-0.021	8.0993	0.946
		17	0.047	0.046	8.3965	0.957
		18	0.045	0.046	8.6805	0.967
		19	-0.001	0.001	8.6806	0.978

		20	0.039	0.032	8.8936	0.984
		21	0.023	0.010	8.9673	0.989
		22	-0.008	-0.017	8.9772	0.993
		23	0.028	0.008	9.0917	0.996
		24	0.001	-0.025	9.0918	0.997
		25	-0.041	-0.056	9.3393	0.998
		26	0.043	0.027	9.6196	0.999
		27	0.045	0.036	9.9261	0.999
		28	-0.027	-0.025	10.039	0.999
		29	0.043	0.038	10.323	0.999
		30	0.026	0.030	10.434	1.000
		31	-0.026	-0.018	10.544	1.000
		32	0.026	0.012	10.654	1.000
		33	0.017	0.007	10.702	1.000
		34	-0.029	-0.032	10.844	1.000
		35	0.008	-0.013	10.854	1.000
		36	-0.035	-0.054	11.061	1.000

Analyse statistique:

On remarque que tous les coefficients d'autocorrélation sont nuls, tous les termes sont à l'intérieur de l'intervalle de confiance et la statistique de Ljung-

Box $Q = 10.43$ est inférieure à la valeur tabulée du khi-deux à 5 degrés de liberté $\chi^2_{0.05, 30} = 43.77$.

Décision statistique :

La statistique $L_b = 10.43 < \chi^2_{0.05, 30} = 43.77$ en accepte l'hypothèse H_0 donc la série est stationnaire.

Test de stationnarité après retirer la composante saisonnière :

Tableau 3.9 : Test de Dikey-Fuller augmenté pour le troisième modèle

Null Hypothesis: LELCSA has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 12 (Automatic - based on Modified AIC, maxlag=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-0.822357	0.9595
Test critical values:		
1% level	-4.050509	
5% level	-3.454471	
10% level	-3.152909	

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

Analyse statistique :

D'après les résultats précédents on accepte hypothèse H_0 de-ce-fait il existe un racine unitaire donc la série ELC est non stationnaire.

$$t\phi_{jcal} = -0.822357 > t\phi_{jtab} \text{ au seuil de coefficient (1\%, 5\%, 10\%)(-4.050509, -3.454471, -3.152909)}$$

Pour rendre la série stationnaire on utilise la première différence ce que nous permis les résultats ci-dessus :

Tableau 3.10 : test ADF première différence pour le modèle [3]

Null Hypothesis: D(LELCSA) has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 0 (Automatic - based on Modified AIC, maxlag=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-17.13809	0.0000
Test critical values:		
1% level	-4.041280	
5% level	-3.450073	
10% level	-3.150336	

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

Analyse statistique :

$$\text{Test d'hypothèse : } \begin{cases} H_0: \phi_j = 1 \\ H_1: \phi_j \neq 1 \end{cases}$$

Dans ce cas on accepte l'hypothèse alternative $t\phi_j \text{ cal} = -17.13809 < t\phi_j \text{ tab}$ au seuil de coefficient (1%, 5%, 10%)(-4.041280,-3.450073,-3.150336)

Donc il n'existe pas une racine unitaire à l'ordre 1

Alors la série ELC est stationnaire à l'ordre 1 $D(ELC) \rightarrow \text{coin}(1)$

Test ADF première différence

A) la modélisation de la série ELC:

Après confirmer que la série est stationnaire on vient à l'étape d'identification des ordres (p,q) du modèle mixte ARMA cette phase est la plus importante et la plus difficile.

3.2.1 Identification du modèle :

La série (LELCSA_t) s'écrit sous la forme :

$$LB13sa_t = \phi_1 LB13sa_{t-1} + \phi_2 LB13sa_{t-2} + \dots + \phi_p LB13sa_{t-p} + \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

Après ajouter le coefficient de retard (1) on trouve :

$$\Rightarrow (1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) L B I 3 s a_t = (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q) \varepsilon_t$$

On préfère le modèle qui nous donne la meilleure combinaison entre les critères (Schwarze, Akaike) on respectent le niveau de R^2 et la statistique de DW.

