

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR



Université Kasdi Merbah
- Ouargla -

N° d'ordre :
N° de série :

Faculté des Mathématiques et des Sciences de la Matière
Département de mathématiques

Mémoire présent en vue de l'obtention du
Diplôme de master
Spécialité : mathématiques
Option : probabilité et statistique

Présenté par :
SEBAA Farah

Thème

Prévision par la méthode de Box–Jenkins
Sur les données du transport d'électricité (GRTE/DTE-HM)

Soutenu publiquement le : Juin 2018

Devant le jury composé de :

Me. Badidja Université de KASDI Merbah - Ouargla Président
Me. Bahadi.H Université de KASDI Merbah - Ouargla Examineur
Mel. Halil.R Université de KASDI Merbah - Ouargla Rapporteur

Remerciements

Nous tenons à exprimer nos plus vifs remerciements à dieu tout puissant pour la volonté le courage, la santé, la patience qu'il nous a donnée durant toutes ces longues années d'étude que nous puissions arriver là.

Nous remercions toute la noble famille enseignante, et particulièrement l'ensemble des enseignants de la département de mathématiques, ainsi que notre encadreur Mlle :H. RAC HIDA
Nos profonds remerciements à tout ceux qui ont contribué de près ou de loin de loin à la réalisation de ce travail.

Farah

Dédicace

*Je dédie le fruit de ce modeste travail pour
Le soleil qui a éclairé ma vie et, qui m'a montré l'amour avec tous ses sens
et qui m'a entouré par ses belles sentiments et sa tendresse « mon chère
Belkhir*

*A mes très chers parents qui je ne saurais jamais remercier ; Assez pour les
sacrifices qu'ils ont consentaient pour moi, ainsi que chaque instant ; De
bonheur de qu'elle m'a procuré.*

A ma fille(Ouâla) .

A mon fils(Ouâil)

Amâ Bell-mâr(S) et mon Beau-pâr(S) .

Mes chères frères, chaque un à son nom (H ,N ,I, I, M, I)

A mes chères sœurs, chaque une à son nom : (T, N ,S, S) et son famille

A mes sours mari (H, M ,S, S, S) et(O, N, N)

A mes frères mari (K' et sa femme A) et H ,H et N

A mes collègues

Aux familles Benâti, Sebââ, Lkhabichi.

Enfin à A toute personne que j'ai connu et j'ai aimé.

S. Farah

Table des matières

Dédicace	I
Remerciement	II
Table des matières.....	III
liste des figures et tableaux	IV
Notation et abréviations	V
Introduction Générale	01
Chapitre I : Présentation de l'entreprise « GRTE » et prévision de la consommation d'électricité.....	02
I.1 Histoire de Sonelgaz.....	03
I.2 Organisation de sonelgaz	04
I.3 Présentation de GRTE.....	05
I.3.1- Consistance du réseau.....	06
I.3.2-Organisation de GRTE	06
I.3.3- Organigramme de la région Hassi Messaoud Messaoud.....	07
I.3.4- Réseau Electrique De La Région De Transport De Hassi Messaoud....	07
I.3.5- Généralité sur le transport de l'électricité.....	08
1- Ligne à haute tension	08
2- Poste électrique.....	08
I.4 Consommation d'électricité	08
I.5 Solution technique adopté à la prévision de la consommation électrique.....	08
I.6 Les concepts de base des prévisions	09
Chapitre II : Généralités sur Série temporelle	10
II.1 Définition	11
II.2 Composants d'un série temporelle.....	11
II.3 Exemples de séries temporelles	11
II.4 Quelques types de décomposition.....	12
II.4.1 Schéma de sélection	14
II.5 Détection de la tendance et saisonnalité.....	14
II.6 Opérateurs de retard et différenciation.....	16
II.7 Modélisation des séries chronologique.....	18
II.7.1 Processus stochastique.....	18
II.7.2 Processus stationnaire.....	18
II.7.3 Autocorrélation simple et partielle.....	19
II.7.3.1 La fonction d'autocorrélation et d'autocovariance.....	19
II.7.3.2 La fonction d'autocorrélation.....	19
II.7.3.3 la fonction d'autocorrélation partielle.....	21
II.7.4 Bruit blanc	22
II.7.5 la non stationnarité.....	23
II.7.6Modèle autorégressif.....	24
II.7.7 Modèle moyenne mobile.....	27
II.7.8Modèle autorégressif – Moyenne mobile.....	31
II.7.8.1 Modèle ARMA.....	31
II.7.8.2 Modèle ARIMA et SARIMA.....	32
II.8 Méthodologie de Box & Jenkins.....	34
II.8.1 Identification du modèle.....	36

II.8.2 Sélection de modèle	37
II.8.3 Estimation.....	38
II.8.4 Validation.....	39
Chapitre III : Application de la méthode de Box et Jenkins	44
III.1. Analyse du graphique et de corrélogramme	46
III.1.1 Représentation graphique	46
III.1.2 Analyse du corrélogramme de la série ELEC	47
III.2 Analyse de saisonnalité	48
III.2.1 Graphique par saison	48
III.2.2 Désaisonnaliser la série	48
III.3 stationnarisation de la série.....	51
III.4 spécification et estimation du modèle.....	53
III.5 Validation du modèle SARIMA (1,1 ,12)	54
III.6 Prévision.....	56
Conclusion Générale	57
Bibliographie.....	58

Liste des figures et tableaux

Figure I.1 : Filiales et participation SONELGAZ.....	05
Figure I.2 : Filiales Organisation de GRTE.....	06
Figure I.3 : Organigramme de la région Hassi Messaoud.....	07
Table II.1.Calculs des moyennes par année et par période.....	15
Table II.2.Analyse de la variance pour détecter une saisonnalité et/ou une tendance..	16
Figure II.1 : la méthode de Bob-Jenkins.....	35
Figure II.2 : stratégie simplifier des racines unitaires	37
Tableau III.1: transport d'électricité par GRTE/ DTE-HM (2008-2016)	45
Figure III.1: représentation graphique de la série brute (ELEC).....	46
Figure III.2: Corrélogramme de la série (ELEC).....	47
Figure III.3: Graphe de la série (ELEC) par saison	48
Figure III.4: Graphie de la série brute (LELECSA).....	49
Figure III.5: Corrélogramme de la série (LELECSA).....	50
Tableau III.2 : Test de Dickey –Fuller augment (avec constant et tendance) de la série (LELECSA)	51
Tableau III.3 : Test de Dickey –Fuller augment (avec constant) de la série (LELECSA)	52
Tableau III.4 : Estimation des paramètres du modèle SARIMA (1,0,12).....	53
Tableau III.5: Corrélogramme des résidus	54
Figure III.6:Caractéristique statistique des résidus de la modélisation ARMA.....	55
Tableau III.6: Valeurs Prévisionnelle et réelle en 2017 (Unité : KWh)	56

Notation et abriviations

ADF	Test de Dickey Fuller Augmenté
ARIMA	Autorégressif moyenne mobile intégré
ARMA	Autorégressif moyenne mobile
BB	Bruit blanc
BJ	Quantité de Jarque-Bera
Cov	Covariance
Ds	Non stationnarité stochastique
E	Espérance
FAC	Fonction d.autocorrélation
FACP	Fonction d.autocorrélation partielle
MA	Moyenne mobile
MCO	Méthode de moindre carée
TS	Non stationnarité déterministe
GRTE	Société de Gestion du Réseau de Transport de l'Electricité
DTE	Direction regionale transport d'electricite

Introduction générale

Une série chronologique, ou série temporelle, est une série d'observations ordonnées chronologiquement. Elles se rencontrent naturellement dans une grande variété de domaines. Notamment l'économie, la finance, l'écologie, le transport, la démographie...

Les objectifs d'étude sont multiples. La prévision est sans doute le but le plus fréquent. Il s'agit de prévoir les valeurs futures d'une variable grâce aux valeurs observées dans le passé de cette même variable.

La méthode de prévision à court terme de box & Jenkins est la méthode la plus élaborée, sa caractéristique est d'effectuer de sélectionner une modélisation stochastique de type ARIMA pour faire les prévisions.

Dans ce travail nous avons appliqué la méthode de Box et Jenkins pour des données mensuelle de l'entreprise GRTE concernant le transport de l'électricité et cela afin d'obtenir des prévisions.

Notre travail est réparti en trois chapitres dont voici le contenu.

Le premier chapitre est consacré à la présentation de l'entreprise La Société de Gestion du Réseau de Transport de l'Electricité « GRTE ».d'où nous avons procuré les données nécessaire à la suite de ce travail.

Chapitre I

Présentation de l'entreprise « GRTE »
Et
Prévision de la consommation
d'électricité

I.1 Histoire de Sonelgaz

Sonelgaz est la Société Nationale de l'Electricité et leur distribution et du GAZ. Elle réalise la production de l'électricité ainsi que le transport et la distribution du gaz. [13]

Elle a été créée en 1947 sous le nom d'EGA (Electricité et Gaz d' Algérie).

Par ordonnance n°6959 du 26 juillet 1969 parue dans le journal officiel du 1^{er} Août 1969, SONELGAZ est créée en substitution à EGA dissout par ce même décret. L'ordonnance lui assigne pour mission générale de s'intégrer de façon harmonieuse dans la politique énergétique intérieure du pays. Le monopole de la production, du transport, de la distribution, de l'importation et de l'exportation de l'énergie électrique attribué à Sonelgaz a été renforcé. De même, Sonelgaz s'est vue attribuer le monopole de la commercialisation du gaz naturel à l'intérieur du pays, et ce pour tous les types de clients (industries, centrales de production de l'énergie électrique, clients domestiques). Pour ce faire, elle réalise et gère des canalisations de transport et un réseau de distribution

Suite au décret exécutif n°91-475 du 14 Décembre 1991, change de nature juridique et devient un Etablissement Public à Caractère Industriel et Commercial. Le décret exécutif n°95-280 du 17 Septembre 1995 confirme la nature de Sonelgaz en tant qu'établissement public à caractère industriel et commercial. Sonelgaz est placée sous tutelle du Ministre chargé de l'énergie. Elle est dotée de la personnalité morale, jouit de l'autonomie financière et est régi par les règles de droit public dans ses relations avec l'Etat. Le même décret définit les missions de Sonelgaz : Assurer la production, le transport et la distribution de l'énergie électrique et assurer la distribution publique du gaz dans le respect des conditions de qualité, de sécurité et au moindre coût, dans le cadre de sa mission de service public.

En juin 2002, Sonelgaz passe au statut d'une Société Par Actions (SPA) suite à la promulgation de la loi du 05 février 2002 (loi n°02-01) relative à l'électricité et à la distribution du gaz par canalisations. Cette loi permet l'ouverture du marché de l'électricité. L'Etat demeure l'actionnaire majoritaire de Sonelgaz SPA.

Comme conséquence à cette nouvelle loi, une institution importante a été créée : l'Autorité de régulation qui est totalement autonome et indépendante pour l'électricité et la distribution du gaz par canalisations dénommée la Commission de Régulation de l'Electricité et du Gaz (CREG).

Les principales missions de la CREG sont :

- Favoriser la compétitivité des secteurs de l'électricité et du gaz ;
- Assurer la transparence des activités dans le seul intérêt des opérateurs et des consommateurs
- Surveiller et contrôler le respect des lois et règlements relatifs aux marchés de l'électricité et du gaz

I.2 Organisation de sonelgaz

SONELGAZ vit, depuis quelques années, une phase particulièrement importante de son histoire. Désormais, la restructuration de Sonelgaz, suite à l'avènement de la loi N°01.02 du 05 février 2002 s'est achevée avec la création de l'ensemble des filiales. [13]

La nouvelle organisation validée en février 2017, a érigé Sonelgaz en Groupe industriel composé de 16 filiales . Ainsi, ses filiales métiers de base assurent la production, le transport et la distribution de l'électricité ainsi que le transport et la distribution du gaz par canalisations.

On compte:

1. La Société de Production de l'Electricité (SPE),
2. Sharikat Kahraba wa takat moutadjadida (SKTM)
3. La Société de l'Engineering de l'électricité et du Gaz (CEEG)
4. La Société de Gestion du Réseau de Transport de l'Electricité (GRTE),
5. La Société de Gestion du Réseau de Transport Gaz (GRTG),
6. L'Opérateur Système électrique (OS), chargée de la conduite du système Production / Transport de l'électricité,
7. La Société de Distribution de l'électricité et du gaz (SDC)

Les sociétés travaux du Groupe Sonelgaz, Kahrif, Kahrakib, Etterkib, Inerga et Kanaghaz, sont spécialisées dans le domaine de la réalisation des infrastructures énergétiques (engineering, montage industriel, réalisation de réseaux...) et c'est grâce à ces sociétés que l'Algérie dispose

aujourd'hui d'infrastructures électriques et gazières répondant aux besoins du développement économique et social du pays.

Par ailleurs, les filiales métiers périphériques sont en charge d'activités annexes, telles que la maintenance d'équipements énergétiques, la distribution de matériel électrique et gazier, le transport et la manutention exceptionnels, ... etc.

De plus, Sonelgaz détient également des participations dans des sociétés, dont le métier est en rapport avec le domaine de l'électricité et du gaz.

Grâce à sa ressource humaine formée et qualifiée, le Groupe occupe une position privilégiée dans l'économie du pays en tant que responsable de l'approvisionnement de plus de six millions de ménages en électricité et de trois millions en gaz naturel, soit une couverture géographique de plus de 99% en taux d'électrification et 59% pour la pénétration gaz.

Filiales et Participations

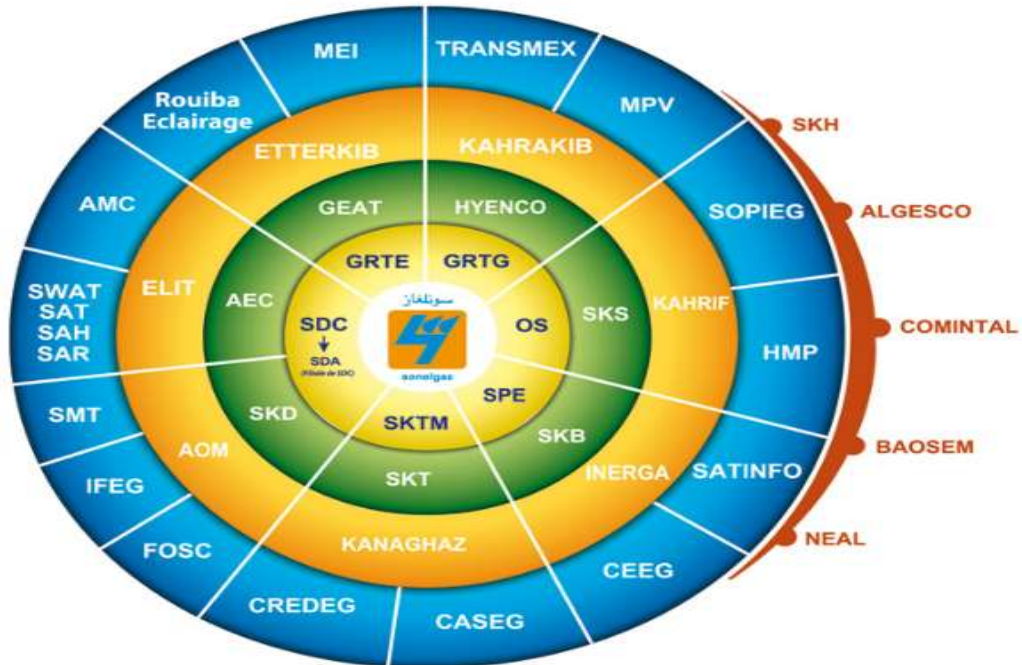


Figure I.1 : Filiales et participation SONELGAZ

I.3 Présentation de GRTE

La loi n° 02-01 du 5 février 2002 relative à l'électricité et à la distribution du gaz par canalisations introduit l'ouverture progressive du marché électrique et impose la séparation des activités de Transport, Production, Distribution et règlemente l'accès des Tiers au réseau.

Le Gestionnaire du Réseau de Transport de l'Electricité, dénommée GRTE.Spa, filiale du Groupe SONELGAZ, a été créée le 1er janvier 2004, conformément à cette loi et enregistrée sous l'appellation « SONELGAZ Transport de l'Electricité, GRTE Spa » en septembre 2004. A cet égard, le Gestionnaire dispose d'une autorisation d'exploiter le réseau de transport délivrée par le ministère de l'énergie et des Mines, après avis de la Commission de Régulation de l'électricité et du gaz (CREG). [14]

La raison sociale de l'entreprise devient en février 2009, « Société Algérienne de Gestion du Réseau de Transport de l'Electricité, dénommée GRTE Spa », Société par Actions au capital de 40.000.000.000 DA. GRTE est chargé de **l'exploitation, de la maintenance et du développement du réseau de transport de l'électricité**, en vue de garantir une capacité adéquate par rapport aux besoins de transit et de réserve. Conformément à la loi n°02-01 du 5 février 2002, GRTE a un monopole naturel sur le transport d'électricité. Ce monopole lui garantit un revenu requis qui provient de la rémunération sur le transit d'énergie, dont la tarification est fixée par la CREG et qui est le même sur l'ensemble du territoire national. Actuellement, le tarif est fixé à 0,66 DA le kWh.

La gestion du système production transport d'électricité SPTE est assurée par l'Opérateur Système Electrique, une société filiale du groupe Sonelgaz, qui doit veiller en permanence au maintien de l'équilibre offre – demande d'énergie électrique. GRTE a mis en place un système de management intégré (SMI) en vue d'une certification totale progressive. La certification ISO 9001 Qualité s'est concrétisée en janvier 2010.

I.3.1- Consistance du réseau

GRTE assure ses activités à travers des directions centrales et cinq directions de Régions Transport de l'Electricité : Alger, Oran, Sétif, Annaba, Hassi Messaoud. Ces régions, à travers 24 services de transport répartis sur le territoire national assurent une maintenance de proximité et la relation directe avec les clients. [14]

Le réseau de transport est un ensemble d'ouvrages lignes et postes haute tension (60, 90, 150, 220 et 400 kV) équipé d'un réseau de télécommunication pour la surveillance, le contrôle et la télécommande des équipements.

Il est constitué d'un réseau interconnecté au nord du pays, avec des interconnexions internationales (Tunisie et Maroc) et d'un réseau isolé au sud.

Les utilisateurs du réseau sont les centrales électriques, les sociétés de distributions de l'électricité et clients HT ainsi que pour les échanges internationaux à travers les interconnexions. GRTE exploite un réseau composé de :

- 25 147 km de lignes dont 2 547 km en 400 kV
- 283 postes (dont 12 en 400 KV) dotés d'une capacité de transformation de 48 806 MVA à travers 773 transformateurs et cabines mobiles
- Un réseau de fibre optique de 16 095,7 km. GRTE assure le transit pour la Société de distribution d'électricité (y compris les clients industriels qui sont clients de ces sociétés de distribution).

En 2013, GRTE a assuré un transit de 52 879 GWH et réalisé un chiffre d'affaires est de 35,554 Milliards de Dinars pour un investissement de 50,232 Milliards de Dinars.

