

République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la
Recherche Scientifique



Ecole Nationale Polytechnique
Département de Génie Industriel

Mémoire du projet de fin d'études d'ingénieur

Thème

**Modélisation et prévision à moyen terme de la
consommation nationale du GPL carburant
2007-2011**

Présenté par :

M^{lle} Amira KETTAL

M^{lle} Lamia OUIS

Dirigé par :

M^{lle} N.ABOUN

M^{me} O.BELMOKHTAR

M. S.HAMDANI

Promotion : juin 2007

DEDICACES

A ma mère Saida et à mon père Haidar
A mes deux frères : Fahed et Khaled
A tous mes amis
A tous ceux que j'ai connu durant ma formation
Je dédie ce travail

Amira

A ma mère Safia et mon père A. Kader
A ma sœur Lydia et mon frère Lamine
A tous ceux que j'estime
A tous ceux qui m'ont soutenue
Je dédie ce travail

Lamia

REMERCIEMENTS

Nous tenons à adresser nos sincères remerciements à nos deux promotrices Melle N. ABOUN et Mme O. BELMOKHTAR pour leur soutien permanent, leurs conseils judicieux, leurs orientations, leur disponibilité et surtout leur patience qui nous ont permis de réaliser ce présent travail.

Nous tenons à remercier Mr. S.HAMDANI qui nous a proposé ce présent sujet, pour toute l'aide qu'il nous a apporté à la réalisation de ce dernier, toutes les fois où il a répondu patiemment à nos questions et nos interrogations, avec ses recommandations, et ses précieux conseils.

Nous tenons à témoigner notre reconnaissance à Mme S. MASSOUT du Ministère de l'Energie et des Mines ainsi qu'à Mr DAIF ALLAH de Naftal qui nous ont si gentiment aidé lors de la collecte des données.

Nous remercions également tous les professeurs du département de Génie Industriel de nous avoir formées.

Nous tenons à remercier les membres de nos familles qui nous ont soutenus tout au long de ces années d'études.

ملخص

إن الطلب على المواد البترولية المستعملة خصيصا في مجال النقل قد عرف في السنوات الأخيرة تطورا سريعا القدرات الوطنية في مجال التصفية تكاد إن لا تكون كافية لتلبية هذا الطلب على المدى المتوسط الطريقة الفعالة من أجل توسيع سوق الوقود في الجزائر تتمثل في ترقية استهلاك وقود "سير غاز" في هذا السياق قمنا بانجاز هذا العمل اهتمنا بشكل خاص بنمذجة و صياغة التنبؤات الخاصة بالاستهلاك الوطني "سير غاز" في قطاع النقل

الكلمات المفتاحية : النمذجة, التنبؤ, الوقود, سير غاز, ارتداد متعدد, شعاع ارتدادي تلقائي, بوكس و جنكنس

Résumé :

La demande en produits pétroliers, utilisés essentiellement dans le domaine des transports, a connu ces dernières années une évolution très rapide. Les capacités nationales de raffinage risquent de ne plus être suffisantes pour satisfaire cette demande à moyen terme. Une manière efficace de détendre le marché des carburants en Algérie serait entre autres, la promotion de la consommation du GPL carburant. C'est dans ce cadre que ce travail a été mené. Nous nous sommes intéressés à la modélisation et à la prévision de la consommation nationale du GPLc dans le secteur de transport.

Mots clés : modélisation, prévision, carburant, gaz de pétrole liquéfié, régression multiple, vecteur autorégressif, Box & Jenkins.

Abstract:

The demand for petroleum products, which are primarily used in the transportation field, knew these last years a very fast evolution. The national refining capacities could be no more sufficient to satisfy the medium-term demand. An effective way to loosen the field market of the fuels in Algeria would be to promote the consumption of the liquefied petroleum gas. It is within this framework that this study has been carried out. For that, we have been interested in modelling and forecasting the national consumption of the liquefied petroleum gas in the transportation field.

Key words: modelling, forecast, fuel, liquefied petroleum gas, multiple regression, autoregressive vector, Box and Jenkins.