Dans le modèle(LELCSA), la fonction d'autocorrélation partielle prouve que ses termes AR(1), AR(4), AR(5) ne sont pas nulle tandis que la fonction d'autocorrélation simple prouve que ses termes MA(4), MA(1) ne sont pas nulle .

Après constater les modèles filtrant :

On peut choisir le modèle $MA(1)$ a cause de plusieurs considérations :

- La valeur minimum des critères (Schwarze, Akaike)
- Niveau maximum de déterminant

D'après ces points la meilleure forme mathématique du modèle déterminatif de la série stationnaire (LELCSA_t) sera sous la forme :

$$LELCSA_t = \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

Et après joindre le coefficient de retard L on trouve la forme ci-dessus:

$$LELCSA_t = \delta + (1 - \theta_1 L) \varepsilon_t$$

$$LELCSA_t \rightarrow MA(1)$$

3.2.2 Estimation du modèle:

Estimation du modèle identifiée à l'aide du programme EVIEWS 7.0

Qui nous fourni les résultats suivantes :

Tableau 3.11 :

Résultat d'estimation du modèle identifier de la série (LELCSA)

Dependent Variable: LELCSA

Method: Least Squares

Date: 06/11/13 Time: 15:37

Sample (adjusted): 2003M10 2013M04

Included observations: 115 after adjustments

Convergence achieved after 8 iterations

Backcast: 2004M01

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-12.05071	10.12758	-1.995625	0.1244
MA(1)	0.389578	0.114472	-2.960360	0.0042
R-squared	0.083252	Mean dependent var		-17.12641
Adjusted R-squared	0.069965	S.D. dependent var		132.7048
S.E. of regression	127.9783	Akaike info criterion		12.56936
Sum squared resid	1130112.	Schwarz criterion		12.63310
Log likelihood	-444.2124	F-statistic		6.266025
Durbin-Watson stat	1.085263	Prob(F-statistic)		0.014678
Inverted MA Roots	00.35			

Donc on peut formuler le modèle $LELCSA_t \rightarrow MA(1)$ mathématiquement par cette forme :

$$LECSA_t = -12.05071 + (1 + 0.389578L)\varepsilon_t$$

$$LELCSA_t = -12.05071 + \varepsilon_t + 0.389578\varepsilon_{t-1}$$

$$(10.12758) \quad (0.114472)$$

$$R^2 = 0.083252$$

$$n = 115$$

$$DW = 1.08$$

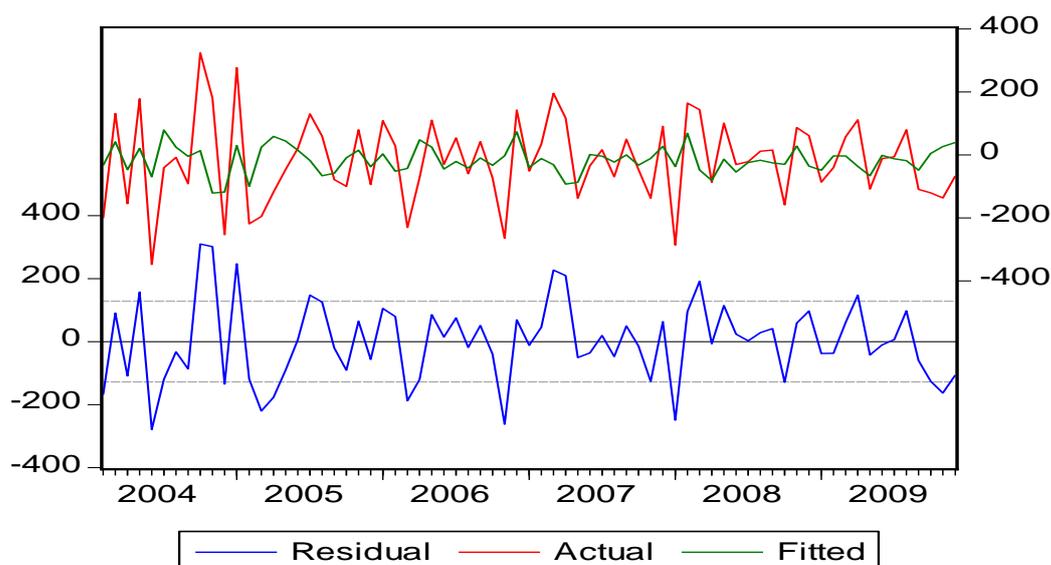
Ou les valeurs lesquelles entre parenthèse représente la variance standard des paramètres estimées