I.3.2-Organisation de GRTE

GRTE est le gestionnaire d'unique et propriétaire du réseau de transport d'électricité. Il est chargé de l'exploitation, la maintenance et le développement du réseau de transport. [12]

L'organisation de GRTE se présente comme suit :

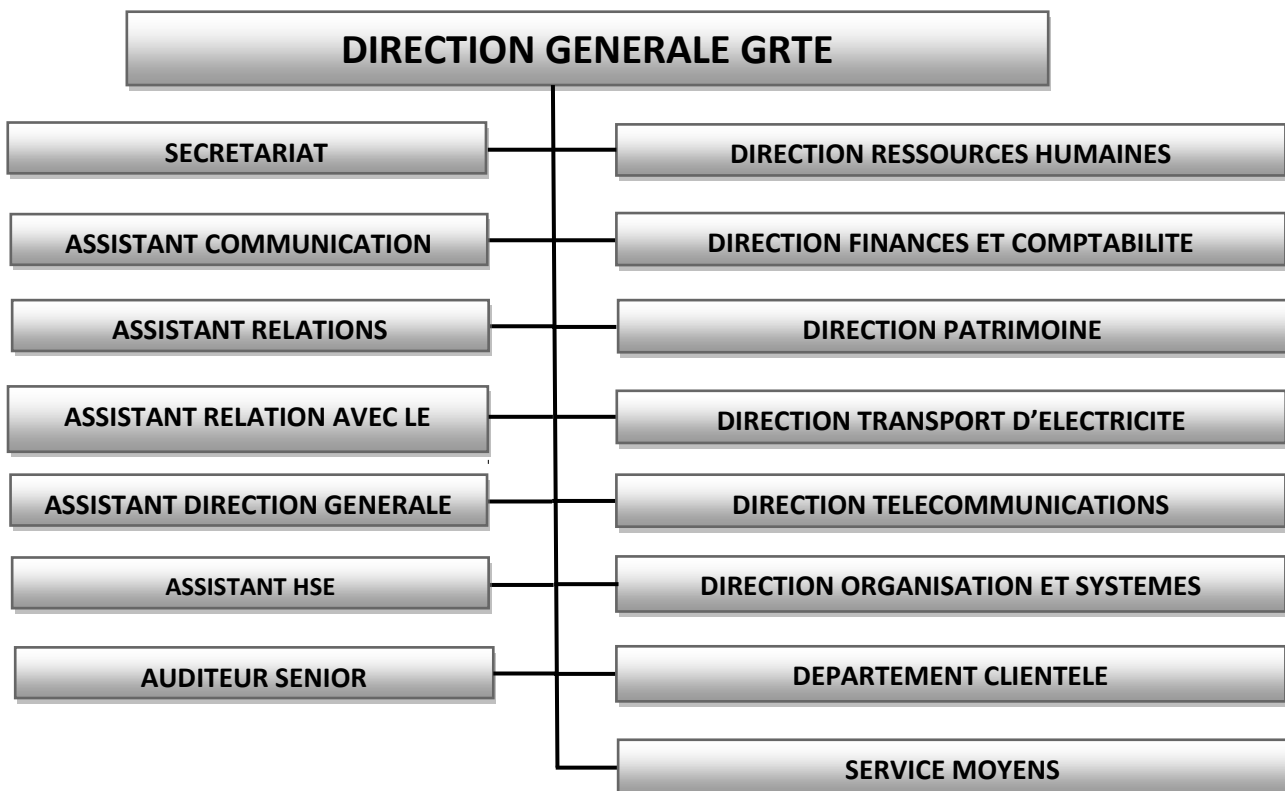


Figure I.2 : Filiales Organisation de GRTE

Afin de pouvoir gérer l'ensemble du réseau national du transport d'électricité, GRTE est organisé en cinq (05) directions régionales qui sont : Alger, Annaba, Sétif, Oran et Hassi Messaoud.

I.3.3- Organigramme de la région Hassi Messaoud

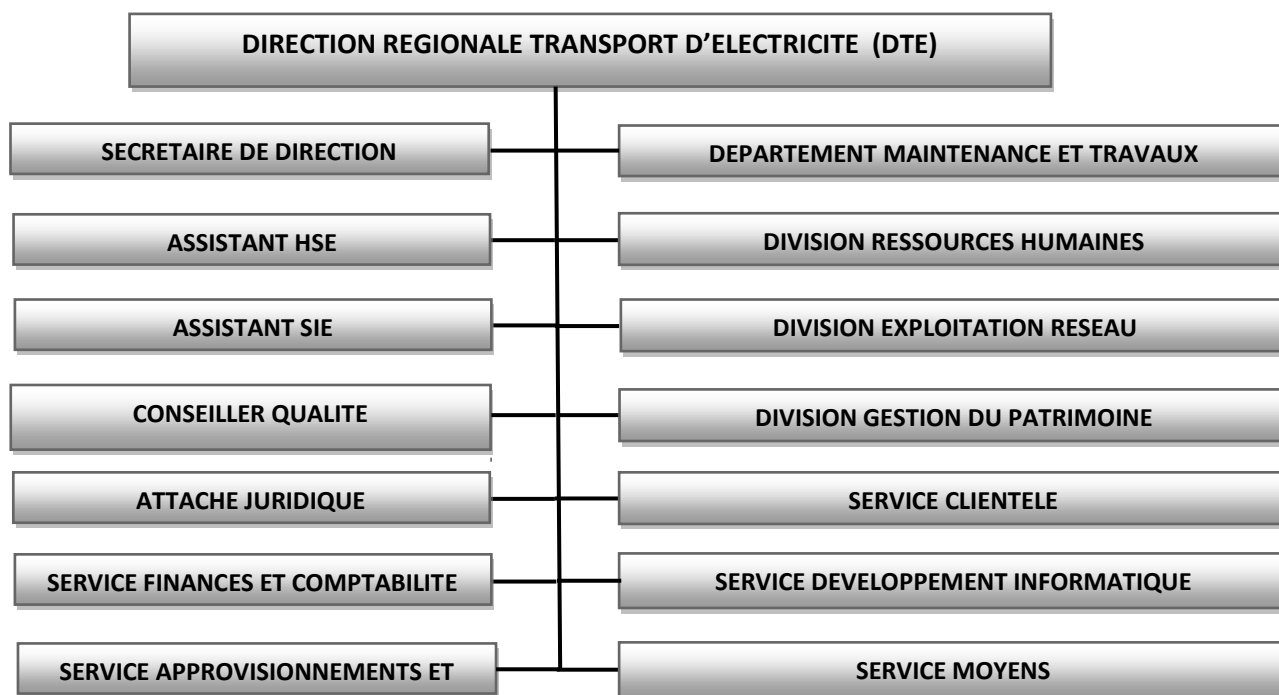


Figure I.3 : Organigramme de la région Hassi Messaoud [13]

I.3.4- Réseau Electrique De La Region De Transport De Hassi Messaoud

Le réseau électrique de Hassi Messaoud est constitué de :

Désignation	Unité	Situation 2015	Situation 2016	Situation 2017
Postes 400 kV	Nombre	02	03	04
Postes 220 kV		24	24	26
Postes 60 kV		11	12	12
Total Postes		37	39	42
Lignes 400 kV	km	325	663,8	1123
Lignes 220 kV		3037	2794,03	3095
Ligne 60 kV		784,96	820,36	821
Total Lignes		4147	4278,19	5039
Fibre Optique	km	2621	2667,75	3 589
Transformateurs	Nombre	88	94	108
Cabines Mobiles	Nombre	15	17	22

Les ouvrages sont interconnectés par des lignes électriques 400kV, 220kV ou 60kV. Comme le montre le schéma, le réseau électrique du pôle d'Adrar est isolé du réseau de Hassi Messaoud.

I.3.5- Généralité sur le transport de l'électricité

Les réseaux de transport sont à haute tension (HTB) (de 50 kV à 400 kV) et ont pour but de transporter l'énergie des grands centres de production vers les régions consommatrices d'électricité. Les grandes puissances transitées imposent des lignes électriques de forte capacité de transit, ainsi qu'une structure maillée (ou interconnectée). Les réseaux maillés garantissent une très bonne sécurité d'alimentation, car la perte de n'importe quel élément (ligne électrique, transformateur ou groupe de production) n'entraîne aucune coupure d'électricité si l'exploitant du réseau de transport respecte la règle dite du "N-1" (possibilité de perdre n'importe quel élément du réseau sans conséquences inacceptables pour les consommateurs).

Le réseau électrique est constitué non seulement de matériel haute tension (dit matériel de puissance), mais également de nombreuses fonctions périphériques telles que la téléconduite ou le système de protection et le plus important [14]:

1- Ligne à haute tension

La ligne à haute tension est le composant principal des grands réseaux de transport d'électricité.

Elle transporte l'énergie par l'intermédiaire de l'électricité de la centrale électrique au consommateur. Ces lignes sont aériennes, souterraines ou sous-marines, quoique les professionnels réservent plutôt ce terme aux liaisons aériennes.

2- Poste électrique

Les postes électriques Hautes Tension servent à transformer la tension, aiguiller le courant électrique et protéger les installations et les personnes.

I.4 Consommation d'électricité :

L'électricité ne se stockant pas, la production globale sur le réseau électrique doit, à tout moment, être strictement égale à la consommation. Toute modification de la demande ou de la production se répercute instantanément sur tout le système électrique. [14]

Prendre de bonnes décisions en terme de gestion d'un système électrique n'est guère possible si l'on ne dispose pas de techniques efficaces permettant d'en modéliser les aléas, de modéliser l'effet des différentes variables intervenant dans le problème, et d'en faire une prévision aussi précise que possibles.

Ce travail se limitera dans le contexte de la prévision à moyen terme, une prévision de la consommation Algérienne d'électricité considérée comme l'une des missions essentielles de SONELGAZ ; des missions qui ont pour objet d'évaluer les alternatives de développement du système de production électrique à mettre en œuvre pour satisfaire l'évolution croissante de la demande d'énergie électrique du pays.

I.5 Solution technique adopté à la prévision de la consommation électrique

Les techniques de prévision, si elles sont massivement utilisées par les différents acteurs du marché de l'électricité dans plusieurs pays, le sont aussi en climatologie, en qualité de l'air, en finance, dans le domaine commercial...; il n'en reste pas moins que la problématique de l'électricité présente bel et bien des spécificités, même si sa résolution technique reste classique. [14]

I.6 Les concepts de base des prévisions :

Les concepts de base des prévisions sont :

- composante,
- saisonnalité,
- horizon de prévision,
- fréquence de rafraîchissement.

La prévision étant en amont de la planification, les concepts de base de la prévision reposent essentiellement sur le planning. [14]:

La composante est de savoir quelle donnée doit être prévue. Elle se décide suivant ce que l'entreprise cherche à améliorer. Pour planifier ses ressources, l'élargissement du plan d'investissement pour une entreprise de production d'électricité a besoin de prévision en unité de demande de cette énergie (consommation).

Le degré de saisonnalité de la chronique affecte le choix des méthodes mathématiques à utiliser.

L'**horizon de prévision** est en relation avec la durée du plan. L'horizon de prévision est défini par la taille du planning qu'on veut en déduire (en unités temporelles).

La méthode de prévision s'appuie sur le caractère cyclique de la consommation. On compte trois horizons identifiant trois cycles temporels :

- prévision à long terme : un cycle annuel (pointe de consommation annuelle en décembre et creux au 15 août),
- prévisions à moyen terme : un cycle hebdomadaire (5 jours ouvrables avec une consommation globalement stable et le week-end où la consommation diminue).
- prévisions à court terme : un cycle journalier. C'est un horizon de prévision opérationnel. Les prévisions sont utilisées pour prendre des décisions continues sur le planning. C'est un horizon stratégique. Les prévisions sont utilisées pour décider de rentrer dans un nouveau marché, établir un programme de développement à long terme.

Si l'entreprise décide d'utiliser un modèle de prévision différent pour chaque niveau temporel de la pyramide des temps, l'**intervalle de temps** ou la maille de temps entre deux périodes dépend de l'unité temps des plannings. Si les plannings se font par jour de semaine, les données prévisionnelles sont au jour. Si l'entreprise décide d'utiliser qu'une seule prévision pour la restituer sur différentes échelles de temps, il lui faut trouver le niveau d'agrégation temporel lui permettant de répartir la prévision de façon homogène sur les mailles de temps supérieures et inférieures à celle choisie pour la prévision.

Plus élevée est la **fréquence de rafraîchissement**, meilleures seront les prévisions.

Chapitre II

Généralités sur Série temporelle

II.1 Définition

Une série temporelle (ou chronologique) est une suite d'observations $\{x_1, x_2, \dots, x_n\}$ indexée par le temps. On supposera qu'il s'agit d'une réalisation d'un processus X , c'est à dire d'une suite X_i de variables aléatoires. [10] [11]

Une série temporelle est généralement constituée de plusieurs éléments.

II.2 Composants d'une série temporelle

- **Tendance** : représente l'évolution à long terme de la série.

Exemples:

Croissance économique, évolution climatologique à long terme (cyclique ou non)

- **Saisonnalité** : évolution se répétant régulièrement tous les ans ou tous les mois ou toutes les semaines.

Exemples :

- En météorologie, températures plus faibles en hiver qu'en été.
- En économie, saisonnalité induite par les périodes de vacances, les périodes de fêtes, le climat...
- Chiffre d'affaire d'un magasin.

- **Composante stationnaire (ou résiduelle)** : ce qui reste lorsque l'on a enlevé les autres composantes. Décrit l'évolution à court terme de la série.

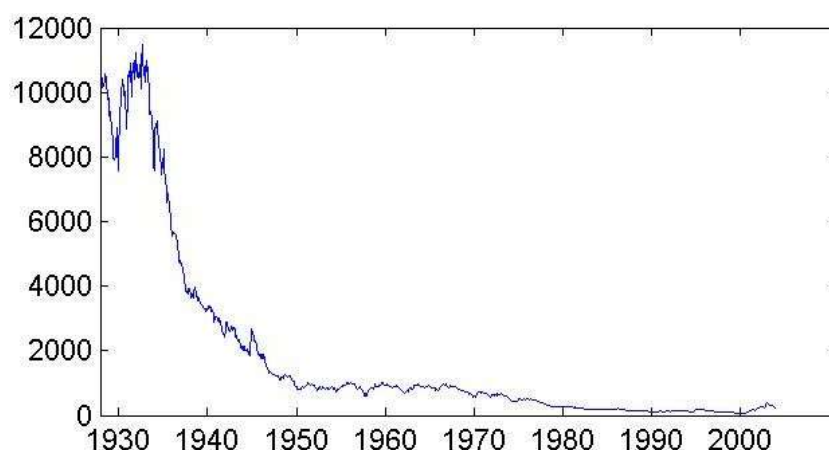
La notion de série temporelle stationnaire définie plus précisément dans la suite. Cette hypothèse jouera un rôle fondamental dans la suite, et remplacera l'hypothèse usuelle des v.a i.i.d. (ici, il peut exister une dépendance entre deux valeurs successives prises par la série observée).

Le modèle le plus courant consiste à supposer que la série initiale s'écrit sous la forme (modèle additif) composante stationnaire.

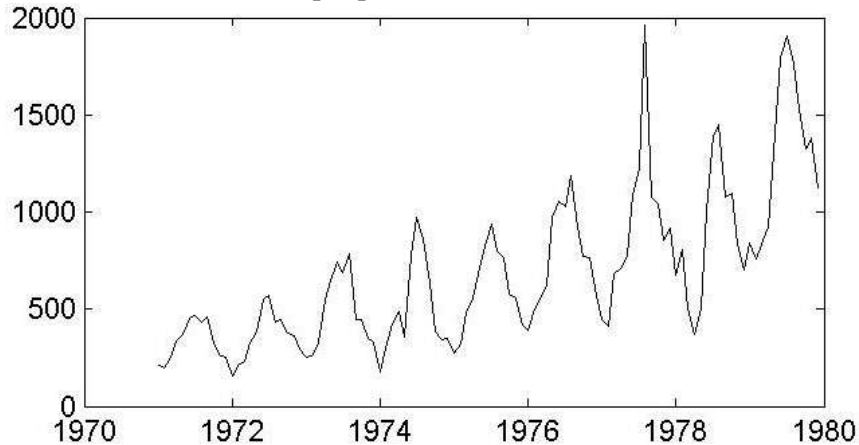
II.3 Exemples de séries temporelles

Economie

Évolution du cours du Dow Jones entre 1928 et 2004, données mensuelles [11]

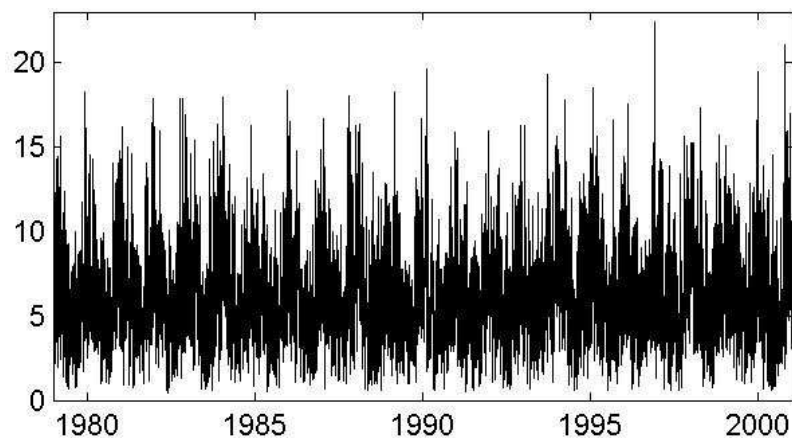


Production de poissons, en milliers de francs, en Finistère nord (Brest, Morlaix, Paim-pol) entre 1971 et 1979, données mensuelles [11]



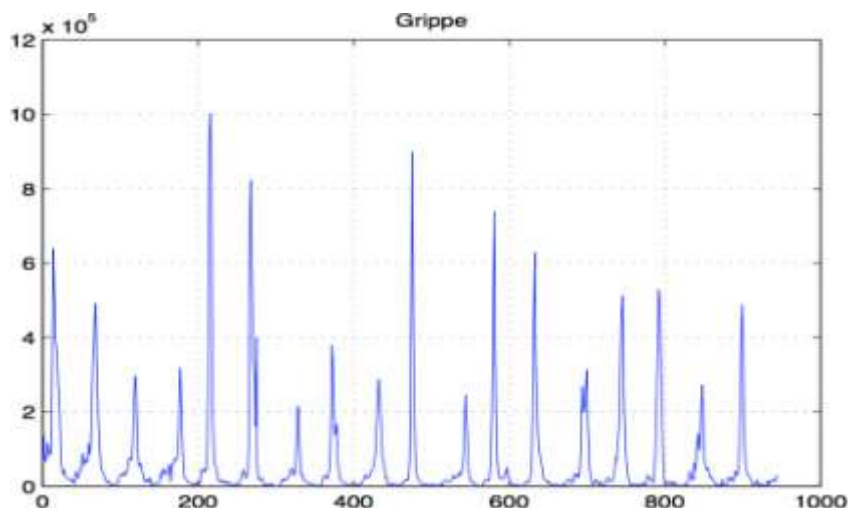
Environnement

Évolution de l'intensité du vent, en m/s, au large de la Bretagne entre 1979 et 2001, données journalières. [11]



Santé

Evolution des cas de grippe en France, par semaine, de 1984 (semaine 44) à 2002 (semaine 50), soit 945 valeurs. [11]



II.4 Quelques types de décomposition

On propose trois types de décomposition : [04] [10]

-Le modèle additif:

$$y_i = f_i + s_i + e_i, \quad 1 \leq i \leq n.$$

-Le modèle multiplicatif:

Revient à un modèle additif pour le log des données.

$$\text{Log}(y_i) = f_i + s_i + e_i \iff y_i = e^{f_i} e^{s_i} e^{e_i}, \quad 1 \leq i \leq n.$$

Là encore, on impose que la somme des facteurs saisonniers soit nulle : $\sum_{j=1}^p s_j = 0$

Remarque:

1. On choisit ce modèle si la courbe qui uni les maximums des observations s'éloigne ou s'approche de celle qui uni les minimums des observations, i.e. si les amplitudes des fluctuations dépendent du niveau.
2. Le modèle multiplicatif est souvent utilisé pour des données de type économique.