TABLE DES MATIERES

INTRODUCTION GENERALE	12
PROBLEMATIQUE	14
CHAPITRE I : ETUDE DE L'EXISTANT	16
I. LE GPL CARBURANT	17
1. Historique et définition	17
2. Sources d'obtention du GPL	18
3. Pourquoi le GPLc ?	18
3.1 Le GPLc est écologique	18
3.2 Le GPLc est économique	19
3.3 Le GPLc est performant	21
3.4 Le GPL c est disponible	21
4. Comment faire pour rouler en GPLc ?	21
4.1 Les modalités	21
4.2 La sécurité	22
II. L'ETUDE DU MARCHÉ	23
1. Etude de la demande	23
1.1 L'évolution de la demande des carburants en Algérie	23
1.2 L'évolution des prix des carburants en Algérie	24
1.3 Evolution du parc automobile en Algérie	25
2.1 Etude de la production en Algérie	27
2.2 Etude de la distribution et de la commercialisation des carburants en Algérie ..	30
CHAPITRE II: LA MODELISATION DE LA CONSOMMATION DU GPLC	33
SECTION I : OUTIL DE MODELISATION : LA REGRESSION MULTIPLE.....	35
I. Présentation du modèle	35
II. Rôle du terme aléatoire	35
III. Estimation des paramètres du modèle	36
1. Hypothèses.....	36
2. Formulation des estimateurs.....	36
IV. Equation d'analyse de la variance	37
1. Le coefficient de détermination R^2	37
2. Le coefficient de corrélation ρ	38
V. Test de spécification de l'ajustement.....	38
1. Test global de Fisher : (le F- Test)	38
2. Test de significativité des variables : (le T- Test)	38
3. Test sur les résidus.....	39
4. Test de stabilité du modèle : Le test de CUSUM	43
VI. La prévision.....	43
SECTION II : MODELISATION DE LA CONSOMMATION DU GPLC PAR LA REGRESSION MULTIPLE.....	44
I. Segmentation de la demande du carburant.....	44
1. Segmentation selon le type d'utilisation.....	44
2. Segmentation selon le type de clients	45
II. Les variables explicatives de la consommation du GPLc.....	46
1. Le prix	46
2. La taille du parc automobile	47

3. Le taux de croissance de la taille du parc automobile	47
4. La population	47
5. La population active	47
6. Le PIB	47
7. La masse salariale	47
8. La distance	47
9. La taille de la population utilisant le Transport en Commun	48
III. Formulation économique	48
IV. Formulation économétrique	49
1. Les modèles testés	49
2. Critère de choix de la meilleure variante	52
V. Application	53
1. Estimation des paramètres du modèle simple	54
2. Estimation des paramètres du modèle logarithmique.....	60
3. Choix du modèle.....	62
4. Tests et validation du modèle simple	62
5. Interprétation des résultats.....	69
CHAPITRE III : LA PREVISION DE LA CONSOMMATION DU GPLC.....	71
SECTION I : ELEMENTS D'ANALYSE DES SERIES CHRONOLOGIQUES	74
I. Définitions et concepts de base	74
1. Série chronologique	74
2. Stationnarité	74
3. Fonction d'autocorrélation simple et fonction d'autocorrélation partielle.....	75
II. Modélisation des séries chronologiques	76
1. Modèle autorégressif AR (p).....	76
2. Modèle MA (q) (moving average : moyenne mobile)	77
3. Modèle mixte ARMA (mélange du processus AR et MA)	78
III. Séries non stationnaires	78
1. Analyse de la tendance	79
2. Test de racine unitaire.....	80
VI. Extension du modèle ARMA : (Le modèle ARIMA (p, d, q))	80
SECTION II : LA METHODOLOGIE DE BOX & JENKINS:	81
I. L'identification du modèle	81
II. Estimation des paramètres	82
III. Adéquation du modèle.....	82
1. Test sur les paramètres du modèle	82
2. Test sur les résidus	82
VI. Choix du modèle	82
1. Critère de pouvoir prédictif	83
2. Critère d'information	83
V. La prévision	83
SECTION III: MISE EN ŒUVRE DE LA METHODE DE BOX & JENKINS POUR LA PREVISION DE LA CONSOMMATION DU GPL C	85
I. Mise en œuvre de la méthodologie Box & Jenkins sur les séries	85
1. Etude de la série « la taille de la population utilisant le Transport en..... Commun » POP_{tc}	85
2. Etude de la série « la taille de la population active » POP_{act}	96
4. Etude de la série PIB	106

5. Etude de la série Prix essence P_{ess}	107
II. La prévision de la consommation du GPLc	107
CHAPITRE IV : LA MODELISATION VECTORIELLE AUTOREGRESSIVE	110
SECTION 1: PRINCIPES DE BASE DE LA MODELISATION VECTORIELLE AUTOREGRESSIVE ..	112
I. Représentation d'un modèle VAR	112
II. Condition de stationnarité du modèle	112
IV. Détermination du nombre de retard p.....	113
V. Dynamique d'un modèle VAR.....	113
V.1 Représentation VMA d'un processus VAR.....	113
V.2 Analyse des chocs	114
V.3 La décomposition de la variance	115
V.4 La causalité	115
SECTION 2 : ANALYSE VAR ENTRE LES VARIABLES	117
($P_{GAS}, C_{GPLC}, C_{ESS}, C_{GAS}$)	117
I. Spécification du modèle VAR	117
1. Stationnarisation des séries	117
2. Détermination de l'ordre du modèle VAR optimal	118
II. Analyse des relations dynamiques entre les variables	123
1. Test de causalité de Granger	123
2 l'Analyse impulsionnelle	126
2.1 La décomposition de la variance de l'erreur de prévision:.....	126
2.2 Les fonctions de réponses impulsionnelles	129
CONCLUSION GENERALE	135
LES ANNEXES	137
ANNEXE 1 :QUESTIONS / REPONSES SUR LE GPL C	138
ANNEXE 2 : LES DONNEES	142
ANNEXE 3 : GUIDE EVIEWS 5	144
ANNEXE 4 : TEST DE DICKEY- FULLER.....	150
ANNEXE 5 : APPLICATION DE LA METHODE DE BOX & JENKINS SUR LES SERIES T_{CR} , PIB ET P_{ESS}	153
ANNEXE 6 : STATIONNARISATION DES SERIES CONSOMMATIONS.....	180
ANNEXE 7 : LE QUESTIONNAIRE.....	191
BIBLIOGRAPHIE	195