3.2. 3 Validation du modèle:

Cette étape a un but d'examiner la fiabilité du modèle statistique qu'on a choisi $LELCSA_t \rightarrow MA(1)$ à travers les points ci-dessus ;

A) Comparaison entre la série originale et la série estimé:

figure 3.4 : représente le graphe des résidus



B) Analyse de la fonction d'autocorrélation des résidu :

Pour tester si les paramètres de la fonction d'autocorrélation du résidus appartient dans l'intervalle de signification , on utilise le test de Lujing-Box-

Pierce ou :
$$Q^* = 115(115 + 2) \sum_{i=1}^k (k - i)r_i^2 \rightarrow \chi_{k-p-q}^2$$

d'après le tableau qui teste la signification des coefficients d'autocorrélation globale et partielle en Remarque:

$$Q^* = 31.2318$$

Et pour que c $k=30$, $p=0$, $q=1$:

$$Q^* = 31.2318 < \chi_{0.05} = 48.278$$

Donc en accepte l'hypothèse qui constate que tous les coefficients de la fonction d'autocorrélation ne tranche pas de 0 ne diffère pas de 0.

C) Tests des parametre des coefficients estimé:

➤ **Tester la signification de: $\hat{\theta}_1$**

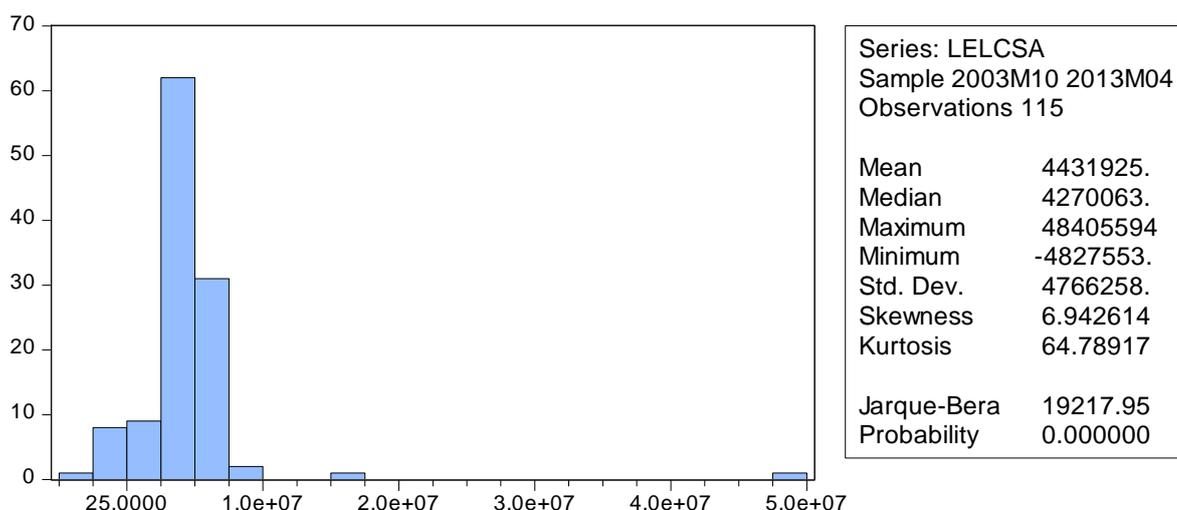
On a la statistique calculé $|t_{calcul}^*| = 1.9956$ supérieure de 1.96(au seuil de coefficient 5%) on peut rejeter l'hypothèse nulle ($H_0 : \hat{\theta}_1 = 0$) donc le paramètre estimé $\hat{\theta}_1$ différent de 0.