-Modèles mixtes:

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

$$y_i = f_i(1 + s_i) + e_i = f_i + f_i s_i + e_i, \quad 1 \leq i \leq n.$$

Remarque:

Une fois un modèle est obtenu, il peut être utilisé pour la prédiction de valeurs futures.

II.4.1 Schéma de sélection

La saisonnalité d'une chronique peut parfois être influencée par l'extra-saisonnier et/ou les résidus c'est de l'existence de ces interactions que sont nés le schéma de décomposition des série chronologique : additif, multiplicatif et mixte. [09] [01]

Nous présentons une simple technique pour sélectionner les schéma

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.

II.5 Détection de la tendance et saisonnalité

Analyse de la variance et test de Fisher

Le test de Fisher à partir de l'analyse de la variance permet de pallier ces deux inconvénients.

Ce test suppose la chronique sans tendance ou encore sans extra-saisonnalité.

Dans le cas contraire cette composante est éliminée par une régression sur le temps (extra-saisonnalité déterministe), ou par une procédure de filtrage (extra-saisonnalité aléatoire). [09]

Soit :

N le nombre d'années, p le nombre d'observations (la périodicité) dans l'année (trimestre $p=4$, mois $p=12$, etc.).

X_{ij} la valeur de la chronique pour la i ème année ($i = 1 \dots, N$) et la j ème période ($j = 1 \dots, p$) supposée telle que $x_{ij} = m_{ij} + e_{ij}$; les e_{ij} les résidus considérés comme aléatoires formés d'éléments indépendants : $e_{ij} \rightarrow N(0; \sigma^2)$

Les m_{ij} sont les éléments d'une composante de la chronique qui s'écrivent : $m_{ij} = a_i + b_j$ avec b_j qui mesure l'effet période en colonne du tableau et a_i qui mesure l'effet année en ligne du tableau.

Deux effets absents sont testés contre deux effets significativement présents :

- si l'effet période (ici trimestre) est significatif, la série est saisonnière ;
- si l'effet année est significatif, ceci suggère deux interprétations.

- La chronique de départ n'a pas été transformée, elle possède alors des paliers horizontaux.
- La chronique a été transformée, des changements de tendance existent dans la chronique.

Le déroulement du test est le suivant :

a. Calcul de la variance totale du tableau. Soit S_T la somme totale des carrés :

$$S_T = \sum_{i=1}^N \sum_{j=1}^P (X_{ij} - \bar{X}_{..})^2$$

Avec $\bar{X}_{..} = \frac{1}{N \times P} \sum_{i=1}^N \sum_{j=1}^P X_{ij}$: moyenne générale de la chronique sur les $N \times p$ observations.

$\bar{X}_i = \frac{1}{P} \sum_{j=1}^P X_{ij}$: moyenne de l'année i

$\bar{X}_j = \frac{1}{N} \sum_{i=1}^N X_{ij}$: moyenne de la période j

Nous obtenons :

$$\begin{aligned}
 S_T &= \sum_{i=1}^N \sum_{j=1}^P (X_{ij} - \bar{X}_{..})^2 \\
 &= \sum_{i=1}^N \sum_{j=1}^P (X_{ij} - \bar{X}_i + \bar{X}_i - \bar{X}_{ij} + \bar{X}_{ij} + \bar{X}_{..} - \bar{X}_{..} - \bar{X}_{..})^2 \\
 &= \sum_{i=1}^N \sum_{j=1}^P [(\bar{X}_i - \bar{X}_j) + (\bar{X}_i - \bar{X}_{..}) + (X_{ij} - \bar{X}_i - \bar{X}_j + \bar{X}_{..})]^2 \\
 &= \sum_{i=1}^N \sum_{j=1}^P (\bar{X}_i + \bar{X}_{..})^2 + \sum_{i=1}^N \sum_{j=1}^P (\bar{X}_j + \bar{X}_{..})^2 + \sum_{i=1}^N \sum_{j=1}^P (\bar{X}_{ij} - \bar{X}_i - \bar{X}_j + \bar{X}_{..})^2 \\
 &= p \sum_{i=1}^N (\bar{X}_i + \bar{X}_{..})^2 + N \sum_{j=1}^p (\bar{X}_j + \bar{X}_{..})^2 + \sum_{i=1}^N \sum_{j=1}^p (\bar{X}_{ij} - \bar{X}_i - \bar{X}_j + \bar{X}_{..})^2
 \end{aligned}$$

$$S_T = S_A \text{ (année)} + S_p \text{ (période)} + S_R \text{ (résidus)}$$

Année/ période	1	.	.	.	J	.	.	.	p	Moyennes années
1	X ₁₁				X _{1j}				X _{1p}	$\bar{X}_i = \frac{1}{p} \sum_{j=1}^p X_{ij}$
.	.				.				.	
.	.				.				.	
i.	X _{i1}				X _{ij}				X _{ip}	
.	.									
N	X _{N1}	.	.	.	X _{N1}	.	.	.	X _{Np}	
Moyennes périodes					$\bar{X}_i = \frac{1}{N} \sum_{i=1}^N X_{ij}$					$\bar{X}_{..} = \frac{1}{N \times p} \sum_{j=1}^p \sum_{i=1}^N X_{ij}$

Table II.1. Calculs des moyennes par année et par période

Somme des carrés	Degré de liberté	Désignation	Variance
$S_p = N \sum_{j=1}^p (\bar{X}_j - \bar{X}_{..})^2$	p-1	Variance périodes	$V_p = \frac{NS_p}{p-1}$
$S_p = p \sum_{j=1}^p (\bar{X}_i - \bar{X}_{..})^2$	N-1	Variance année	$V_A = \frac{S_A}{N-1}$
$S_p = \sum_{i=1}^N \sum_{j=1}^p (X_{ij} - \bar{X}_i - \bar{X}_j + \bar{X}_{..})^2$	(p-1)(N-1)	Variance résidu	$V_R = \frac{S_A}{(p-1)(N-1)}$
S_T	Nxp-1	Variance totale	$V_T = \frac{S_T}{N \times p - 1}$

Table II.2. Analyse de la variance pour détecter une saisonnalité et/ou une tendance

Nous utilisons les résultats du tableau précédente pour effectuer l'analyse de la variance de la série. A partir de ce tableau, nous pouvons construire les tests d'hypothèses.

b. Test de l'influence du facteur colonne (période, mois ou trimestre ; H_0 : pas d'influence). Fisher empirique $F_c = \frac{V_p}{V_R}$ que l'on compare au Fisher lu dans la table : F_{V_1, V_2}^a à $V_1 = p - 1$ et $V_2 = (N - 1)(p - 1)$ degrés de liberté.

Si le Fisher empirique est supérieur au Fisher lu dans la table, on rejette l'hypothèse H_0 , la série est saisonnière.

c. Test de l'influence du facteur ligne (année ; H_0 : pas d'influence). Fisher empirique $F_c = \frac{V_A}{V_R}$ que l'on compare au Fisher lu dans la table : F_{V_3, V_2}^a à $V_3 = (N - 1)(p - 1)$ degrés de liberté.

Si le Fisher empirique est supérieur au Fisher lu, on rejette l'hypothèse H_0 , la série est donc affectée d'une tendance.

II.6 Opérateurs de retard(B) et différenciation(Δ) :

Pour étudier les processus (et donc les séries chronologiques), on définit des opérateurs retard et de différenciation. [06]

- Opérateur de retard

L'opérateur B décale le processus d'une unité de temps vers le passé :

$$B(X_t) = X_{t-1} .$$

Si on applique h fois cette opérateur, on décale le processus de h unités de temps :

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

Opérateur différence

L'opérateur Δ fait la différence entre le processus et sa version décalée de une unité de temps. Cet opérateur se construit en utilisant l'opérateur précédent :

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

$$\Leftrightarrow \Delta = I - B$$

ou I est l'opérateur identité:

$$I(X_t) = X_t.$$

Dans la littérature, cet opérateur est parfois noté ∇ .

Propriétés d'opérateur:

- Elimination de la tendance :

L'opérateur Δ élimine les tendances linéaires. Pour un processus de la forme

$$X_t = at + b + \varepsilon_t$$

Ou ε_t est stationnaire, on a

$$\Delta(X_t) = X_t - X_{t-1} = (at + b + \varepsilon_t) - (a(t-1) + b + \varepsilon_{t-1}) = a + (\varepsilon_t - \varepsilon_{t-1}).$$

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

Par exemple, pour une tendance de degré

$$X_t = at^2 + bt + c + \varepsilon_t$$

On a

$$\Delta^2 = (I-B)^2 = I-2B+B^2$$

Et donc

$$\Delta^2(X_t) = X_t - 2X_{t-1} + X_{t-2} = \dots = 2a + (\varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2}).$$

Il faut bien noter que l'opérateur Δ^d raccourcit la série des x_t de d valeurs puisque $\Delta^d X_t$ n'est défini que pour $t > d$.

- Elimination de la saisonnalité :

L'opérateur $\Delta_\omega = (I-B^\omega)$ élimine une saisonnalité de période ω (ω entier).

Si on a un processus de la forme

$$X_{k\omega+j} = s_j + \varepsilon_{k\omega+j}$$

Ou ε_t est stationnaire, on a

$$\begin{aligned} \Delta(X_{k\omega+j}) &= (I - B^\omega)(X_{k\omega+j}) \\ &= X_{k\omega+j} - X_{(k-1)\omega+j} \\ &= (s_j + \varepsilon_{k\omega+j}) - (s_j + \varepsilon_{(k-1)\omega+j}) \\ &= \varepsilon_{k\omega+j} - \varepsilon_{(k-1)\omega+j} \end{aligned}$$

On utilise fréquemment les opérateurs Δ_{12} pour les données mensuelles, Δ_7 pour les données journalières, Δ_{24} pour les données horaires,etc.

Cet opérateur est également coûteux en information puisqu'il raccourcit la série de ω valeurs, les ω premières servant de références pour la saisonnalité.

II.7 Modélisations des séries chronologiques

II.7.1 Processus stochastique

Définition:

On appellera ici processus stochastique (à temps discret) une suite $\{X_t\}$ de variables aléatoires indicée par le temps. [06]

Il faut bien noter que

- les X_t ne sont pas forcément indépendants les uns des autres,
- la loi de X_t dépend de t :

$$X_t \sim F_t$$

La notion de processus généralise celle de variable aléatoire : la réalisation d'un processus est une série chronologique (on emploie également le terme de trajectoire).

II.7.2 Processus stationnaire

Un processus est dit stationnaire si la loi de tout k-uplets est invariante par translation dans le temps [06]:

$$\forall k, \forall (t_1, t_2, \dots, t_k), \quad L(X_{t_1}, X_{t_2}, \dots, X_{t_k}) = L(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_k+h})$$

Cette hypothèse de stabilité temporelle de la loi du processus facilite évidemment bien des analyses. Pratiquement, cette hypothèse implique notamment que la covariance entre les valeurs prises en 2 dates distinctes par le processus ne dépend que de la longueur de la période qui les sépare (et pas de la date initiale) :

$$E(X_t), V(X_t) \text{ et } \text{Cov}(X_t, X_{t-h}) \text{ ne dépendent pas de } t.$$

soit

$$E(X_t) = \mu$$

$$V(X_t) = \sigma^2$$

$$\text{Cov}(X_t, X_{t-h}) = \gamma_h$$

Processus faiblement stationnaire : Un processus est dit faiblement stationnaire si ses moments d'ordre 1 et 2 sont stationnaires :

$$E(X_t) \equiv \mu, \quad V(X_t) \equiv \sigma^2, \quad \text{Cov}(X_t, X_{t-h}) = \gamma_h.$$

Remarque:

- La stationnarité entraîne la faible stationnarité.
- Un processus gaussien faiblement stationnaire est stationnaire.

II.7.3 Autocorrélations simple et partielle

Les principes caractéristiques temporels d'un processus sont donnés par autocorrélation (simple) et l'autocorrélation partielle. [03]

II.7.3.1 La fonction d'autocorrélation et d'autocovariance

La fonction d'autocorrélation $\{\gamma(h)\}$ mesure la covariance entre une variable et cette même variable à dates différentes, pour un délai h :

$$\gamma(h) = \text{cov}(Y_t, Y_{t-h}) = E(Y_t - E(Y_t))(Y_{t-h} - E(Y_{t-h}))$$

$$\text{Ainsi } \gamma(0) = \text{var}(Y_t) = E[(Y_t - E(Y_t))^2] = \sigma_Y^2$$

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

La fonction d'autocovariance d'un processus stationnaire est une fonction:

- paire: $\gamma(-h) = \gamma(h) \forall h$

- semi-définie positive:

$$\sum_{j=1}^n \sum_{k=1}^n a_j a_k \gamma(t_j - t_k) > 0 \quad \forall n \in \mathbb{N}, \forall a_j \in \mathbb{R}, \forall t_j \in \mathbb{Z}$$

Puisque cette quantité est égale à $V(\sum_{j=1}^n a_j Y_{t_j})$

II.7.3.2 La fonction d'autocorrélation

Est définie par :

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}, h \in \mathbb{Z}$$

Avec $\rho(0) = 1$ et $|\rho(h)| < 1$ (donc $|\gamma(h)| \leq \gamma(0)$)

On appelle coefficient d'autocorrélation d'ordre 1 (resp. d'ordre k) le coefficient de corrélation linéaire $\rho(1)$ (resp. $\rho(k)$) calculé entre la série et cette série décalée d'une période (resp. k périodes).

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

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

Puisque la fonction $\rho(h), h \in \mathbb{Z}$ est de type positif, on a la propriété suivante :

$$\det R(m) \geq 0, \forall m \in \mathbb{N}^*$$

Ainsi, on a les contraintes suivantes :

$$\begin{aligned} \det R(1) &\geq 0 \\ \det R(2) &\geq 0 \Leftrightarrow \rho(1)^2 \leq 1 \\ \det R(3) &\geq 0 \Leftrightarrow [1 - \rho(2)][1 + \rho(2) - 2\rho(1)^2] \geq 0 \end{aligned}$$

Ainsi, comme $\rho(2) < 1$, on a $\rho(2) \geq 2\rho(1)^2 - 1$. Si la corrélation d'ordre 1 est élevée, il en est de même de la corrélation d'ordre 2 . Il ne peut donc y avoir de chute brutale de valeur entre $\rho(1)$ et $\rho(2)$ lorsque $\rho(1)$ est grand.

L'équivalent empirique de la fonction d'autocorrélation, noté $\hat{\rho}(h)$, est obtenu à partir de l'estimation suivant pour l'autocovariance $\hat{\gamma}(h)$ à l'ordre de h:

$$\hat{\gamma}(h) = \frac{1}{T-h-1} \sum_{j=h+1}^T (y_t - \bar{y})(y_{t-h} - \bar{y})$$

Ou

$$\bar{y} = \frac{1}{T} \sum_t y_t$$

Afin de tester la nullité du coefficient d'autocorrélation d'ordre h, on calcule la variance de ce coefficient. On peut montrer qu'elle est donnée par:

$$V(\hat{\rho}(h)) = \frac{1}{T} \sum_{j=-(h-1)}^{h-1} \hat{\rho}(j)^2$$

Soit en utilisant la symétrie des $\rho(j)$, on obtient:

$$V(\hat{\rho}(h)) = \frac{1}{T} \left(1 + 2 \sum_{j=-(h-1)}^{h-1} \hat{\rho}(j)^2 \right)$$

La statistique de teste de nullité du coefficient d'autocorrélation est :

$$t = \frac{\hat{\rho}(h)}{\sqrt{\hat{V}(\hat{\rho}(h))}}$$

Elle suit une loi de Student. La variance dépendant de h, l'intervalle de coefficient associé au corrélogramme (ensemble des coefficients d'autocorrélation quand h varie) augmente avec h.

II.7.3.3 La fonction d'autocorrélation partielle

Elle mesure la liaison (linéaire) entre Y_t et Y_{t-h} une fois retirés les transitants par les variables intermédiaires $Y_{t-1}, \dots, Y_{t-h+1}$.

Le coefficient d'autocorrélation partielle d'ordre h, noté $r(h)$, est le coefficient de corrélation entre :

$$Y_t - E(Y_t/Y_{t-1}, \dots, Y_{t-h+1})$$

Et

$$Y_{t-h} - E(Y_{t-h}/Y_{t-1}, \dots, Y_{t-h+1})$$

On a donc :

$$r(h) = \text{corr}(Y_t, Y_{t-h}/Y_{t-1}, \dots, Y_{t-h+1})$$

C'est donc le coefficient de Y_{t-h} dans la régression de Y_t sur $Y_{t-1}, \dots, Y_{t-h+1}, Y_{t-h}$

Si Y_t est un processus stationnaire centré, la prédiction optimale de Y_t sachant son passé jusqu'à $t-h$ est donné par:

$$E\left(\frac{Y_t}{Y_{t-1}}, \dots, Y_{t-h}\right) = a_1 Y_{t-1} + \dots + a_h Y_{t-h}$$

Que l'on peut réécrire matriciellement :

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_h \end{pmatrix} = \begin{pmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_{h-1} \\ \gamma_1 & \gamma_0 & \dots & \gamma_{h-2} \\ & & \ddots & \\ & & & \gamma_{h-1} & \gamma_{h-2} & \dots & \gamma_0 \end{pmatrix}^{-1} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_h \end{pmatrix} = \begin{pmatrix} 1 & \rho_1 & \dots & \rho_{h-1} \\ \rho_1 & 1 & \dots & \rho_{h-2} \\ & & \ddots & \\ & & & \rho_{h-1} & \rho_{h-2} & \dots & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_h \end{pmatrix}$$

Le coefficient d'autocorrélation partielle d'ordre h d'un processus stationnaire est alors a_h et calcul de la manière suivante :

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

Avec

$$R(h) = \begin{pmatrix} 1 & \rho_1 & \dots & \rho_{h-1} \\ \rho_1 & 1 & \dots & \rho_{h-2} \\ & & \ddots & \\ & & & \rho_{h-2} & \dots & 1 \end{pmatrix}$$

Et $|R(h)^*|$ la matrice $R(h)$ dans laquelle on a remplacé la colonne h par $\begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_h \end{pmatrix}$, soit:

$$R(h)^* = \begin{pmatrix} 1 & \rho_1 & \dots & \rho_1 \\ \rho_1 & 1 & \dots & \rho_2 \\ & & \ddots & \\ & & & \rho_{h-2} & \dots & \rho_h \end{pmatrix}$$

Ainsi,

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

De manière empirique, les autocorrélations soit partielle s'estiment:

A partir de la régression MCO de Y_t sur les h retards et en prenant le dernier coefficient.