➤ **Tester la signification de: $\hat{\delta}$**

On a aussi la statistique calculée $|t_{calcul}^*| = 2.9603$ supérieure de 1.96 donc en rejette l'hypothèse ($H_0 : \hat{\delta} = 0$) et dans ce cas le paramètre estimé $\hat{\delta}$ se diffère de 0

D) Test de normalité du résidus (test de Jarque et Bera):

Figure 3.5 : test de normalité du résidus



• **Test de Skewness :**

Pour tester l'hypothèse nulle $H_0 : v_1=0$ on va calculer la statistique :

$$v_1 = \frac{|\beta_1^{1/2} - 0|}{\sqrt{6/n}} = \frac{|6.942614 - 0|}{\sqrt{6/115}}$$

On a $v_1 < 1.96$ donc on accepte l'hypothèse $H_0 : v_1=0$

Donc la série des résidus est symétrique

- **Test de Kurtosis :**

Dans ce cas on va tester l'hypothèse d'aplatissement normale $H_0 : v_2=0$

$$v_2 = \frac{|\beta_2 - 3|}{\sqrt{24/n}} = \frac{|64.78917 - 3|}{\sqrt{24/115}}$$

On a $v_2 < 1.96$: on accepte l'hypothèse d'aplatissement normale de la série des résidus.

- **Test de Jarque- Bera :**

Pour tester l'hypothèse nulle (la série des résidus suit une distribution normale) on va calculer la statistique de Jarque et Bera :

$$S = \frac{n}{6}\beta_1 + \frac{n}{24}(\beta_2 - 3)^2 \rightarrow \chi^2_{1-\alpha}(2)$$

On a $S = 2.1028 < \chi^2_{0.05}(2) = 5.99$ de ce fait on accepte l'hypothèse de distribution normale du résidu (H_0) au seuil de signification 5%.

Alors ε_t est un Bruit Blanc qui suit une distribution normale.

3.2.4 Prévision:

Peut-être parvenir que la meilleure représentation de la série $LELC_t$

$$LELCS \longrightarrow MA(1)$$

On peut construire un modèle prévisionnel au niveau de la série $LELCSA$ canonique à la fin de **Mai 2013** par l'équation rétrogradée estimée :

$$\begin{aligned} LELCSA_t &= \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1} \\ LELCSA_t &= -12.05071 + \varepsilon_t - (0.389578)\varepsilon_{t-1} \\ LELCSA_t &= -12.05071 + \varepsilon_t + 0.389578\varepsilon_{t-1} \\ LELCSA_{MAI2013} &= -12.05071 + \varepsilon + 0.338878\varepsilon_{MAI2013} \end{aligned}$$

Considérons tous les erreurs qui viendront au futur être nulles ($\varepsilon_t = 0$)

Et pour le but de recouvrer la valeur prévisionnelle de la série $LELC$ de mois de **Mai 2013** par l'équation rétrogradée par première différence :

$$LELCSA_t = LELCSA_t - LELCSA_{t-1}$$

$$LELCSA_t = LELCSA_{t-1} + LELCSA_t$$

$$LELCSA_{MAI2013} = LSA_{AVR2013} + DLB13SA_{MAI2013}$$

$$LB13SA_{MAI2013} = 1097 + (-51.7471)$$

$$LB13S_{MAI2013} = 1045.25$$

On peut maintenant calculer la valeur de la consommation d'électricité prévu au mois de **Mai 2013** d'après la relation :

$$LELCSA_{i,j} = LELC_{i,j} - CS_i$$

$$\forall (i = 1, 2, \dots, 12)(j = 2003, 2004, \dots, 2013)$$

Ou CS_i : représente le coefficient saisonnier du mois i

$LELC_i$: Proportion de la consommation du mois i canonique à l'année j

Avec les même étapes précédentes qu'on a adopté pour calculer la valeur prévu de la consommation d'électricité à propos du mois de **Mai 2013** en peut calculer les valeurs prévisionnelles de la consommation d'électricité avec un horizon de 12 mois.

Le tableau suivant définir les valeurs recouvrable:

Tableau 3.12- Prévission par la méthode de Box & Jenkins sur la consommation d'électricité

Valeurs prévisionnelles	Mois
6025413	01/2014
6523541	02/2014
6158962	03/2014
6589313	04/2014
7085696	05/2014
7152365	06/2014
7589621	07/2014
7612895	08/2014
8032697	09/2014
8546921	10/2014
8912458	11/2014
9063854	12/2014

D'après le tableau précédent en remarque que les résultats recouvrables par la méthode de Box & Jenkins rapprochent les résultats constaté

La consommation d'électricité en 2014 pour la distribution publique pour la ville de SAIDA connaîtra des augmentations vénielles ainsi elle va connaître une stationnarité surtout dans le deuxième trimestre par rapport à 2013.

Conclusion :

Considérons la consommation d'électricité a connue des perturbations ces dernières années il est intéressant pour la SONELGAZ de faire une étude scientifique et rigoureuse sur l'évolution dans le temps de cette matière stratégique dans ses déférents aspects, et ce tant pour des prises de décisions que pour les planifications futures. Les résultats de notre étude économétrie pour but de mettre on place un modèle de prévision à cour terme appliqué à la consommation d'électricité pour la distribution publique c'est un modèle qui nous a donnée des prévisions constatent que la consommation d'électricité connaîtra des augmentations vénielles ; aussi elle connaîtra une stationnarité surtout dans le deuxième trimestre de 2014.

Conclusion générale :

L'électricité représente une des sources d'énergie, les plus utilisées dans le monde, voir la plus prometteuse puisqu'elle bénéficie avec la montée des préoccupations environnementales d'une considération importante. L'Algérie est un pays producteur et exportateur de cette énergie, il a été donc primordial pour la SONELGAZ, d'analyser et prévoir l'évolution temporelle de la consommation d'électricité, afin de comprendre le mécanisme qui gère cette consommation et de tirer par la suite des informations pertinentes qui serviront à la phase de prise de décisions.

Notre étude a porté sur l'analyse des séries chronologiques représentant l'évolution de la consommation trimestrielle d'électricité pour la distribution publique. Dans la phase descriptive, nous avons remarqué que la consommation d'électricité pour la distribution publique est directement liée aux facteurs météorologiques. Ce qui voulait dire que nous allions étudier une série chronologique qui comprenant 115 observations.

Nous avons appliquée la méthode de Box & Jenkins sur nos données. Comme notre série étaient affectées d'une forte saisonnalité, nous étions donc obligé de les désaisonnaliser à l'aide l'opérateur de dessaisonalisation. Après avoir effectué la dessaisonalisation, les séries obtenues étaient stationnaires, nous avons donc choisi de les modéliser par les modèles ARMA. Une fois le modèle choisi, estimé et validé, nous avons calculé nos prévisions.

Au cours de ce travail nous avons eu à illustrer l'emploi du modèle d'analyse d'interventions, un des modèles de la méthode de BOX & JENKINS, pour la prévision de la consommation d'électricité, et ce, en utilisant une série allant d'octobre2003 à avril 2013.

Cela nous a conduit aux résultats suivant lesquels la série de la consommation d'électricité est générée par un modèle ARMA(1,2) expliquant le fait que les valeurs de la consommation observées au temps t dépendent des valeurs et des chocs aléatoires de la période $t-1$.

C'est donc non seulement dans la mémoire de notre série chronologique sous étude, mais également dans les variables binaires par lesquelles nous avons saisi l'information qualitative que nous avons dégagé le modèle qui représente le mieux le processus stochastique par lequel notre série est générée.

Aussi, nous recommandons son utilisation dans toutes les prévisions portant sur des variables d'intérêt stratégique.

Toutefois, il sera intéressant de faire une étude explicative sur la consommation d'électricité pour la distribution publique, en utilisant une méthode de prévision économétrique incluant comme variables explicatives : les facteurs météorologiques, le nombre d'abonnés, le taux de pénétration du gaz naturel... etc.