En estimant les autocorrélations simples et en calculant $\hat{r}(h)$ à partir de la formule ci-dessus.

A partir de l'algorithme de DURBIN: il permet de calculer récursivement les divers coefficient de régression en évitant l'inversion des matrice de corrélation $R(h)$. Il est basé sur une formule de calcul des coefficients $a_h(H)$ à partir des $a_h H - 1$ et $a_1(1) = \rho(1)$.

Afin de tester la nullité du coefficient d'autocorrélation partielle d'ordreh, on donne la variance de l'autocorrélation partielle estimée:

$$V(\hat{r}(h)) = \frac{1}{T} \quad \forall h$$

Ainsi, l'intervalle de confiance du corrélogramme partiel est le même pour tout h .

II.7.4 Bruit blanc

Un processus $\{X_t\}$ est un bruit blanc s'il constitue un échantillon i.i.d. d'espérance nulle [06] :

$$\begin{aligned} \forall t: X_t \sim F \quad E(X_t) &= 0 \\ \forall t \neq s: (X_t, X_s) &\text{ indépendants} \end{aligned}$$

Un tel processus n'a ni tendance ni mémoire : la connaissance de la valeur du processus à une date donnée n'apporte aucune information pour la prédiction de sa valeur à une date ultérieure:

$$\forall t > s: L(X_t/X_s) = L(X_t)$$

Bruit blanc faible : Un processus est un bruit blanc faible s'il constitue un échantillon non corrélé :

$$\begin{aligned} \forall t : X_t &\sim F, & E(X_t) &= 0 \\ \forall t \neq s : cov(X_t, X_s) &= 0 \end{aligned}$$

- Un bruit blanc est un bruit blanc faible
- Un bruit blanc faible gaussien est un bruit blanc

II.7.5 Non stationnarité

La plupart des séries économiques sont non stationnaires, c'est-à-dire que le processus qui les décrit ne vérifie pas au moins une des conditions de la définition d'un processus stationnaire du second ordre, donnée par [05]:

- $E(Y_t) = m$ indépendant du temps
- $V(Y_t) = \gamma(0) < \infty$, $\gamma(0)$ indépendant du temps
- $Cov(Y_t, Y_{t-h}) = \gamma(h)$ ne dépend pas de t

Ceci nous conduit à définir deux types de non stationnarité selon que c'est plutôt la condition portant sur le moment d'ordre 1 qui n'est pas vérifiée (non stationnarité déterministe) ou les conditions portant sur les moments du second ordre qui ne sont pas vérifiées (non stationnarité stochastique).

Non stationnarité déterministe:

Le processus TS (Trend Stationary) s'écrit :

$$y_t = \alpha + \beta t + \varepsilon_t$$

Où ε_t représente l'erreur du modèle à la date t .

Il présente un non stationnarité de nature déterministe.

Le processus TS est non stationnaire car $E(y_t) = \alpha + \beta t$ dépend du temps t .

Le processus y_t peut être stationnarisé en retranchant à y_t la valeur estimée $\hat{\alpha} + \hat{\beta}t$ par la méthode des Moindres Carrés Ordinaires (voir document économétrie [1] pour MCO).

Non stationnarité stochastique:

Le processus DS (Differency Stationary) avec dérive ($\beta \neq 0$) s'exprime comme suit :

$$y_t = y_{t-1} + \beta + \varepsilon_t.$$

Le processus DS avec dérive est appelé aussi marche au hasard (ou marche aléatoire) avec dérive.

Il présente un non stationnarité de nature stochastique.

Par récurrence, on obtient (dans le cas avec dérive) :

$$\begin{aligned} y_1 &= y_0 + \beta + \varepsilon_1 \\ y_2 &= y_1 + \beta + \varepsilon_2 = y_0 + \beta + \varepsilon_1 + \beta + \varepsilon_2 \\ &= y_0 + 2\beta + \varepsilon_1 + \varepsilon_2 \\ y_t &= y_0 + \beta t \sum_{i=1}^t \varepsilon_i \text{ où } \varepsilon_i \sim \text{iid}(0, \sigma_i^2) \end{aligned}$$

ε_i est identiquement et indépendamment distribuée.

Le processus DS avec dérive est non stationnaire car on a $E(y_t) = y_0 + \beta t$ qui dépend du temps t . Plus $t \rightarrow \infty$ et plus $E(y_t) \rightarrow \infty$.

Le processus DS (Differency Stationary) sans dérive ($\beta = 0$) s'écrit :

$$y_t = y_{t-1} + \varepsilon_t$$

Le processus DS sans dérive est appelé aussi marche au hasard (ou marche aléatoire).

Par récurrence, on obtient (dans le cas sans dérive) :

$$\begin{aligned} y_1 &= y_0 + \varepsilon_1 \\ y_2 &= y_1 + \varepsilon_2 = y_0 + \varepsilon_1 + \varepsilon_2 \dots \\ y_t &= y_0 + \varepsilon_2 \sum_{i=1}^t \varepsilon_i \text{ ou } \varepsilon_i \sim \text{iid}(0, \sigma_i^2) \end{aligned}$$

Le processus DS sans dérive est non stationnaire car on a :

$$\text{var}(y_i) = \text{var}\left(\sum_{i=1}^t \varepsilon_i\right) = \sum_{i=1}^t \text{var}(\varepsilon_i) = \sum_{i=1}^t \sigma_i^2 = t\sigma_i^2$$

On constate que la variance du processus DS sans dérive dépend du temps t . Plus $t \rightarrow \infty$ et plus $\text{Var}(y_t) \rightarrow \infty$.

Pour stationnariser le processus DS (avec ou sans dérive), il suffit de le passer en différence première :

$$y_t - y_{t-1} = \beta + \varepsilon_t \text{ (cas avec dérive)}$$

Ou

$$y_t - y_{t-1} = \varepsilon_t \text{ (cas sans dérive).}$$

II.7.6 Modèle autorégressif

Dans l'étude d'une série chronologique, il est naturel de penser que la valeur de la série à la date t peut dépendre des valeurs prises aux dates précédentes [06] :

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

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

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

Ou $\{\varepsilon_t\}$ est un bruit blanc.

Processus autorégressif AR(1)

Le processus autorégressif d'ordre 1, noté AR(1), est stationnaire et vérifie l'équation

$$X_t = \varphi X_{t-1} + \varepsilon_t = \varphi B(X_t) + \varepsilon_t$$

Où ε_t est un bruit blanc de variance σ^2 .

On a donc, à toute date t:

$$E(X_t) = \mu, \quad V(X_t) = \sigma_X^2.$$

Espérance, variance, stationnarité : On a

$$E(X_t) = \varphi E(X_{t-1}) + E(\varepsilon_t) = \varphi E(X_{t-1}),$$

$$\text{et } V(X) = \varphi^2 V(X_{t-1}) + V(\varepsilon_t) = \varphi^2 V(X_{t-1}) + \sigma^2.$$

La stationnarité implique que

$$\begin{cases} \mu = \varphi \mu \\ \sigma_X^2 = \varphi^2 \sigma_X^2 + \sigma^2 \end{cases} \Rightarrow \begin{cases} \mu = 0 \text{ ou } \varphi = 1 \\ (1 - \varphi^2) \sigma_X^2 = \sigma^2 \end{cases} \Rightarrow \begin{cases} \mu = 0 \\ |\varphi| < 1 \end{cases}$$

Remarque : La contrainte $\mu = 0$ n'est pas très forte puisqu'on passe d'un processus d'espérance μ à un processus d'espérance nulle par simple translation.

Autocorrélation : On a

$$\begin{aligned} X_t &= \varphi X_{t-1} + \varepsilon_t = \varphi(\varphi X_{t-2} + X_{t-1}) + \varepsilon_t = \varphi^2 X_{t-2} + \varphi \varepsilon_{t-1} + \varepsilon_t \\ &= \varphi^h X_{t-h} + \sum_{k=0}^{h-1} \varphi^k \varepsilon_{t-k} \end{aligned}$$

Donc

$$\begin{aligned} \gamma_1 = cov(X_t, X_{t-h}) &= E(X_t, X_{t-h}) = \rho_h = \left(\left(\varphi^h X_{t-h} + \sum_{k=0}^{h-1} \varphi^k \varepsilon_{t-k} \right) X_{t-h} \right) \\ &= \varphi^h E(X_{t-k}^2) = \varphi^h V(X_{t-h}^2) = \varphi^h \sigma_x^2 \end{aligned}$$

Le coefficient d'autocorrélation vaut donc

$$\rho_h = \frac{\gamma_h}{\gamma_0} = \frac{\varphi^h \sigma_x^2}{\sigma_x^2} = \varphi^h$$

Ce dernier résultat implique que, puisque $|\varphi| < 1$, les corrélogrammes auront des allures décroissantes amorties :

$$\begin{array}{ll} \varphi = 0.9 & \varphi = 0.3 \\ \varphi = -0.9 & \varphi = -0.3 \end{array}$$

Le processus "oublie" progressivement les valeurs passées.

Autocorrélation partielle : En ce qui concerne les autocorrélations, la définition même du processus nous montre que

$$\tau_1 = \rho_1 = 1,$$

$$\tau_h = 0 \text{ pour } h > 1 : \text{ la régression de } X_t \text{ sur } \{X_{t-1}, X_{t-2}, \dots, X_{t-h+1}\} \text{ vaut}$$

$$X_t^* = \phi X_{t-1}$$

Et donc

$$X_t - X_t^* = \varepsilon_t,$$

Or les ε_t sont non corrélés par hypothèse.
Les corrélogrammes partiels auront donc la forme suivante :

$$\varphi = 0.9 \quad \varphi = -0.9$$

Cette forme de corrélogramme partiel rappelle que, conditionnellement à X_{t-1} , la connaissance de X_{t-2} n'apporte aucune information sur X_t .

Exemple : Processus AR (1), $\phi = 0,2$ [02]

Sample: 1 1500
Included observations: 1500

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
■	■	1	0.181	0.181	49.418	0.000
■	■	2	0.022	-0.011	50.138	0.000
■	■	3	-0.005	-0.007	50.170	0.000
■	■	4	-0.009	-0.007	50.289	0.000
■	■	5	-0.036	-0.034	52.207	0.000
■	■	6	-0.017	-0.005	52.644	0.000
■	■	7	0.008	0.013	52.753	0.000
■	■	8	0.026	0.023	53.805	0.000
■	■	9	0.025	0.016	54.778	0.000
■	■	10	0.024	0.016	55.645	0.000
■	■	11	0.009	0.001	55.759	0.000
■	■	12	-0.061	-0.064	61.313	0.000
■	■	13	-0.085	-0.063	72.326	0.000
■	■	14	-0.043	-0.015	75.180	0.000
■	■	15	0.002	0.015	75.189	0.000
■	■	16	0.028	0.026	76.418	0.000
■	■	17	0.018	0.003	76.887	0.000
■	■	18	-0.011	-0.024	77.084	0.000
■	■	19	0.010	0.008	77.244	0.000
■	■	20	-0.019	-0.014	77.821	0.000
■	■	21	0.003	0.017	77.837	0.000
■	■	22	-0.009	-0.005	77.950	0.000

Processus autoregressif AR(2)

Un processus AR(2) est un processus stationnaire qui vérifie une équation de la forme

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \varepsilon_t = \varphi_1 B(X_t) + \varphi_2 B^2(X_t) + \varepsilon_t.$$

Dans un tel modèle, l'influence du passé se manifeste par une régression linéaire sur les deux valeurs antérieures.

Selon les valeurs de φ_1 et φ_2 , il n'est pas toujours possible de trouver un processus stationnaire vérifiant cette équation.

Dans le cas des processus stationnaires, on peut montrer que

$$\rho_h \rightarrow 0.$$

$$h \rightarrow \infty$$

Plus précisément, l'autocorrélation décroît à vitesse exponentielle ; on a par exemple

$$\rho_1 = \frac{\varphi_1}{1 + \varphi_1}, \quad \rho_2 = \frac{\varphi_1^2 - \varphi_2(1 + \varphi_2)}{1 + \varphi_2}$$

Le corrélogramme d'un AR(2) est assez semblable à celui d'un AR(1).

D'autre part, la régression sur les deux dates précédentes donne

$$X_t^* = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} \text{ et donc} \\ X_t - X_t^* = \varepsilon_t.$$

L'autocorrélation partielle τ_h est donc nulle dès que h est supérieur à 2 : $h > 2 \Rightarrow \tau_h = 0$.

Processus autorégressif AR(p)

De façon générale, un processus AR(p) est un processus qui dépend linéairement des p valeurs antérieures :

$$X_t = \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \varepsilon_t$$

Où $\{\varepsilon_t\}$ est un bruit blanc.

On peut, sans restriction de généralité, supprimer le terme constant et obtenir un modèle de la forme

$$X_t = \sum_{k=1}^p \varphi_k X_{t-k} + \varepsilon_t \Leftrightarrow \Phi(B)X_t = \varepsilon_t$$

Où φ est le polynôme de degré p dont les coefficients sont $(1, -\varphi_1, \dots, -\varphi_p)$.

Autocorrélation : On montre que les autocorrélations sont solutions des équations Yule et Walker

$$\rho_h + \sum_{k=1}^p \varphi_k \rho_{h-k} = 0$$

Comme pour les processus AR(1) et AR(2), on montre que ρ_h décroît exponentiellement.

Autocorrélation partielle : Dans un tel processus, X_t et X_{t-p+1} sont indépendants conditionnellement aux valeurs intermédiaires $\{X_{t-1}, X_{t-2}, \dots, X_{t-p}\}$ et donc

$$h > p \Rightarrow \tau_h = 0.$$

La valeur à la date t dépend des p dates précédentes et pas des autres.

Cette propriété sert à l'identification des modèles et à déterminer l'ordre p d'un processus AR(p) au vu du corrélogramme partiel.

Estimation des coefficients : Par ailleurs, les coefficients φ_k du modèle peuvent s'estimer à partir de l'estimation des ρ_h en utilisant le système de **Yule** et **Walker** :

$$\rho_h = \sum_{k=1}^p \varphi_k \rho_{h-k} \text{ pour } h \geq 1$$

Ou par d'autres méthodes comme le maximum de vraisemblance ou les moindres carrés.

II.7.7 Modèle moyenne mobile

Le terme ε_t est souvent présenté comme un "choc" (une innovation pour les économètres) : il rend compte d'un élément nouveau dans l'évolution du processus. On peut envisager que ces chocs (non-corrélés et d'espérances nulles) aient des effets sur l'évolution du processus non seulement à la date à laquelle ils se produisent mais aussi à des dates ultérieures [06].

Processus moyenne mobile MA(1)

Un processus moyenne mobile d'ordre 1, noté MA(1), est un processus stationnaire de la forme:

$$X_t = \varepsilon_t + \theta \varepsilon_{t-1}$$

Où $\{\varepsilon_t\}$ est un bruit blanc de variance σ^2 .

Pour un tel processus, on a

$$\begin{aligned} E(X_t) &= E(\varepsilon_t) + \theta E(\varepsilon_{t-1}) = 0, \\ V(X_t) &= V(\varepsilon_t) + \theta^2 V(\varepsilon_{t-1}) = (1 + \theta^2)\sigma^2. \end{aligned}$$

Autocorrélation : On a

$$\gamma_1 = \text{cov}(X_t, X_{t-1}) = \text{cov}(\varepsilon_t, \theta\varepsilon_{t-1}, \varepsilon_{t-1} + \theta\varepsilon_{t-2}) = \theta V(\varepsilon_{t-1}) = \theta\sigma^2$$

Le coefficient d'autocorrélation d'ordre 1 vaut donc

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\theta\sigma^2}{(1 + \theta^2)\sigma^2} = \frac{\theta}{(1 + \theta^2)}$$

Pour $h > 1$, on a

$$\gamma_h = \text{Cov}(X_t, X_{t-h}) = \text{Cov}(\varepsilon_t + \theta \varepsilon_{t-1}, \varepsilon_{t-h} + \theta \varepsilon_{t-h-1}) = 0$$

Donc les coefficients d'autocorrélations d'ordre supérieur à 1 sont nuls :

$$h > 1 \Rightarrow \rho_h = 0.$$

Si on inverse la formule de ρ_1 , on obtient

$$(1 + \theta^2)\rho_1 = \theta \Leftrightarrow \theta^2\rho_1 - \theta + \rho_1 = 0$$

Équation en θ qui n'admet de solution que si

$$\Delta = 1 - 4\rho_1^2 > 0 \Leftrightarrow \rho_1^2 < 1/4 \Leftrightarrow |\rho_1| < 1/2.$$

Pour un processus MA(1), l'autocorrélation d'ordre 1 est inférieure à 1/2 en valeur absolue. Les corrélogrammes d'un processus MA(1) seront donc de la forme

$$\theta = 0.9 \quad \theta = -0.9$$

Autocorrélation partielle :

Le calcul du coefficient d'autocorrélation partielle est plus complexe, il se résout en manipulant l'équation

$$X_t = (I + \theta B)\varepsilon_t \Leftrightarrow \frac{1}{I + \theta B}X_t = \varepsilon_t \Leftrightarrow \left(\sum_{k=0}^{\infty} (-\theta)^k B^k \right) X_t = \varepsilon_t$$

Ce qui permet de donner une autre équation d'un processus MA(1) sous la forme

$$X_t = \varepsilon_t - \left(\sum_{k=0}^{\infty} (-\theta)^k B^k \right) X_t = \varepsilon_t - \sum_{k=0}^{\infty} (-\theta)^k X_{t-k}$$

On peut ainsi montrer que

$$\tau_h = \frac{(-\theta)^h (\theta^2 - 1)}{1 - \theta^{2(h+1)}}$$

Qui nous donne des corrélogrammes partiels de la forme

$$\begin{array}{ll} \theta = 0.9 & \theta = 0.3 \\ \theta = -0.9 & \theta = -0.3 \end{array}$$

Processus moyenne mobile MA(2)

Un processus MA(2) est défini par une équation du type

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} = (I + \theta_1 B + \theta_2 B^2) \varepsilon_t$$

On a

$$E(X_t) = E(\varepsilon_t) + \theta_1 E(\varepsilon_{t-1}) + \theta_2 E(\varepsilon_{t-2}) = 0$$

$$V(X_t) = V(\varepsilon_t) + \theta_1^2 V(\varepsilon_{t-1}) + \theta_2^2 V(\varepsilon_{t-2}) = (\theta_2^2 + \theta_1^2 + 1) \sigma^2$$

$$\text{Cov}(X_t, X_{t-1}) = \text{Cov}(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}, \varepsilon_{t-1} + \theta_1 \varepsilon_{t-2} + \theta_2 \varepsilon_{t-3})$$

$$= \theta_1 V(\varepsilon_{t-1}) + \theta_2 \theta_1 E(\varepsilon_{t-2}) = \theta_1 (1 + \theta_2) \sigma^2$$

$$\begin{aligned} \text{Cov}(X_t, X_{t-2}) &= \text{Cov}(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}, \varepsilon_{t-2} + \theta_1 \varepsilon_{t-3} + \theta_2 \varepsilon_{t-4}) \\ &= \theta_2 V(\varepsilon_{t-2}) + \theta_2 \sigma^2 \end{aligned}$$

$$\begin{aligned} \text{Cov}(X_t, X_{t-h}) &= \text{Cov}(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}, \varepsilon_{t-h} + \theta_1 \varepsilon_{t-h-1} + \theta_2 \varepsilon_{t-h-2}) \\ &= 0 \text{ pour } h > 2 \end{aligned}$$

On a donc

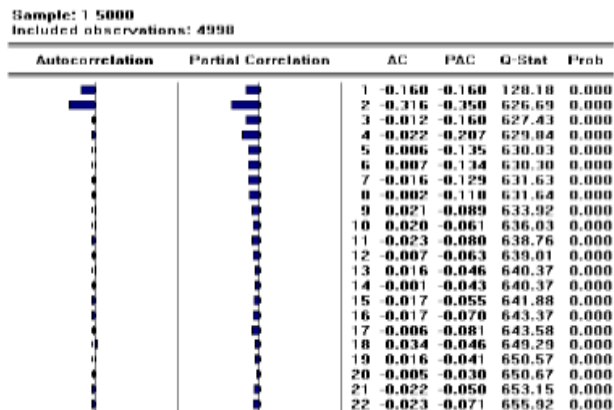
$$\rho_1 = \frac{\theta(1 + \theta_2)}{(\theta_2^2 + \theta_1^2 + 1)}, \quad \rho_2 = \frac{\theta_2}{(\theta_2^2 + \theta_1^2 + 1)}, \quad \rho_h = 0 \text{ pour } h > 2$$

Enfin, on peut montrer que l'autocorrélation partielle décroît de façon exponentielle :

$$\tau_h \rightarrow 0. \\ h \rightarrow \infty$$

Exemple :

Processus MA(2), les deux premières autocorrélations sont négatives (et nulles ensuite) [02]



Processus moyenne mobile MA(q)

Il s'agit d'un processus vérifiant l'équation

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

Où Θ est le polynôme de degré q dont les coefficients sont $\{1, \theta_1, \dots, \theta_q\}$.

Dans un tel modèle, on suppose de plus que l'influence des chocs passés se manifeste au travers d'une fonction linéaire.

Un tel modèle est appelé moyenne mobile d'ordre q car X_t est une moyenne mobile (en anglais moving average) appliquée aux variables aléatoires $\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q}$. Le terme moyen est à prendre dans un sens très large dans la mesure où la somme des coefficients θ_k n'est pas nécessairement égale à 1.

Autocorrélation : Pour un tel processus, on peut montrer que l'autocorrélation ρ_h est nulle pour $h > q$:

$$\rho_h = \begin{cases} \frac{\theta_h + \sum_{k=1}^{k=q-h} \theta_k \theta_{h-k}}{1 + \sum_{k=1}^{k=h} \theta_k^2} & \text{si } h \leq q \\ 0 & \text{si non} \end{cases}$$

Cette propriété est évidemment très précieuse pour l'identification du modèle et la détermination de l'ordre q d'un processus MA(q).

Autocorrélation partielle :

Les autocorrélations partielles τ_h d'un processus moyenne mobile d'ordre q ont un comportement semblable à celui des autocorrélations ρ_h d'un processus autoregressif de même ordre : elles s'amortissent à vitesse exponentielle.

Remarque :

Un processus autorégressif d'ordre 1 peut s'exprimer sous forme de moyenne mobile en inversant l'équation :

$$(I - \varphi B)X_t = \varepsilon_t \Leftrightarrow X_t = \frac{1}{I - \varphi B} \varepsilon_t = \left(\sum_{k=0}^{\infty} \varphi^k B^k \right) \varepsilon_t$$

On obtient ainsi une moyenne mobile d'ordre q infini dont les coefficients décroissent exponentiellement

$$AR(1) \sim MA(\infty).$$

II .7.8 Modèles autorégressif – Moyenne mobile

II.7.8.1 Modèle ARMA (p, q)

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

- une dépendance du processus vis-a-vis de son passé : modèle AR(p),
- un effet retardé des chocs : modèle MA(q).

Un tel modèle, appelé autorégressif - moyenne mobile (ARMA), est caractérisé par le paramètre p de la partie autorégressive et le paramètre q de la partie moyenne mobile. Un processus ARMA(p, q) vérifie l'équation [02] [06]

$$X_t = \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

Soit

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

Le traitement d'un tel processus est plus complexe que celui des 2 précédents. On peut cependant montrer que ses autocorrélations et ses autocorrélations partielles sont des fonctions amorties tendant vers 0 en valeur absolue à vitesses exponentielles.

Propriétés des autocorrélations

1. Soit (X_t) un processus ARMA(p, q), alors les autocovariances $\gamma(h)$ satisfont

$$\gamma(h) - \sum_{i=1}^p \phi_i \gamma(h - i) = 0 \text{ pour } h \geq q + 1 \quad (I)$$

Preuve

La forme ARMA(p,q) de (X_t) est

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

En multipliant par X_{t-h} , où $h \geq q + 1$, et en prenant l'espérance, on obtient (I).

2. Soit (X_t) un processus ARMA(p,q), alors les autocorrélations $\gamma(h)$ satisfont

$$\gamma(h) - \sum_{i=1}^p \phi_i \gamma(h-i) = \sigma^2 [\theta_h + h_1 \theta_{h+1} + \dots + h_{q-h} \theta_q] \text{ pour } 0 \leq h \leq q$$

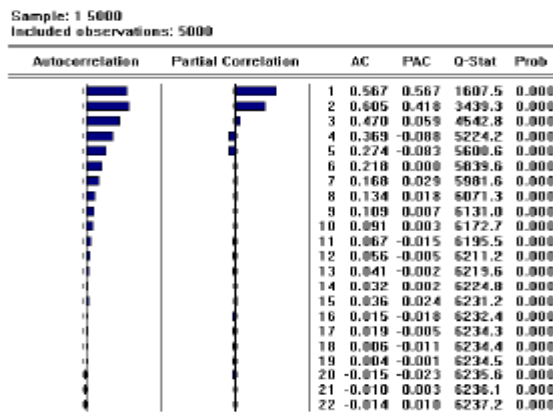
où les h_i correspondent au coefficients de la forme MA(∞) de (X_t) ,

$$X_t = \sum_{j=0}^{+\infty} h_j \varepsilon_{t-1}$$

On peut ainsi dresser un tableau comparatif des corrélogrammes et corrélogrammes partiel des processus

Processus	Autocorrélation ρ_h	Autocorrélation partielle τ_h
AR(p)	Amortie	nulle pour $h > p$
MA(q)	nulle pour $h > q$	Amortie
ARMA(p, q)	Amortie	Amortie

Exemple: des autocorrélogrammes de processus ARMA(2,1) [02]



II.7.8.2 Modèles ARIMA et SARIMA

La condition de stationnarité des modèles présentés ici n'est évidemment pas toujours convenable. On peut intégrer certains types de non-stationnarités en élargissant le modèle ARMA. Ces élargissements consistent en une série d'opérations préalables visant à éliminer la tendance ou la saisonnalité. [02] [06]

Modèle ARIMA :

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

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

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

$$Y_t = \text{ARMA}(p, q) \Leftrightarrow X_t = \text{ARIMA}(p, d, q).$$

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

$$\Phi(B)Y_t = \Theta(B) \varepsilon_t \Leftrightarrow \Phi(B)\Delta^d X_t = \Theta(B) \varepsilon_t$$

Où Φ et Θ sont deux polynômes de degrés respectifs p et q .

Le 'I' de ARIMA signifie l'intégrate "d" comme réciproque de la différenciation.

Evidemment, le degré "d" n'est généralement pas connu. Pour le déterminer on peut agir par tâtonnements ou avoir recours à des tests de stationnarité : puisqu'un processus ARMA(p, q) est stationnaire, on cherche d tel qu'on puisse accepter l'hypothèse de stationnarité pour le processus $Y_t = \Delta^d X_t$. De façon générale, on se réfère à un principe de parcimonie et cherche la valeur satisfaisante minimale de d .

Modèle SARIMA :

Les modèles SARIMA peuvent vus comme une généralisation des modèles ARIMA, contenant une partie saisonnière.

Définition :

de façon générale, soient s_1, \dots, s_n n entiers, alors un processus (X_t) est un processus SARIMA(p, d, q)

- autorégressif moyenne mobile intégré saisonnier - s'il vérifie une équation du type

$$\Phi(L)(1 - L^{s_1}) \dots (1 - L^{s_n})X_t = \Theta(L)\varepsilon_t \text{ pour tout } t \geq 0$$

Où

$$\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 + \dots - \phi_p L^p \text{ ou } \phi_p \neq 0 \text{ et } \Theta(L) = 1 + \theta_1 L - \theta_2 L^2 + \dots - \theta_q L^q \text{ ou } \theta_q \neq 0$$

Sont des polynômes dont les racines sont de module supérieur à 1 et ou les conditions initiales

$$Z_{-1} = \{X_{-1}, \dots, X_{-p}, \varepsilon_{-1}, \dots, \varepsilon_{-q}\}$$

Sont non-corrélées avec $\varepsilon_0, \dots, \varepsilon_t \dots$ et ou le processus (ε_t) est un bruit blanc de variance σ^2

Cette forme incluse les modèles ARIMA puisque 'il suffit de prendre $n=d$ et $s_1 = \dots = s_n = 1$.

Toutefois, les deux formes les plus utilisées sont les suivantes,

$$\begin{aligned} \Phi(L)(1 - L^s)X_t &= \Theta(L)\varepsilon_t \quad \text{pour tout } t \geq 0 \\ \Phi(L)(1 - L^s) \dots (1 - L)^d X_t &= \Theta(L)\varepsilon_t \quad \text{pour tout } t \geq 0 \end{aligned}$$

Où un seul facteur saisonnier "s" intervient, soit applique à un processus ARMA dans le premier cas, soit appliqué à un processus ARIMA dans le second cas.

Exemple :

Soit $s \in N \setminus \{0\}$ correspondant à la saisonnalité, et considérons le processus défini par

$$(1 - \phi L^s)X_t = (1 + \alpha L)(1 + \beta L^s)\varepsilon_t \text{ ou } X_t - \phi X_{t-s} = \varepsilon_t - \alpha\varepsilon_{t-1} - \beta\varepsilon_{t-s} + \alpha\beta\varepsilon_{t-s-1}$$

Les autocorrélations sont données par

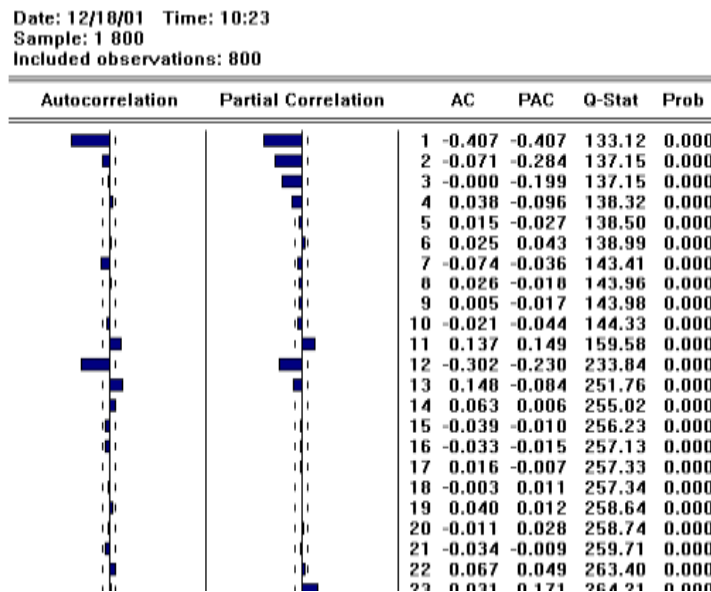
$$\begin{aligned} \rho(1) &= \frac{-\alpha(1+\beta^2)}{(1+\alpha^2)(1+\beta^2)} = \frac{-\alpha}{1+\alpha^2} \\ \rho(S-1) &= \frac{\alpha[\beta-\phi-\phi(\beta-\phi)^2/(1-\phi^2)]}{(1+\alpha^2)[1+(\beta-\phi)^2/(1-\phi^2)]} \\ \rho(S) &= \frac{-(1+\alpha^2)}{\alpha} \rho_{S-1} \end{aligned}$$

Avec

$\rho(h) = 0$ pour $2 \leq h \leq S-2$ puis $\rho(S+1) = \rho(h-1)$ et $\rho(h) = \phi\rho(h-S)$ pour $h \geq S+2$ en particulier

$$\rho(kS) = \phi^{k-1}\rho(S)$$

Le graphique suivant montre l'autocorrélogramme (empirique) d'un tel processus simulé [02]



II.8 Méthode de Box & Jenkins

Dans la méthodologie d'analyse des séries chronologiques synthétisée par Box et Jenkins en 1976, on utilise ces trois types de processus AR, MA, ARMA pour construire un modèle restituant le mieux possible le comportement d'une série chronologique selon une procédure en quatre étapes [01] [02] [06] [08]:

- Identification
- Sélection de modèle
- Estimation
- Validation

On peut expliquer les étapes précédentes dans ce schéma:

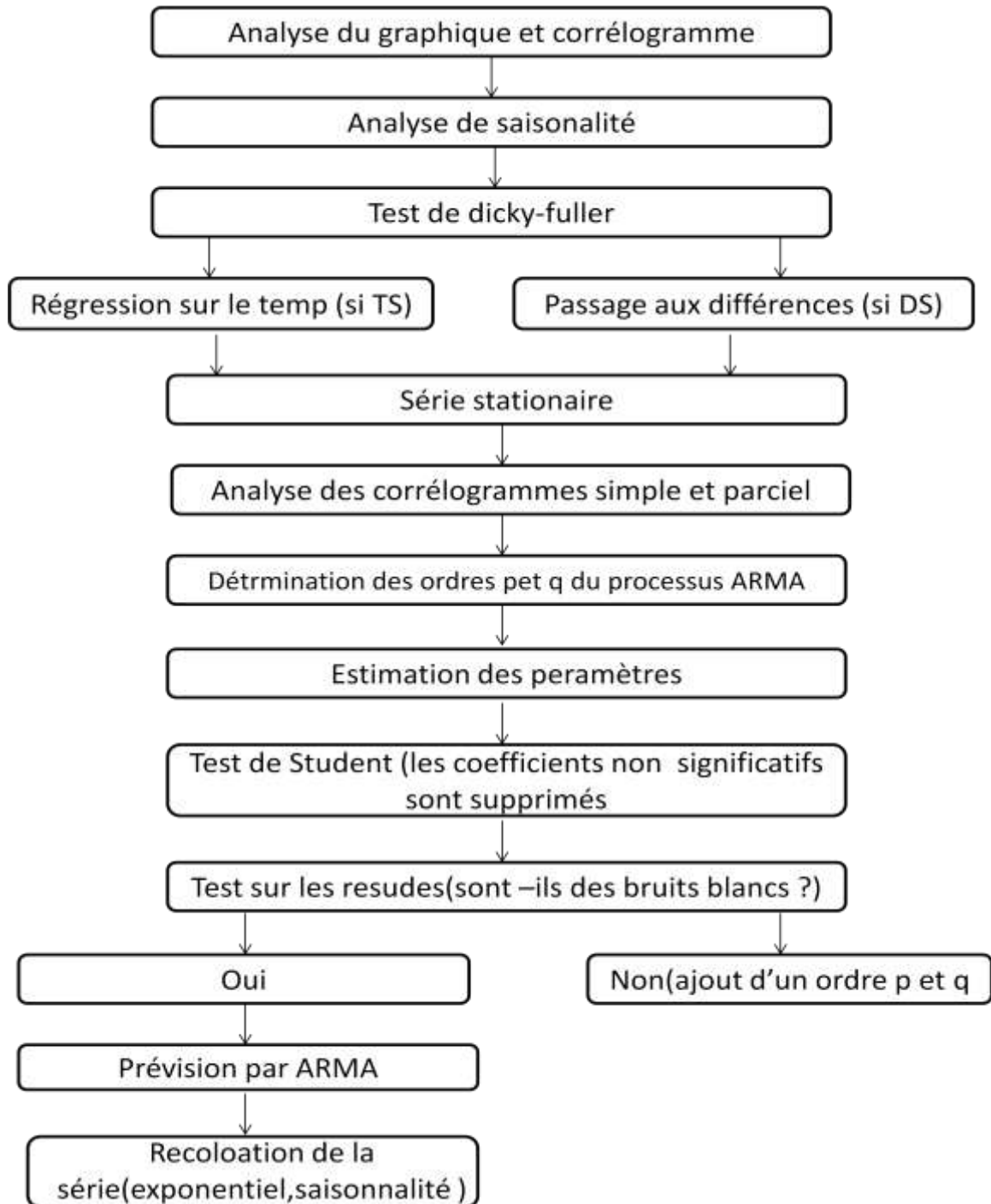


Figure II.1 : la méthode de Bob-Jenkins.

Source : Steven C. WHEELWRIGHT, Spyros MAKRIDAKIS, 1974.

II.8.1 Identification

Préparation des données: Stationnarisation

La première étape consiste à se ramener à une série stationnaire. Elle se décompose en deux phases.