Sommaire

Introduction générale

Chapitre 1 : prévision de la consommation d'électricité

1 .1<u>Généralité sur la prévision:</u>	02
1.1.1 Définition de la prévision.....	02
1.1.2 Objectif de la prévision.....	02
1.1.3 Horizon de prévision.....	02
2.1Classification des techniques de prévision.....	03
2.1.1 Les méthodes informelles.....	03
2.1.2 Les méthodes extrapolatives.....	04
2.1.3 Les méthodes explicatives.....	05
2.1.4 Les méthodes systémiques.....	05
1 .2 <u>Consommation d'électricité :</u>	
1.2.1 Définition de la consommation.....	06
1.2.2 Les types de la consommation d d'électricité.....	07
1.2.3 la consommation finale d'électricité	07
1.3 <u>Présentation de la SONELGAZ ;</u>	
1.3.1 Historique	08
1.3.2 Macro structure de la DR-SAÏDA.....	09
1.4 <u>Présentation de la direction Analyse et Prévision (DAP):</u>	
1.4.1Principales activités de la DAP.....	12

Chapitre 2 : étude analytique des séries temporelles

2.1 concepts de base sur les séries chronologiques :

2.1.1 Définition d'une série chronologique.....	14
2.1.2 Les composantes d'une série chronologique.....	15
2.1.3 Schémas de décomposition d'une série chronologique.....	18
2.1.3.1 Procédure de choix d'un schéma de décomposition.....	19
2.1.4 Détections des composantes d'une série chronologique.....	20
2.1.4.1 Analyse des graphes.....	21
2.1.4.2 Tests statistiques.....	21
2.1.4.2.1 Détection de la tendance.....	22
2.1.4.2.2 Détection de la saisonnalité.....	22

2.2 Elément d'analyse des séries temporelles:

2.2.1 Processus stochastique	23
2.2.2 Stationnarité.....	24
2.2.2.1 Définition et propriétés	24
2.2.2.2 Fonction d'autocorrelation simple et partielle.....	25
2.2.2.3 Tests de « bruit blanc » et de stationnarité	25
2.2.2.3.1 Analyse des fonctions d'autocorrelation.....	26
2.2.2.3.2 Statistique de Box-Pierce et Ljung-Box.....	26
2.2.2.3.3 Les tests de normalité	27
2.2.3 Les processus aléatoires non stationnaires.....	28
2.2.3.1 Description des processus TS et DS.....	28
2.2.3.2 Les tests de racines unitaires et la stratégie séquentielle de test.....	30

2.2.3.2.1 Les tests de Dickey-Fuller simple (DF).....	31
2.2.3.2.2 Les tests de Dickey-Fuller Augmentés (ADF).....	33
2.2.3.2.3 Le test de Philips et Perron.....	34
2.2.3.2.4 Stratégie de tests.....	35

2.3 Processus linéaires des séries temporelles :

2.3.1 Processus Moyenne Mobile (Moving Average) q : $MA(q)$	37
2.3.1.1Caractéristiques des corrélogrammes.....	38
2.3.2 Processus Auto Régressif d'ordre p : $AR(p)$	39
2.3.2.1 Caractéristiques des processus $AR(p)$	39
2.3.3 Processus mixte Auto Régressif Moyenne Mobile d'ordre p et q : $ARMA(p, q)$	40
2.3.3.1 Caractéristiques des processus $ARMA(p, q)$	41
2.3.4 Processus mixte Auto Régressif Moyenne Mobile Intégré: $ARIMA$...	45
2.3.5 Processus Saisonnier Auto Régressif Moyenne Mobile Intégré d'ordre: $SARIMA$	46.

2.4 La méthodologie de Box & Jenkins :

2.4.1 Recherche de la représentation adéquate.....	47
2.4.1.1L'Identification.....	47
2.4.2 Estimation.....	59
2.4.3 Tests d'adéquation du modèle	50
2.4.4 Prévision.....	51

Chapitre 3 : application de la méthode de Box et Jenkins

3.1 Analyse descriptive :

3.1.1 Présentation des données.....	59
-------------------------------------	----

3.2 Application de la méthode de Box & Jenkins :

3.2.1 Identification et étude de la stationnarité.....	61
3.2.2 Estimation du modèle.....	73
3.2.3 Validation du modèle	75
3.2.4 Prévision	77

Conclusion générale.

Bibliographie.

Annexes.