- Utilisation d'une transformation
- Différentiation (simple et/ou saisonnière)

Test de Dicky- Fuller simple

Dicky et Fuller sont les premiers à fournir un ensemble d'outils statistique formels pour détecter la présence d'une racine unitaire dans un processus autorégressif du premier ordre, ce test permet de tester l'hypothèse.

$$\left. \begin{array}{l} H_0: \text{le modèle a une racine unitaire} \\ H_1: \text{le modèle n'apas de racine unitaire} \end{array} \right\}$$

Ce test est regroupé en 4 cas :

$$y_t = \rho y_{t-1} + \sum_{i=1}^p \alpha_i y_{t-i} + \varepsilon_t \quad H_0: \rho = 1$$

$$y_t = \alpha + \rho y_{t-1} + \sum_{i=1}^p \alpha_i y_{t-i} + \varepsilon_t \quad H_0: \alpha = 0 \text{ et } \rho = 1$$

$$y_t = \alpha + \rho y_{t-1} + \sum_{i=1}^p \alpha_i y_{t-i} + \varepsilon_t \quad H_0: \alpha \neq 0 \text{ et } \rho = 1$$

$$y_t = \alpha + \beta t + \rho y_{t-1} + \sum_{i=1}^p \alpha_i y_{t-i} + \varepsilon_t \quad H_0: \alpha = 0, \beta = 0 \text{ et } \rho = 1$$

Pour simplifier, on écrira:

$$\Delta y_t = \phi y_{t-1} + \sum_{i=1}^p \alpha_i \Delta y_{t-i} + \varepsilon_t \text{ avec } \phi = \rho - 1 \text{ "modèle[1]"}'$$

$$\Delta y_t = \alpha + \phi y_{t-1} + \sum_{i=1}^p \alpha_i \Delta y_{t-i} + \varepsilon_t \text{ avec } \phi = \rho - 1 \text{ "modèle[2]"}'$$

$$\Delta y_t = \alpha + \beta t + \phi y_{t-1} + \sum_{i=1}^p \alpha_i \Delta y_{t-i} + \varepsilon_t \text{ avec } \phi = \rho - 1 \text{ "modèle[3]"}'$$

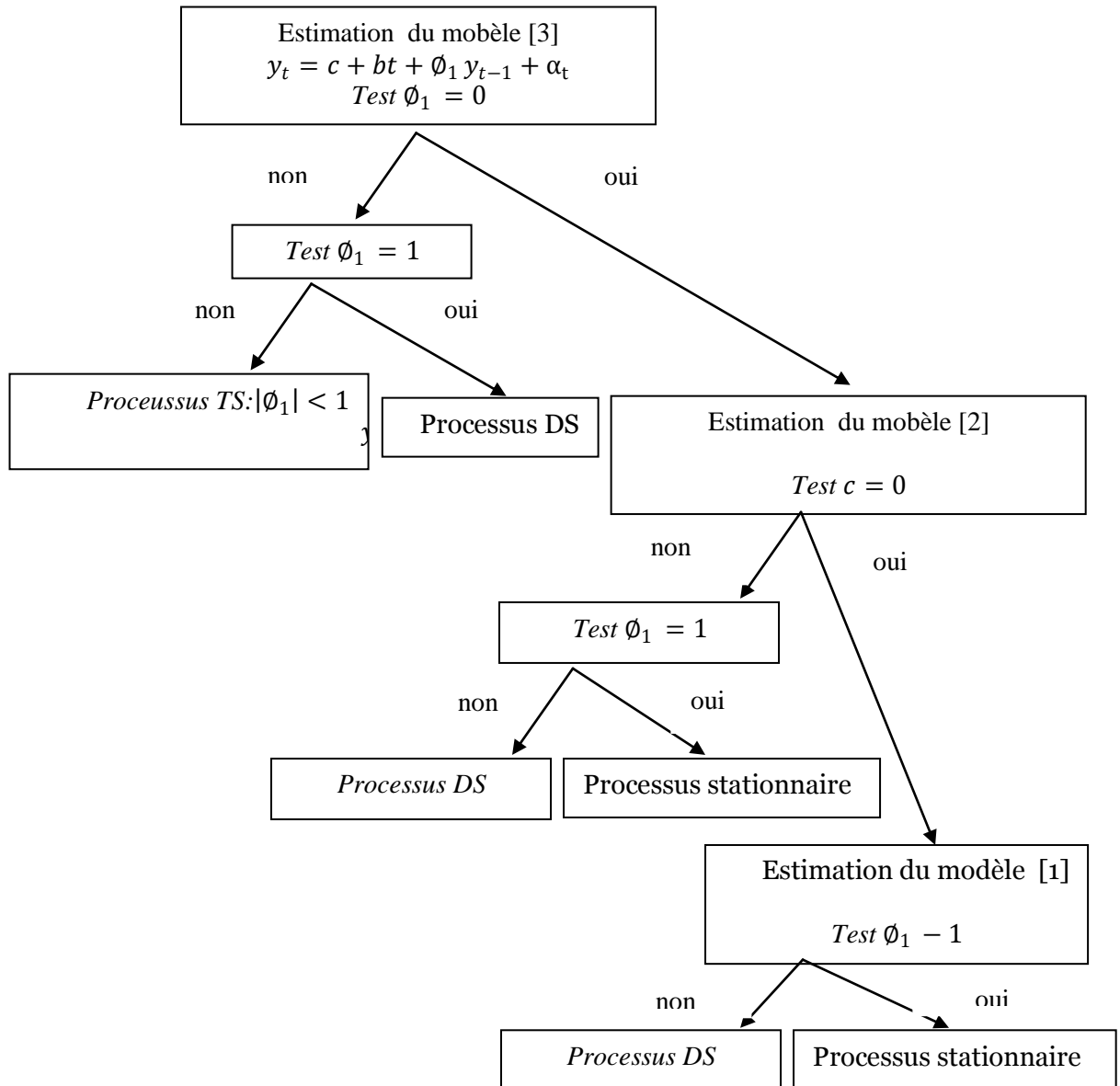


FIG II.2 stratégie simplifiée des racines unitaires

II.8.2 Sélection d'un modèle

Il s'agit de choisir pour les données transformées y_1, \dots, y_n un modèle de type ARMA (p,q) ou de type saisonnier SARIMA (p,0,q)(P,0,Q). L'examen des autocorrélations empirique « totales » (ACF) ou partielles (PACF) permet souvent à lui seul de sélectionner un bon candidat.

ACF et PACF

On se rappelle qu'un processus MA(q) est caractérisé par le fait que ses autocorrélations théoriques $\rho(k)$ sont nulles pour $k > q$ par ailleurs les autocorrélations partielles $\pi(k)$ d'un processus AR(p) sont nulles pour $k > p$. On peut donc espérer reconnaître au moins les processus autorégressif ou moyenne mobile purs. Cependant, on doit se contenter d'estimations de $\rho(k)$ et $\pi(k)$ (puisque les processus sont inconnus), $\hat{\rho}(k)$, $\hat{\pi}(k)$. La sélection s'effectue en examinant les corrélations significatives.

Les estimateurs usuels de $\rho(k)$ et $\pi(k)$ sont naturels compte-tenu de leur définition. Pour le coefficient d'autocorrélation d'ordre k, on utilise :

$$\hat{\rho}(k) = \frac{\sum_{t=1}^{n-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2}$$

Le coefficient d'autocorrélation partielle empirique $\hat{\pi}(k)$ s'obtient par exemple comme l'estimation $\hat{\varphi}_{k,k}$ du coefficient de y_{t-k} , $\varphi_{k,k}$, dans la régression linéaire de y_t sur y_t, \dots, y_{1-k} :

$$y_t = \varphi_{k,1}y_{t-1} + \varphi_{k,2}y_{t-2} + \dots + \varphi_{k,k}y_{t-k} + e_t$$

Les autocorrélations empiriques « complètes » ou partielles sont données par les logiciels de statistique.

Même lorsqu'un coefficient d'autocorrélation théorique est nul, le coefficient estimé correspondante est non nulle et l'on a recours à des tests statistiques permettant de décider si la valeur obtenue est significative ou non. On ne détaillera pas ici les propriétés statistiques sous-jacentes des estimateurs $\hat{\rho}(k)$ et $\hat{\pi}(k)$. Disons simplement que celles-ci dépendent du processus ARMA mais dans la pratique. On considère qu'une valeur à l'intérieur des bornes $\pm 1.96n^{-1/2}$ n'est pas significative quel que soit le modèle. L'interprétation est la suivante : k étant fixé, sous l'hypothèse de nullité de $\rho(k)$ ($\pi(k)$), la probabilité que $\hat{\rho}(k)$, ($\hat{\pi}(k)$) soit à l'intérieur des bornes est approximativement de 95% .

Critères d'information

Une fois choisis l'ensemble des processus ARMA permettant de modéliser les données, il y a donc $(p_{\max} + 1)(q_{\max} + 1)$ modèles possibles et il ne faut en retenir qu'un seul. Pour cela, on dispose de critères de comparaison de modèles que l'on souhaite être le plus petit possible. Les critères classiques sont les suivants:

- Critère d'Akaike (AIC) :

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

- Critère de Schwarz (SBC ou BIC):

$$SBC(p, q) = \log \hat{\sigma}^2 + \frac{2(p + q) \log T}{T}$$

- Critère de Hannan :

$$\varphi(p, q) = \log \hat{\sigma}^2 + \frac{c(p + q) \log \log T}{T}$$

Avec $c > 2$

Pour ces trois critères, $\hat{\sigma}^2$ est l'estimation de la variance σ^2 du bruit blanc. Ces critères sont basés sur le principe de la pénalisation. En effet, on peut montrer que si on passe d'un modèle ARMA(p, q) à un modèle ARMA(p + 1, q) ou à un modèle ARMA(p, q + 1), alors la variance estimée diminue.

Ces critères corrigent donc ce phénomène et on cherche un compromis entre une faible variance et un faible nombre de paramètres à estimer, respectant ainsi le principe de parcimonie (retenir le modèle le moins complexe pour espérer une qualité d'ajustement acceptable).

II.8.3 Estimation

Estimation des paramètres

Méthodes des moments: On a vu dans le paragraphe sur les méthodes empiriques comment estimer les autocorrélations et a donné dans ce chapitre leurs valeurs théoriques pour les processus classiques. La méthode des moments est une des plus utilisées: pour un modèle donné, on sait que les autocorrélations ρ_h et τ_h dépendent des paramètres $\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_d$ selon des équations connues:

$$(\rho_1, \rho_2, \dots, \tau_1, \tau_2, \dots) = F(\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_d)$$

Or on sait facilement estimer les autocorrélations ρ_h et τ_h , il suffit donc d'inverser les équations pour estimer les paramètres:

$$(\hat{\varphi}_1, \dots, \hat{\varphi}_p, \hat{\theta}_1, \dots, \hat{\theta}_d) = F^{-1}(\hat{\rho}_1, \hat{\rho}_2, \dots, \hat{\tau}_1, \hat{\tau}_2, \dots).$$

Il n'est pas nécessaire d'explicitier ici la forme des résultats.

Maximum de vraisemblance:

L'autre méthode classique d'estimation est le maximum de vraisemblance. Elle nécessite de faire des hypothèses sur la loi du bruit blanc $\{E_t\}$. L'hypothèse la plus fréquemment retenue est qu'il s'agit d'un bruit blanc gaussien de variance σ^2 .

On obtient alors des estimateurs du maximum de vraisemblance des paramètres $\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_d$ et de σ .

Cette méthode peut donner lieu à des calculs très lourds aussi bien mathématiquement que numériquement à cause de la forme très « tourmentée » des fonctions de vraisemblance.

Pour éviter des temps de calculs colossaux ; on a le plus souvent recours à des algorithmes qui utilisent des approximations plus ou moins fortes.

II.8.4 Validation

Tous les résultats d'estimation obtenus, toutes les prévisions que l'on pourra faire, dépendent de façon cruciale du fait que les résidus forment un bruit blanc gaussien.

Il est donc capital d'accorder un soin particulier à la vérification de cette hypothèse.

Teste du bruit blanc: Dans l'analyse des séries chronologiques par processus ; le bruit blanc joue un rôle particulier puisque c'est un processus sans aucune structure. Quand, pour un processus $\{X_t\}$, on a éliminé toute tendance, toute saisonnalité et toute dépendance vis-à-vis du passé ; il reste un processus $\{E_t\}$ complètement imprévisible sur lequel il n'y a plus grand-chose à dire.

Le test préliminaire concerne justement ce processus. Quand on étudie une série x_t la première hypothèse testée est :

$$H_0: "\{X_t\} \text{ est un bruit blanc}"$$

Si on accepte cette hypothèse, l'analyse de la série est virtuellement achevée: la série étudiée n'a aucune structure.

On peut citer:

- le teste "Portmanteau" fondé sur la somme des carrés des autocorrélations :

$$Q_1(\{x_t\}) = T \sum_{h=1}^H \hat{\rho}_h^2(x)$$

Ou T est le nombre d'observations, $\hat{\rho}_h(x)$ l'autocorrélation d'ordre h estimée sur la série $\{x_t, 1 \leq t \leq T\}$ et H une quantité suffisamment grande.

- le test fondé sur la statistique de **Ljung et Box** :

$$Q_2(\{x_t\}) = T(T + 2) \sum_{h=1}^H \frac{1}{T - h} \hat{\rho}_h^2(x)$$

Sous l'hypothèse d'un bruit gaussien, ces deux statistiques suivent des lois du χ^2 à $H - p - q$ degrés de liberté.

Remarque:

Les tests de bruit blanc permettent également de “ valider ” le modèle. Si le processus est bien un ARMA(p, q) , le processus résiduel également $\{\hat{E}_t\} = \{X - \hat{X}_t\}$ doit être un bruit blanc. On accepte (ou rejette) cette hypothèse en effectuant un des tests fondés sur une statistique $Q(\{\hat{e}_t\})$ pour cela, il faut donc estimer les autocorrélations du processus $\{\hat{E}_t\}$.

On ne peut envisager d’effectuer des prédictions qu’une fois qu’on a accepté l’hypothèse $H_0: \{E_t\} = \{X_t - \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \theta_1 E_{t-1} + \dots + \theta_q E_{t-q}\}$ est un bruit blanc“

Tests le Normalité des erreurs:

Hypothèses du test :

H_0 : les erreurs suivent une loi Normale

H_1 : les erreurs ne suivent pas une loi Normale

Sous l'hypothèse H_1 la loi des erreurs est donc inconnue.

On caractérise la loi normale $N(m, \sigma^2)$ par le fait

- qu'elle est symétrique \Rightarrow son moment centré d'ordre 3 est nul $\mu_3 = 0$
- que son moment centré d'ordre 4 est $\mu_4 = 3 \mu_2^2 = 3(\sigma^2)^2 \Rightarrow$ sa Kurtosis $K = \mu_4 / \sigma^4 = 3$

Les hypothèses du test peuvent alors s'écrire

$$H_0 : \mu_3 = 0 \text{ et } \mu_4 = 3\sigma^4$$

H_1 : l'une au moins de ces deux propriétés n'est pas vérifiée

Le moment σ^2 est estimé à l'aide des résidus par $s^2 = \sum e_t^2 / (n - k)$

Le moment $\mu_3 = E[(\epsilon - E(\epsilon))^3] = E[\epsilon^3]$ est estimé à l'aide des résidus par $\hat{\mu}_3 = \sum e_t^3 / n$

Le moment $\mu_4 = E[(\epsilon - E(\epsilon))^4] = E[\epsilon^4]$ est estimé à l'aide des résidus par $\hat{\mu}_4 = \sum e_t^4 / n$

- on définit le coefficient de symétrie (skewness)

$$\alpha_3 = \frac{\mu^3}{\sigma^3} \text{ estimé par } \hat{\alpha}_3 = \frac{\sum e_t^3 / n}{s^3}$$

qui suit asymptotiquement une loi $N(0, 3!/n)$ sous l'hypothèse H_0

- on définit le coefficient d'aplatissement (Kurtosis)

$$\alpha_4 = \frac{\mu^4}{\sigma^4} \text{ estimé par } \hat{\alpha}_4 = \frac{\sum e_t^4 / n}{s^4}$$

qui suit asymptotiquement une loi $N(3, 4!/n)$ sous l'hypothèse H_0

Les hypothèses du test peuvent donc s'écrire

$$H_0 : \alpha_3 = 0 \text{ et } \alpha_4 = 3$$

H_1 : l'une au moins de ces deux propriétés n'est pas vérifiée

Il existe deux tests de normalité l'un testant séparément les deux parties de H_0 , l'autre global

Tests de Skewness et Kurtosis

Test de Skewness

$$H_{01} : \mu_3 = 0 \rightarrow \alpha_3 = 0$$

$$H_{11} : \mu_3 \neq 0 \rightarrow \alpha_3 \neq 0$$

sous l’hypothèse H_0 l’estimateur $\hat{\alpha}_3$ de α_3 suit asymptotiquement une loi $N(0, 3!/n)$ où $(3!=6)$, sa variable centrée et réduite t

$$t = \sqrt{n/6} (\hat{\alpha}_3) \text{ suit asymptotiquement une } N(0,1)$$

Si $-1.96 < t_{\text{estimations}} < 1.96$ on décide H_{01} sinon H_{11}

Si on décide H_{11} le test est terminé car la loi n'étant pas symétrique elle ne peut être normale.

Si la décision est H_{01} on passe au test suivant.

Test de Kurtosis

$$H_{02} : \mu_4 = 3 \sigma^4 \rightarrow \alpha_4 = 3$$

$$H_{12} : \mu_4 \neq 3 \sigma^4 \rightarrow \alpha_4 \neq 3$$

sous l'hypothèse H_0 l'estimateur $\hat{\alpha}_4$ de α_4 suit asymptotiquement une loi $N(3, 4!/n)$ ou $(4!=24)$

$$t = \sqrt{n/24} (\hat{\alpha}_4 - 3) \text{ suit asymptotiquement une } N(0,1)$$

Si $-1.96 < t_{\text{estimation}} < 1.96$ on décide H_{02} sinon H_{12}

si la décision est H_{02} , on a donc les deux propriétés vérifiées, on décide alors normalité des erreurs

Si on décide H_{12} le test est terminé car la loi n'étant pas un coefficient d'aplatissement égal à 3 nous ne sommes pas dans le cadre de la loi normale.

Test Global de Jarque et Berra

Il teste donc globalement

$$H_0 : \mu_3 = 0 \text{ et } \mu_4 = 3\sigma^4$$

$$H_1 : \text{l'une au moins de ces deux propriétés n'est pas vérifiée}$$

Sous l'hypothèse H_0 vraie la variable aléatoire S somme des carrés des deux précédents résultats centrés réduits suit donc la loi du **khideux** à deux degrés de liberté.

$$S = \frac{n}{6} \hat{\alpha}_3^2 + \frac{n}{24} (\hat{\alpha}_4 - 3)^2$$

Si $S_{\text{estimation}}$ est inférieur à la borne du khideux on décide H_0 sinon on décide H_1 .

Prévisions à l'aide des modèles ARIMA: Box-Jenkins

Etant donnée une série stationnaire (X_t) , observée entre 1 et T, on cherche à faire de la prévision à horizon h, et donc à prévoir X_{T+1}, \dots, X_{T+h} . Tous les processus AR, MA et ARMA seront supposés mis sous forme canonique, et n'avoir aucune racine unité. Aussi, toutes les racines des polynômes autorégressifs Φ et des polynômes moyennes-mobiles Θ auront leurs racines à l'extérieur du disque unité. Ainsi, pour tous les processus X_t tels que

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

ε_t sera l'innovation du processus X_t . [02]

Prévisions à l'aide d'un modèle AR(p)

Le modèle s'écrit, quitte à recentrer le processus,

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t \text{ ou } \Phi(L)X_t = \varepsilon_t$$

La prévision optimale à la date T + 1, faite à la date T est $TX_{T+1}^* = EL(\{X_{T+1} | X_T, X_{T-1}, \dots\})$
Aussi,

$$TX_{T+1}^* = \phi_1 X_T + \dots + \phi_p X_{T-p}$$

Car (ε_t) est l'innovation. De façon analogue,

$$X_{T+h} = \phi_1 X_{T+h-1} + \dots + \phi_p X_{T+h-p} + \varepsilon_{T+h} \text{ et donc } TX_{T+1}^* = \text{EL}(\{X_{T+h} | X_T, X_{T-1}, \dots\})$$

Est donné, de façon récursive par

$$TX_{T+1}^* = \begin{cases} \phi_1 \cdot TX_{T+h-1}^* + \dots + \phi_{h-1} \cdot TX_{T+1}^* + \phi_h X_T + \dots + \phi_p X_{T+h-p} & \text{pour } h \leq p \\ \phi_1 \cdot TX_{T+h-1}^* + \dots + \phi_p TX_{T+h-p}^* & \text{pour } h > p \end{cases}$$

Prévisions à l'aide d'un modèle MA(q):

On supposera là aussi que l'on s'est ramené à un processus centré (X_t) , satisfaisant

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

La prévision optimale à la date T+1, faite à la date T est

$${}_T X_{T+1}^* = \text{EL}(X_{T+1} | X_T, X_{T-1}, \dots) = \text{EL}(X_{T+1} | \varepsilon_T, \varepsilon_{T-1}, \dots)$$

Car (ε_t) est le processus d'innovation. Aussi,

$${}_T X_{T+1}^* = 0 + \theta_1 \varepsilon_T + \dots + \theta_q \varepsilon_{T+1-q}$$

De façon analogue, X_{T+h} est estimé par

$${}_T X_{T+1}^* = \text{EL}(X_{T+1} | X_T, X_{T-1}, \dots) = \text{EL}(X_{T+h} | \varepsilon_T, \varepsilon_{T-1}, \dots) \text{ et donc}$$

$${}_T X_{T+1}^* = \begin{cases} \theta_h \cdot {}_T X_T^* + \dots + \theta_q X_{T+h-q}^* & \text{pour } h \leq p \\ 0 & \text{pour } h > p \end{cases}$$

Toutefois, cette méthode présente le désavantage d'estimer X_{T+h} à partir des résidus passés, a priori non observables, et non pas du passé de la variable.

Prévisions à l'aide d'un modèle ARMA(p; q):

On supposera là aussi que l'on s'est ramené à un processus centré (X_t) , satisfaisant

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

Remarque : Dans le cas d'un processus non centré, $\Phi(L)X_t = \mu + \Theta(L)\varepsilon_t$, on peut noter que $\text{EX}_t = \mu/\Phi(1) = m$ et que $\Phi(L)(X_t - m) = \Theta(L)\varepsilon_t$. Il est donc toujours possible, en translatant le processus, de se ramener à un processus centré (pour des processus ARMA seulement).

Sous cette forme ARMA, alors

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

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

On a alors ${}_{T+1}X_{T+h}^* = EL(X_{T+h} | X_T, X_{T-1}, \dots) = EL(X_{T+h} | \varepsilon_T, \varepsilon_{T-1}, \dots)$ car ε_t est le processus d'innovation. On peut noter que pour $h > q$

$${}_{T+1}X_{T-h}^* = \begin{cases} \phi_1 \cdot {}_T X_{T-h}^* + \dots + \phi_{h-1} \cdot {}_T X_{T-h}^* + \phi_1 X_T + \dots + \phi_p X_{T+h-p}, & \text{pour } h \leq p \\ \phi_1 \cdot {}_T X_{T-h-1}^* + \dots + \phi_p \cdot {}_T X_{T+h-p}^*, & \text{pour } h > p \end{cases}$$

La forme générale des solutions est connue (comme dans le cas des AR(p)). Toutefois, il y a ici un problème d'initialisation des calculs.

Prévision dans le cas d'un processus ARIMA(p, d, q):

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

$$\Phi(L)(1-L)^d X_t = \Theta(L)\varepsilon_t \text{ Avec les conditions initial } z = (X_{-1}, \dots, X_{-p-d}, \varepsilon_{-1}, \dots, \varepsilon_{-q})'$$

Posons alors $\psi(L) = \Phi(L)(1-L)^d$. la forme ARIMA(p, d, q) peut s'écrire

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

et donc

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

Notons ${}_T X_{T+h}^*$ la prévision faite à la date T

$${}_T X_{T+h}^* = EL(X_{T+h} | X_T, X_{T-1}, \dots, X_0, Z). \text{ Alors}$$

$${}_T X_{T+h}^* = \sum_{i=1}^{p+d} \psi_i {}_T X_{T+h-i}^* + 0 + \sum_{j=1}^q \theta_j {}_T \varepsilon_{T+h-j}^* \text{ où } \begin{cases} {}_T X_{T+h-i}^* = X_{T+h-i} \text{ pour } i \leq h \\ {}_T \varepsilon_{T+h-j}^* = \begin{cases} 0 & \text{pour } j < h \\ \varepsilon_{T+h-j} & \text{pour } j \geq h \end{cases} \end{cases}$$

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

$${}_T X_{T+h}^* = \sum_{i=1}^{p+d} \psi_i {}_T X_{T+h-i}^*$$

Chapitre III

Application de la méthode de Box et Jenkins Sur la série ELEC

Chapitre III : Application de la méthode de Box et Jenkins

Nous nous proposons dans ce chapitre d'appliquer la méthodologie de Box et Jenkins exposée dans le chapitre précédant à notre série de données réelles.

Nos données présentent la quantité d'électricité transporté par la société Algérienne de gestion de réseaux de transport de l'électricité (Hassi-Messaoud).

Les données sont mensuelles et la période retenue pour notre étude est entre janvier 2008 et décembre 2016, ce qui nous donne un total de 108 observations.

Pour le traitement économétrique de notre série on utilise le logiciel Eviews (8.0)

Notre série de données est présentée dans le tableau suivant:

	Janvier	Février	Mars	Avril	Mai	Juin
2008	177558568	159568430	153090004	154748304	195730989	221693165
2009	196905509	168450119	169112289	166192295	206987505	288215437
2010	205311755	186865580	186865580	190391948	206836701	300811456
2011	235940346	208107256	212817603	207931462	243404715	325968688
2012	297562961	283244996	250715256	251863340	315771476	481796664
2013	313255992	277579256	267680338	263696248	318501284	407591582
2014	321474423	270531505	279193511	272732801	353860012	463886551
2015	378099577	324680526	306444635	322708815	430357886	567309219
2016	373966636	326885733	336500737	336769491	445991153	678493085,5

	Juillet	Aout	Septembre	Octobre	Novembre	Décembre
2008	299824508	281645361	261181730	166216190	168305837	199502446
2009	346804253	349907383	239618474	190240008	188129984	205380157
2010	375510849	419612776	292530837	222214670	206547931	229995088
2011	435299412	494143696	396124631	234899863	225984343	260126458
2012	578643846	559250475	372399412	301349881	257916242	308825933
2013	608965054	531664688	422243785	357941113	294155375	331940329
2014	683425775	637289762	570075491	343495057	284910833	345826612
2015	703934240	696838079	539549477	367249494	326367434	380572851
2016	760283192	682747112	581336527	401570408	327679514	362987524

Tableau III.1: Transport d'électricité par GRTE/ DTE-HM (2008-2016) (Unité : KWh)

III.1. Analyse du graphique et de corrélogramme

III.1.1 Représentation graphique

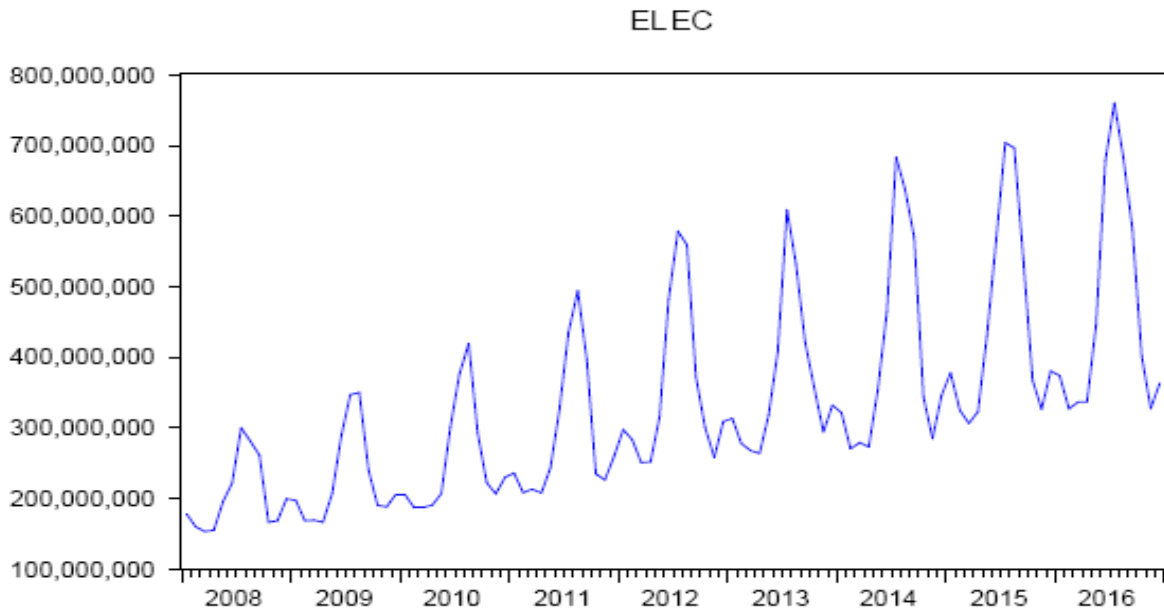


Figure III.1: Représentation graphique de la série (ELEC)

D'après l'analyse graphique de la série (ELEC) on observe 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.

III.1.2 Analyse du corrélogramme de la série ELEC

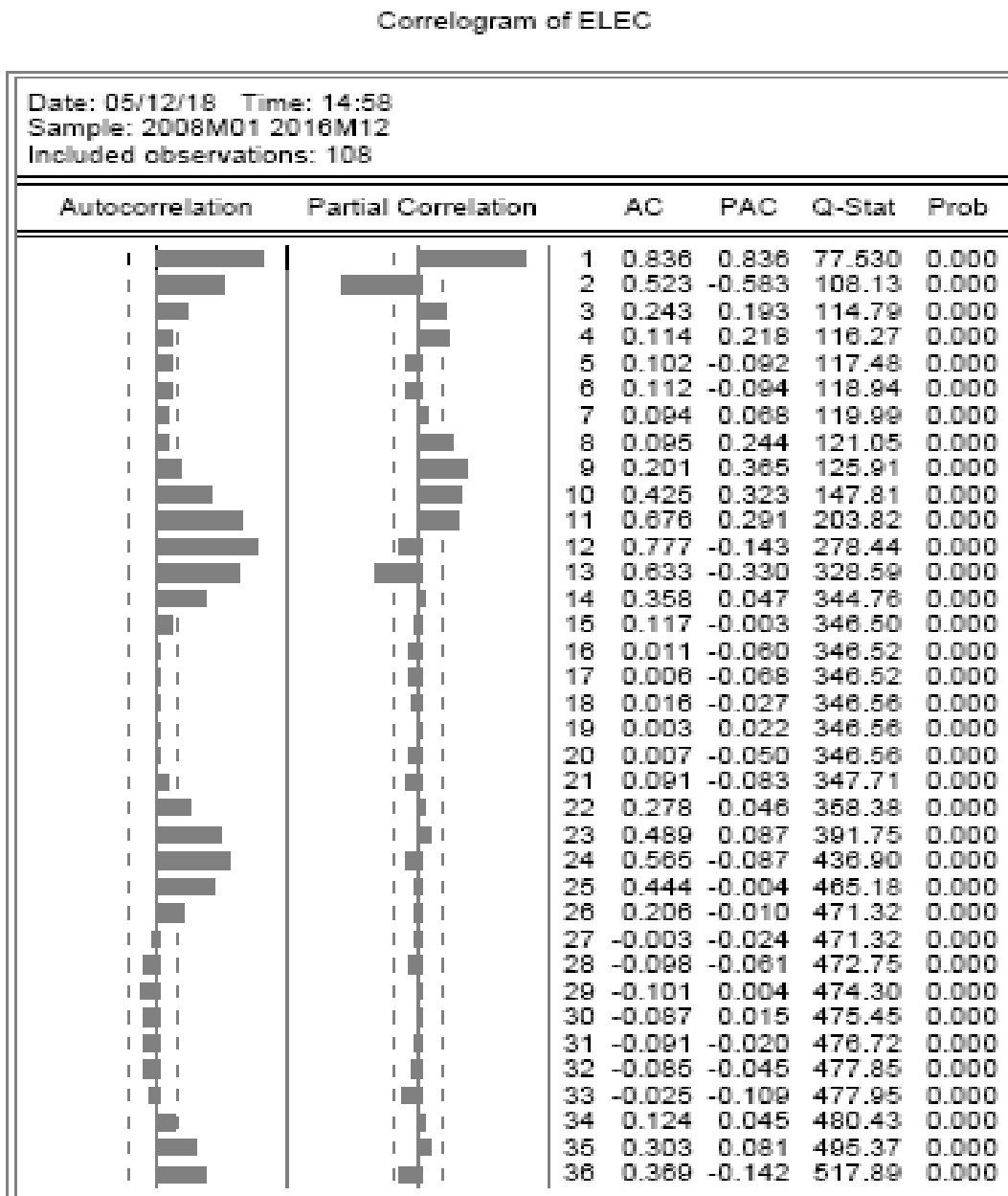


Figure III.2: Corrélogramme de la série (ELEC)

D'après l'analyse statistique du graphe on remarque qu'il y à beaucoup de coefficients de Fonction d'autocorrélation qui sont significatifs comme, c.à.d. qu'il vu la présence de Saisonnalité, ce qui entraine à dire que notre série est non stationnaire.

III.2. Analyse de saisonnalité

III.2.1 Graphique par saison

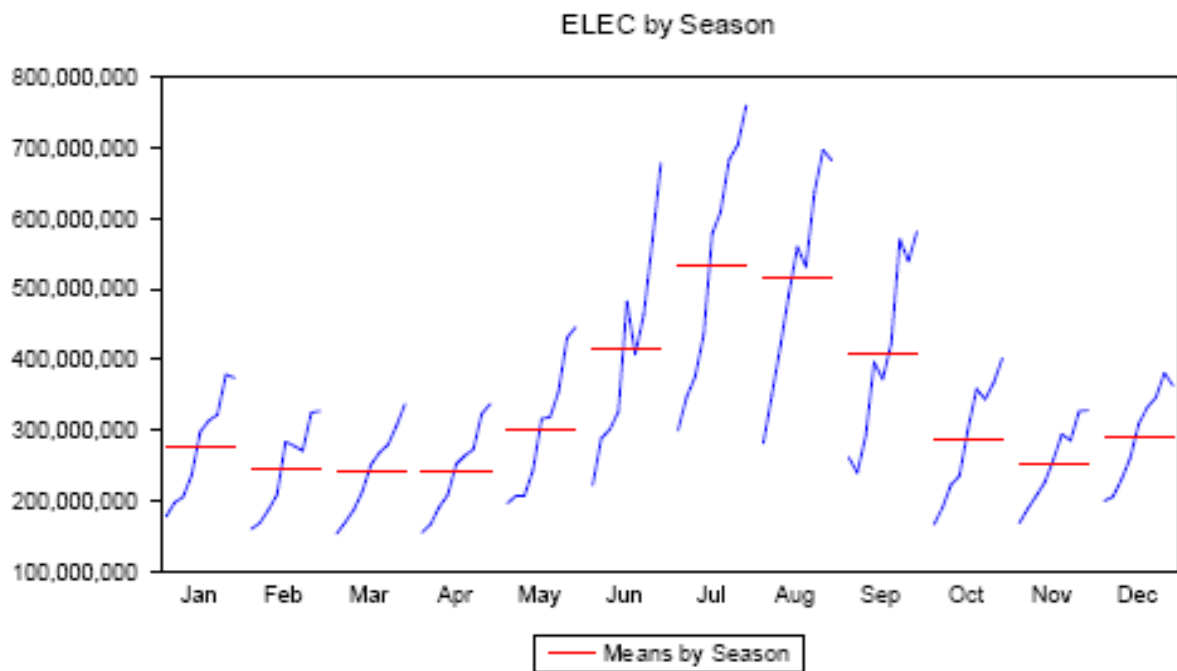


Figure III.3: Graphede la série (ELEC) par saison

Cette représentation graphique par saison montre un mouvement saisonnier répétitif trimestriel.

III.2.2 Désaisonnaliser la série

On procède à la désaisonnalisation de la série par l'application de différence saisonnière, on obtient une nouvelle série.

Puisque notre série est mensuelle la saisonnalité est présumée $S=12$, La différence ce fait sur 12 valeurs.

$$\text{Genre } LELECSA = LELEC - LELEC(-12)$$

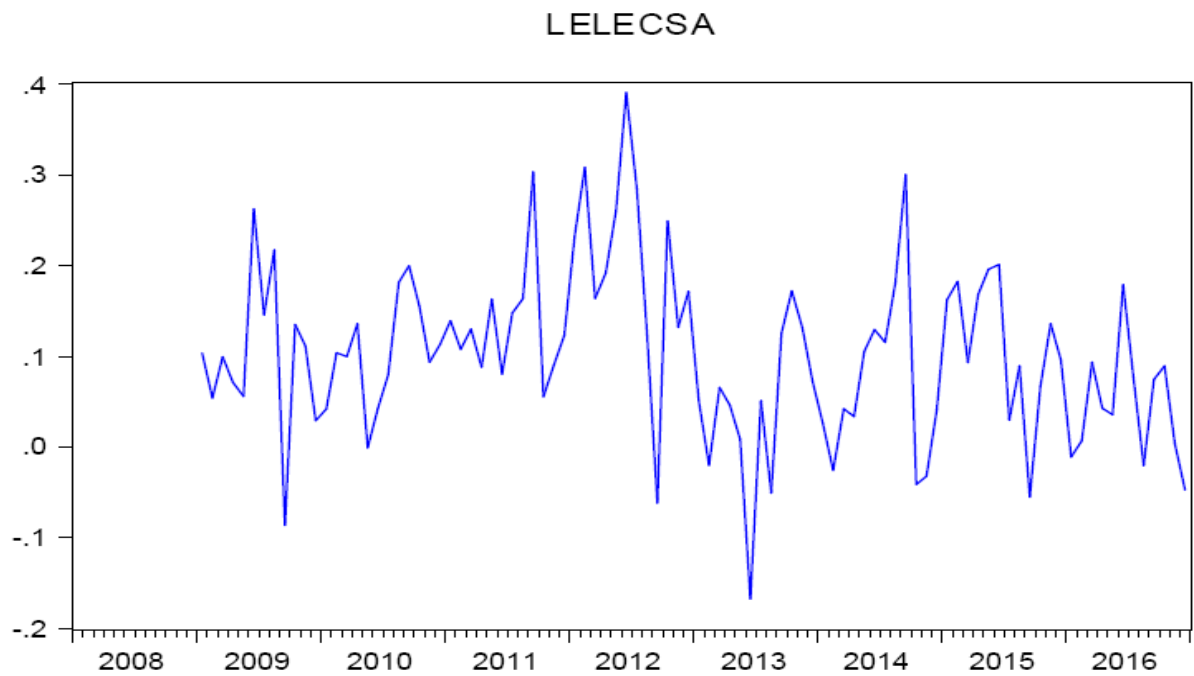


Figure III.4: Graphe de la série brute (LELECSA)

Ce graphique représente la série désaisonnalisée, mais pour s'assurer de la désaisonnalité de notre nouvelle série on trace le corrélogramme.