المستخلص:

يشهد استهلاك الطاقة الكهربائية في السنوات الأخيرة تغيرات هامة حيث تولي شركة سون لغاز اهتمام كبير لهذه المادة الإستراتيجية وذلك من خلال بناء أنظمة تنبؤ بتطورات استهلاك الكهرباء بغية اتخاذ قرارات فعالة و رشيدة واجتناب اكبر قدر ممكن من الخسارة

لقد خلصت هذه الدراسة إلى استنتاج نموذج تنبؤي لاستهلاك الطاقة الكهربائية في مدينة سعيدة وهذا بتطبيق إحدى المناهج الكمية للتنبؤ وهي منهجية بوكس و جنكينز التي أثبتت فعاليتها في مجال التنبؤ قصير المدى لما تمتاز به من دقة ومرونة عاليتين في تحليل السلاسل الزمنية و النموذج المقترح هو ARMA (1, 2).

الكلمات المفتاحية :

استهلاك الطاقة الكهربائية, السلاسل الزمنية , الاستقرارية , التشخيص , التقدير والتنبؤ

Abstract:

In recent years, the electric energy consumption has changed widely, so the *SONALGAZ* Company gives a very big importance to this strategic material, and built an estimation programs for the evolution of electricity consumption in order to take an effective and rational decisions and to avoid the maximum possible loses.

This study gives as a result an effective estimation model for electric energy consumption in Saida town by the implication of one of the quantitative models, it's the *BOX* and *JENKINS* approach that proved its effectiveness in short run estimation for what it characterized by, its high flexibility and precision in analyzing the time series and the proposed model is ARMA (1, 2)

Key words:

Electric energy consumption, time series, stationary, diagnostic, prevision and estimation

Liste des tableaux

Tableau 2-1 typologie des modèles.....	32
Tableau 2-2 : résumé des propriétés des fonctions d'autocorrélation simple et partielle.....	41-42
Tableau 3.1 : Consommation d'électricité en GWH DP pour la série de la ville de SAIDA.....	60
Tableau 3-2 :fonction d'autocorrélation simple et partiel.....	63
Tableau 3.3 : Critères d'Akaike et Schwartz pour la série.....	65
Tableau 3.4 : test de Dickey-Fuller estimé par le modèle[3].....	66
Tableau 3.5 : test de Dickey-Fuller augmenté.....	66
Tableau 3-6 : test de Philips-perron.....	67
Tableau 3.7 : représente la série LELC après retirer la saisonnalité.....	69
Tableau 3.8 : Fonction d'autocorrélation de la série désaisonnalisé.....	70
Tableau 3.9 : Test de Dikey-Fuller augmenté pour le troisième modèle.....	71
Tableau 3.10 : test ADF première différence pour le modèle [3].....	72
Tableau 3.11 : Résultat d'estimation du modèle identifier de la série (LELCSA).....	74
Tableau 3.12- Préviation par la méthode de Box & Jenkins sur la consommation d'électricité.....	79

Liste des figures

Figure 2-1 : représente la tendance.....	15
Figure2-2 : représente la composante saisonalité.....	16
Figure 2-3 :représente la composane résiduelle.....	17
Figure 2.4 : Exemple de schéma additif.....	19
Figure 2.5 : Exemple de schéma multiplicatif.....	19
Figure 3-1 : série consommation d'électricité ville de Saida.....	61
Figure 3-2 : Evolution de la consommation d'électricité depuis 2003	61
Figure 3-2 : Evolution de la consommation d'électricité avec LOG a partir d'octobre 2003.....	68
Figure3.3 :la série désaisonnalisé (LELCSAt).....	69
figure 3.4 : représente le graphe des résidus.....	75
Figure 3.5 : test de normalité du résidus.....	76

Liste des schémas

Schéma 1.1 : Macro structure de la DR-SAÏDA.....	09
Schéma 2-1 : Stratégie simplifiée des tests de racine unitaire.....	35
Schéma 2-2 : les étapes de la méthodologie de Box et Jenkins....	55

ANNEXES

Null Hypothesis: ELC has a unit root

Exogenous: Constant

Lag Length: 11 (Automatic - based on SIC, maxlag=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	0.913624	0.9954
Test critical values:		
1% level	-3.495021	
5% level	-2.889753	
10% level	-2.581890	

Tableau 3.4 : test de Dickey-Fuller estimé par le modèle[1]

Null Hypothesis: ELC has a unit root

Exogenous: None

Lag Length: 11 (Automatic - based on SIC, maxlag=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	2.975354	0.9992
Test critical values:		
1% level	-2.587607	
5% level	-1.943974	
10% level	-1.614676	