Correlogram of LELECSA

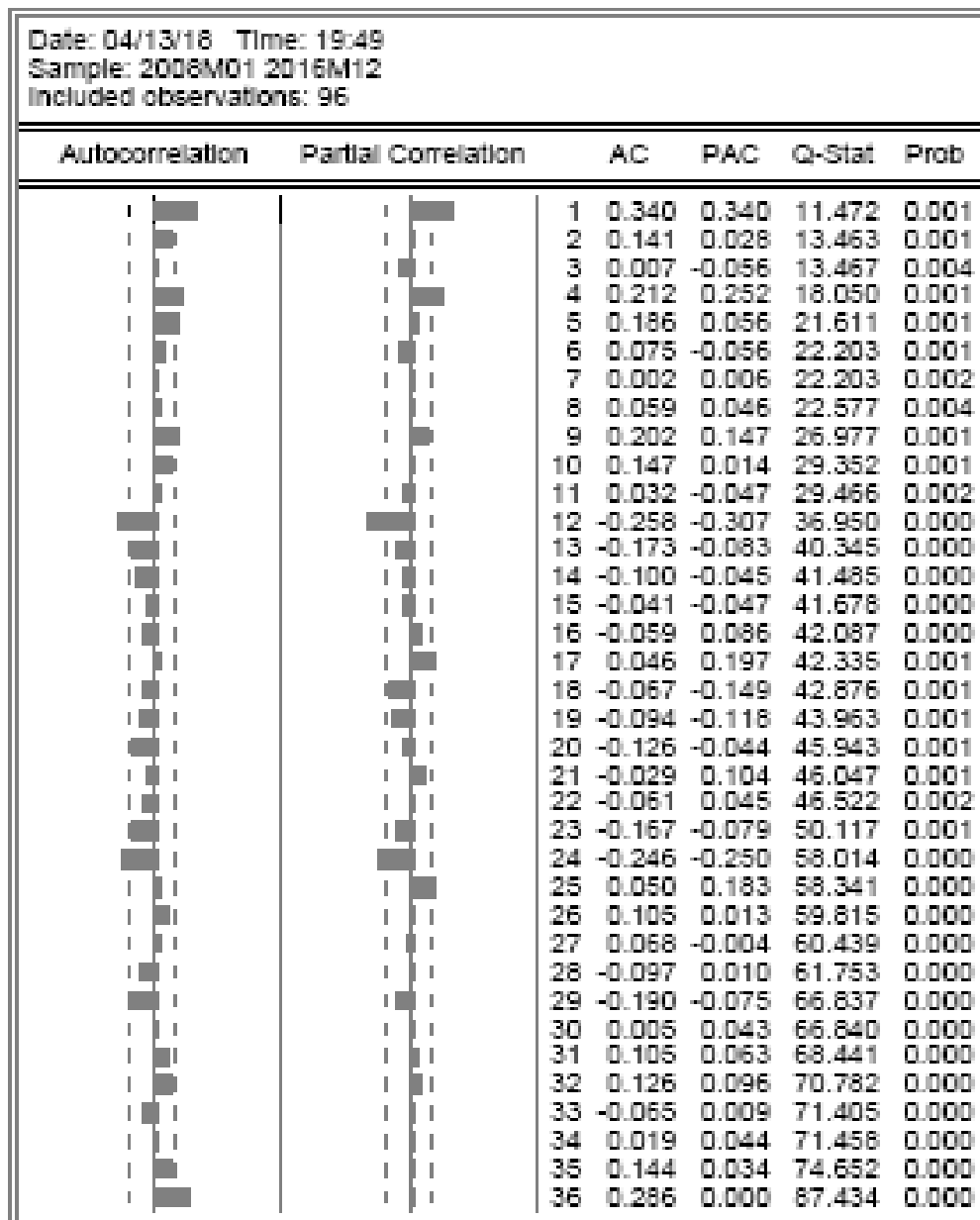


Figure III.5:Corrélogramme de la série (LELECSA)

D'après l'analyse statistique du graphe on remarque presque tous les coefficients de fonction

D'autocorrélation ne sont pas significatifs.

De ce fait, on peut dire que notre nouvelle série est non saisonnière

Il reste à tester la stationnarité.

III.3 stationnarisation de la série

Pour tester la stationnarité de notre série on utilise le test de Dickey –Fuller augment (ADF)

Avec constant et tendance :

Augmented Dickey-Fuller Unit Root Test on LELECSA

Null Hypothesis: LELECSA has a unit root				
Exogenous: Constant, Linear Trend				
Lag Length: 0 (Automatic - based on SIC, maxlag=11)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-6.798048	0.0000
Test critical values:	1% level		-4.057528	
	5% level		-3.457808	
	10% level		-3.154859	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(LELECSA)				
Method: Least Squares				
Date: 04/13/18 Time: 19:52				
Sample (adjusted): 2009M02 2016M12				
Included observations: 95 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LELECSA(-1)	-0.676538	0.099549	-6.798048	0.0000
C	0.098355	0.026048	3.775861	0.0003
@TREND("2008M01")	-0.000505	0.000340	-1.485792	0.1408
R-squared	0.334649	Mean dependent var	-0.001587	
Adjusted R-squared	0.320185	S.D. dependent var	0.108451	
S.E. of regression	0.089419	Akaike info criterion	-1.959908	
Sum squared resid	0.735602	Schwarz criterion	-1.879259	
Log likelihood	96.09564	Hannan-Quinn criter.	-1.927320	
F-statistic	23.13649	Durbin-Watson stat	1.996160	
Prob(F-statistic)	0.000000			

Tableau III.2 : Test de Dickey –Fuller augment (avec constant et tendance) de la série (LELECSA)

- On a t-statistique < t-tabulée au risque de 5% et p-value < 0.05

Donc notre nouvelle série n'admet pas de racine unitaire

-On a p-val > 0.05 alors notre série ne contient pas de tendance

-On a p-val < 0.05 alors elle contient de constant

Sans tendance avec constant

Augmented Dickey-Fuller Unit Root Test on LELECSA

Null Hypothesis: LELECSA has a unit root				
Exogenous: Constant				
Lag Length: 0 (Automatic - based on SIC, maxlag=11)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-6.595495	0.0000
Test critical values:	1% level		-3.500669	
	5% level		-2.892200	
	10% level		-2.583192	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(LELECSA)				
Method: Least Squares				
Date: 04/13/18 Time: 19:53				
Sample (adjusted): 2009M02 2016M12				
Included observations: 95 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LELECSA(-1)	-0.650511	0.098630	-6.595495	0.0000
C	0.065379	0.013724	4.763868	0.0000
R-squared	0.318684	Mean dependent var		-0.001587
Adjusted R-squared	0.311358	S.D. dependent var		0.108451
S.E. of regression	0.089997	Akaike info criterion		-1.957249
Sum squared resid	0.753253	Schwarz criterion		-1.903483
Log likelihood	94.96931	Hannan-Quinn criter.		-1.935523
F-statistic	43.50056	Durbin-Watson stat		2.001760
Prob(F-statistic)	0.000000			

Tableau III.3 : Test de Dickey –Fuller augment (avec constant) de la série (LELECSA)

-On a t-statistique < t-tabulée au risque 5% et p-value < 0.05 donc notre nouvelle série n'admet pas de racine unitaire

D'après le test ADF la série (lelecsa) ne possède pas de racine unitaire, donc elle est stationnaire.

-On a p-value < 0.05 de constant alors elle contient de constant

Donc Les paramètres sera estimer par un modèle avec constant

III.4 Spécification et estimation du modèle:

comme la série initiale est affectée d'une saisonnalité de période $S=12$ on va modéliser par SARIMA(p,d,q)

A partir du correlogramme du variable stationnaire on détermine les retards du modèle ARMA AR(1) AR(4) AR(12) MA(1) MA(4) MA(12) alors on trouve ARMA(1,4) ARMA(1,12).....ect

D'après l'estimation des modèles précédant, on choisit le modèle qui sont coefficients significativement différents de 0 est SARIMA(1,1,12)

Dependent Variable: LELECSA				
Method: Least Squares				
Date: 04/13/18 Time: 19:56				
Sample (adjusted): 2009M02 2016M12				
Included observations: 95 after adjustments				
Convergence achieved after 11 iterations				
MA Backcast: 2008M02 2009M01				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.103609	0.006140	16.87407	0.0000
AR(1)	0.480632	0.096576	4.976713	0.0000
MA(12)	-0.878374	0.028776	-30.52409	0.0000
R-squared	0.464355	Mean dependent var	0.101356	
Adjusted R-squared	0.452710	S.D. dependent var	0.095369	
S.E. of regression	0.070553	Akaike info criterion	-2.433836	
Sum squared resid	0.457950	Schwarz criterion	-2.353187	
Log likelihood	118.6072	Hannan-Quinn criter.	-2.401248	
F-statistic	39.87772	Durbin-Watson stat	2.023688	
Prob(F-statistic)	0.000000			
Inverted AR Roots	.48			
Inverted MA Roots	.99	.86-.49i	.86+.49i	.49-.86i
	.49+.86i	.00+.99i	-.00-.99i	-.49+.86i
	-.49-.86i	-.86+.49i	-.86-.49i	-.99

Tableau III.4 : Estimation des paramètres du modèle SARIMA(1,1,12)

III.5 Validation du modèle SARIMA (1,1,12)

Dans cette étape, il convient maintenant d'analyser le résidu de notre modèle par le corrélogramme et teste de normalité (jak-pera, skwenss, kurtosis.....)

Teste sur les résidus

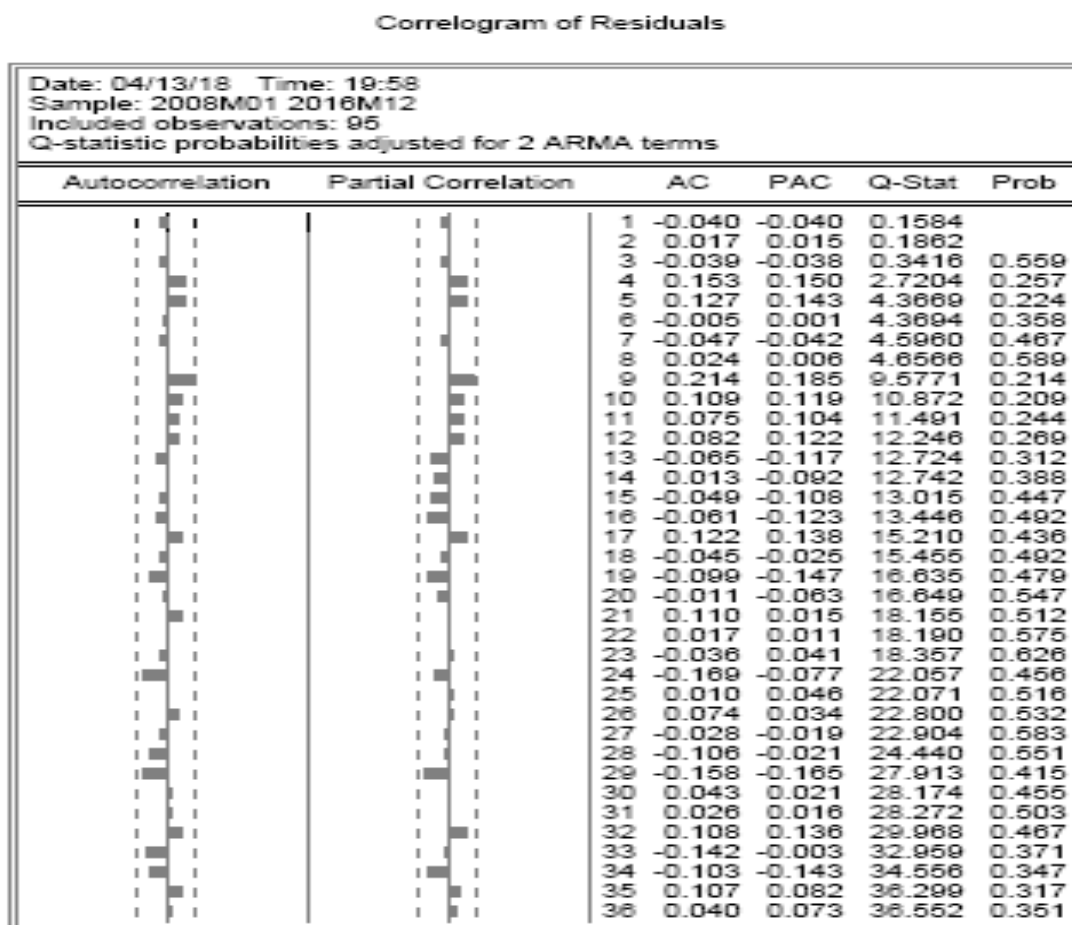


Figure III.5:Corrélogramme des résidus

La statistique Q à une probabilité très supérieur a 0.05 quelque soit la période, et on remarque que tous les pics sont à l'intérieur de l'intervalle de confiance, le résidu peut être assimilé à un bruit blanc.

Test de normalité des résidus

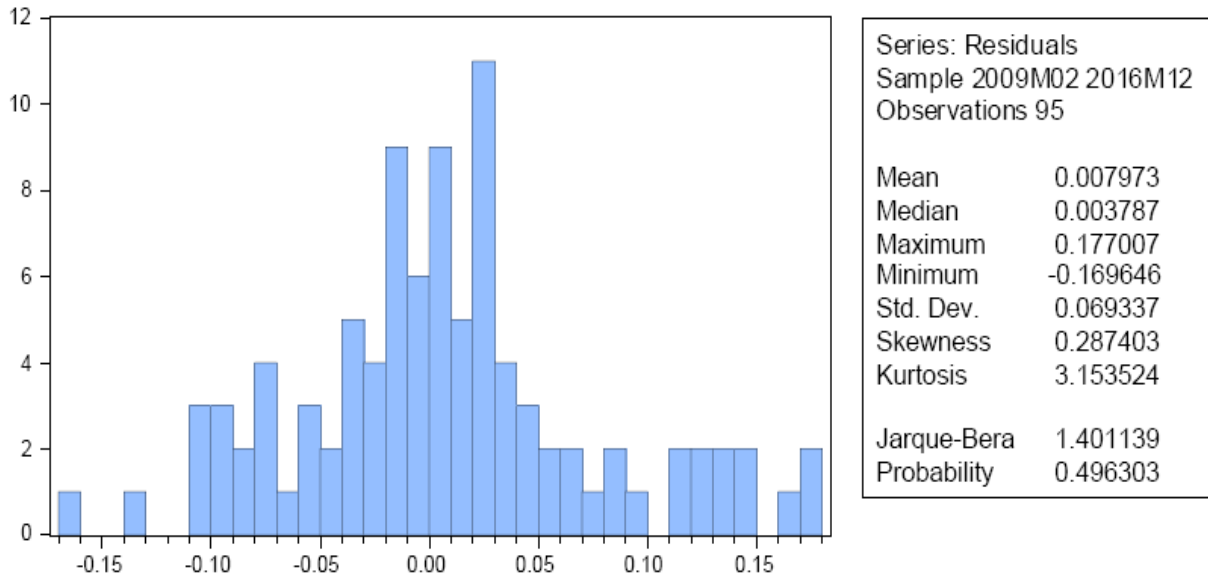


Figure III.6:Caractéristique statistique des résidus de la modélisation SAIRMA

Pour skewness:

$$\begin{cases} H_0: V_1 = 0 \text{ (la distribution des résidus est symétrique)} \\ H_1: V_1 \neq 0 \text{ (la distribution des résidus est Asymétrique)} \end{cases}$$

$$V1 = \frac{|B_1^{1/2} - 0|}{\sqrt{6/108}} = \frac{0.1437015}{0.375} = 0.38 < 1.96 \text{ donc accepte les hypothèses de symétrie}$$

Pour kurtosis:

$$\begin{cases} H_0: V_2 = 0 \text{ (la distribution des résidus est relativement aplatie)} \\ H_1: V_2 \neq 0 \text{ (la distribution des résidus n'est pas aplatie)} \end{cases}$$

$$V2 = \frac{|B_1^{1/2} - 3|}{\sqrt{24/108}} = \frac{3.153524}{0.111} = 1.38 < 1.96 \text{ donc accept les hypothèses de symétrie}$$

Pour jaque-bera

$$JB = 1,40 < X_{0,05}^2(2) = 5,99$$

Donc le résidé de notre série est un bruit blanc gaussien.

III.6 La prévision

d'après l'estimation des paramètres et la validation du modèle SARIMA(1,1,12) la série (lelec) peut être valablement représentée par un processus de type SARIMA(p,d,q) et les valeurs prévis sont les suivants:

	Valeurs Prévisionnelles en 2017	Valeurs réelle en 2017
Janvier	455681468	422 067 224
Février	326033397	329 925 091
Mars	345648918	322 720 179
Avril	376133902	338 676 344
Mai	466635543	526 741 082
Juin	643148960	701 277 739
Juillet	902306987	802 858 527
Aout	744355134	779 698 507
Septembre	572141296	545 058 922
Octobre	434350587	364 717 230
Novembre	368379299	362 468 801
Décembre	397709127	441 727 165

Tableau III.6: Valeurs prévisionnelles et réelles en 2017 (Unité : KWh)

Une simple comparaison avec les valeurs réelles indique que les résultats obtenus sont satisfaisant (il y'a pas beaucoup de différence entre les valeurs réelles et les valeurs prévisionnelles)

Conclusion générale

Notre étude a porté sur l'analyse des séries chronologiques représentant l'évolution du transport mensuelle de l'énergie électrique de l'entreprise (GRTE/DTE-HM).

Nous avons appliqué la méthode de Box & Jenkins sur nos données.

Comme notre série était affectée d'une saisonnalité, nous étions donc obligés de la désaisonnaliser à l'aide de l'opérateur de différence saisonnière. Ensuite nous avons fait recours au test de Dickey-Fuller pour tester la stationnarité.

Nous avons modélisé nos données par les modèles SARIMA. Une fois le modèle choisi, estimé et validé nous avons calculé nos prévisions.

En effet, les valeurs obtenues par notre méthode de prévision ne sont loin des valeurs réelles. Ce qui confirme que la méthode utilisée est la plus adéquate pour effectuer notre étude.

Bibliographie

- AGNÉS LAGNAUX, ISMAG
[1] Renforcement Statistique séries chroniques
- ARTHUR CHRBIENIER,
[2] Cours de séries temporelles théorie et applications, volume1
- CORIENNE PERRAUDIEN,
Université Paris I
[3] Magistère d'économie - Deuxième année séries chronologiques quelques
éléments du cours Année 2004_2005
- FALORIN ARRAM
[4] Séries temporelles, Régression, Interpolation et Géostatistique
- HÉLÈNE HAMISULTANE
[5] Econométrie des séries temporelles
- J.J.DOUDIN, C. DUBY, S.ROBIN & T. Trécourt,
[6] Analyse de séries chronologiques
(INA_PG, Mathématiques) Mai1996.
- JULIEN JACQUES,
[7] Introduction aux séries temporelles
- OLIVIER ROUSTANT
[8] Introduction aux séries chronologiques Axe Méthodes des Statistiques et
Applications, notes des cours Novembre 2008
- BOURBONNAIS ET MICHEL TERRAZA
[9] Analyse des séries temporelles, dunod (paris 2004)
S. FREDERIC
[10] Médicalisation et prévision (série chronologique),
école des mines de Nancy
[11] V.MONBET, Modélisation des séries temporelles mastère statistique et
économétrie Notes des cours Mastère1-2017
[12] Organigramme GRTE

Sites et adresses web

- [13] www.sonelgaz.dz
- [14] [www. Wikipedia.com](http://www.Wikipedia.com)

Résumé

Dans ce travail, on a étudié une série temporelle pour prédiction, un utilisant la méthode de méthode de Box-Jenkins, le type de cette sèrié est de la forme ARMA et ARIMA.

summary

In this Works, we studied a time series to predict the prediction using.The Box-Jenkins the type of this series is ARMA or ARIMA form.

ملخص

تطرقنا في هذا العمل الى دراسة السلاسل الزمنية من اجل التنبؤ و ذلك باستخدام منهجية Box-Jenkins و التي تهدف الى نمذجة سلسلة زمنية باستخدام نماذج ARMA و ARIMA