Tableau 3.4 : test de Dickey-Fuller estimé par le modèle[2]

Tableau 3-6 : test de Philips-Perron avec les trois modèles

L'HYPOTHESE H_0 : LA SERIE ELC EST STATONNAIRE				
Les Modèles	PP T statistique	Valeurs critères du test		
		1%	5%	10%
Modèle 03	- 13 .20047	- 4.040532	-3.449716	-3.150127
Modèle 01	-10.74567	-3.488585	-2.886959	-2.580402
Modèle 02	-1.398571	-2.585405	-2.943626	-1.614866

Référence par la langue française :

1-BOURBOUNNAIS R., Econométrie, 3ème édition. Ed. DUNOD, Paris, 2000.

2-BOURBOUNNAIS R., TERRAZA M., Analyse des séries temporelles en Economie, Ed. PUF, juin 1998.

3-BOURBOUNNAIS R., TERRAZA M., Analyse des séries temporelles, Application à l'économie et à la gestion, Ed. DUNOD, Paris, 2004.

4-CARNOT, N. et TISSOT B. (2002). La prévision économique, Economica.

5-COURS DE SERIES TEMPORELLES .THEORIE ET APPLICATIONS
VOLUME 1· Introduction à la théorie des processus en temps discret.
Modèles ARIMA et méthode Box & Jenkins - ARTHUR CHARPENTIER
UNIVERSITÉ PARIS DAUPHINE.

6-Distribution of the Estimators for Autoregressive Time Series with a Unit Root,” Journal of the American Statistical Association.

7-Econometrie des séries temporelles- cour- Héliène Hamisultane consulter au site: hélènehamisultane. voila.net/travaux/SERIES_TEMPORELLES.pdf.

8-Jarque, Carlos M. & Anil K. Bera (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals.

9-PY B., Statistique descriptive, Nouvelle méthode pour bien apprendre et réussir. 4ème édition. Ed. ECONOMICA, Paris, 1999.

10-REGIS BOURBONNAIS -Manuel Et Exercices Corrigeés Économétrie 6e édition Dunod-paris2007.

11-TIME-SERIES FORECASTING by Chris Chatfield Reader in Statistics
Department of Mathematical Sciences
University of Bath, UK © 2000 by Chapman & C Hall/CRC.

مراجع باللغة العربية

1 الأساليب الكمية التطبيقية في إدارة الأعمال المؤلف العلمي الثلاثي
د- عبد الحميد عبد المجيد البلداوي
د- نجم عبد الله الحميدي دار وائل للنشر

2 تومي صالح مدخل لنظرية القياس الاقتصادي دراسة نظرية ج 2 الطبعة الثانية 2010

3 مولود حشمان السلاسل الزمنية و تقنيات التنبؤ القصير المدى

ديوان المطبوعات الجامعية الطبعة الثالثة 2010

Sites et adresses web :

1- www.sonelgaz.dz

2- www.wikipedia.org

3- <http://www.unige.ch/ses/sococ/eda/bernard/box.pdf>

Dur, dur ! Les séries chronologiques, RAPACCHI B., Centre Interuniversitaire de Calcul de Grenoble, 1993.

4- <http://lumimath.univ-mrs.fr/~boutahar/AE2PRO.pdf>

Master M2-AE2 PRO, Econométrie Bancaire et Financière, Analyse des séries chronologiques, BOUTAHAR M., octobre 2005.

5- <http://www.unctad.org/infocomm/francais/gaz/descript.htm>

MEMOIRES :

1-PASTEELS, Jean-Michel (1997). Extrait de la thèse de doctorat; "L'expertise dans la prévision à court terme de variables économiques: contributions méthodologiques et empiriques". Université Libre de Bruxelles

2التنبؤ بسعر البترول الخام بالجزائر باستعمال طريقة بوكس و جنكينز

من إعداد الطلبة حميدي زقاي -رماس مراد-درقاوي عبد القادر

تحت اشراف د-بن قدور علي

Chapitre 01:
Prévision de la
consommation du gaz
naturel

Chapitre 2 :
étude analytique des
séries temporelles

Chapitre 03 :
Application de la
méthode de
Prévision

Conclusion générale :

Bibliographie

Annexes