LISTE DES FIGURES

CHAPITRE I :

FIGURE I. 1 : EVOLUTION DE LA PRODUCTION NATIONALE DU GPL PAR SOURCE	18
FIGURE I. 2 : EMISSION DU CO ₂ AU COURS DU CYCLE DE VIE DU CARBURANT (GASOIL, ESSENCE, GPLC).....	19
FIGURE I. 3 : ADMINISTRATION DES PRIX ET DES MARGES DES CARBURANTS	20
FIGURE I. 4 : EVOLUTION DE LA CONSOMMATION DES CARBURANTS	23
FIGURE I. 5 : EVOLUTION DES PRIX DES CARBURANT EN ALGERIE.....	24
FIGURE I. 6 : EVOLUTION DU PARC AUTOMOBILE TOTAL EN ALGERIE	25
FIGURE I. 7 : EVOLUTION DU PARC AUTO ESSENCE EN ALGERIE	26
FIGURE I. 8 : EVOLUTION DU PARC AUTO DIESEL EN ALGERIE.....	26
FIGURE I. 9 : LA REPARTITION DES DIFFERENTES RAFFINERIES EN ALGERIE.....	28
FIGURE I. 10 : REPARTITION DE LA PRODUCTION TOTALE PAR RAFFINERIE (2005)	29
FIGURE I. 11 : REPARTITION DE LA PRODUCTION PAR PRODUIT (2005).....	29

CHAPITRE II:

FIGURE II. 1 : AJUSTEMENT D'UN NUAGE DE POINTS PAR UNE DROITE	36
FIGURE II. 2 : SEGMENTATION SELON LE TYPE D'UTILISATION	45
FIGURE II. 3 : SEGMENTATION SELON LE TYPE DE CLIENTS	45
FIGURE II. 4 : SCHEMA SIMPLIFIE DE LA PROCEDURE DE LA SELECTION DU MEILLEUR MODELE	53
FIGURE II. 5 : REPRESENTATION GRAPHIQUE DE LA SERIE RESIDUELLE, ACTUELLE ET ESTIMEE.....	63
FIGURE II. 6 : CORRELOGRAMME DES RESIDUS.....	64
FIGURE II. 7 : TEST DE NORMALITE DES RESIDUS.....	65
FIGURE II. 8 : EVOLUTION DES COEFFICIENTS DE REGRESSION	67
FIGURE II. 9 : TEST DE <i>CUSUM</i> POUR UN SEUIL DE + OU - 5%	68

CHAPITRE III:

FIGURE III. 1 : LES METHODES DE PREVISIONS	72
FIGURE III. 2 : LES ETAPES DE LA DEMARCHE PREVISIONNELLE DE BOX & JENKINS	84
FIGURE III. 3 : REPRESENTATION GRAPHIQUE DE LA SERIE POP _{TC}	85
FIGURE III. 4 : CORRELOGRAMME DE LA SERIE LA SERIE POP _{TC}	86
FIGURE III. 5 : REPRESENTATION GRAPHIQUE DE DPOP _{TC}	87
FIGURE III. 6 : REPRESENTATION GRAPHIQUE DE LA SERIE DDPOP _{TC}	89
FIGURE III. 7: CORRELOGRAMME DE LA SERIE DDPOP _{TC}	90
FIGURE III. 8 : REPRESENTATION GRAPHIQUE DE LA SERIE RESIDUELLE, ACTUELLE ET ESTIMEE DE POP _{TC}	94
FIGURE III. 9 : CORRELOGRAMME DES RESIDUS DU MODELE ARIMA (1, 2, 6).....	94
FIGURE III. 10 : HISTOGRAMME DE LA DISTRIBUTION DES RESIDUS DU MODELE ARIMA (1, 2, 6)	95
FIGURE III. 11 : LA PREVISION DE LA POP _{TC}	96
FIGURE III. 12 : REPRESENTATION GRAPHIQUE DE LA SERIE POP _{ACT}	96
FIGURE III. 13 : CORRELOGRAMME DE LA SERIE POP _{ACT}	97
FIGURE III. 14: REPRESENTATION GRAPHIQUE DE LA SERIE DPOP _{ACT}	98
FIGURE III. 15 : CORRELOGRAMME DE LA SERIE DPOP _{ACT}	100
FIGURE III. 16 : REPRESENTATION GRAPHIQUE DE LA SERIE RESIDUELLE, ACTUELLE ET ESTIMEE DE POP _{ACT} ... 103	
FIGURE III. 17 : CORRELOGRAMME DES RESIDUS DU MODELE ARIMA (4, 1, 11).....	103
FIGURE III. 18 : HISTOGRAMME DE LA DISTRIBUTION DES RESIDUS DE ARIMA (4, 1, 11).....	104
FIGURE III. 19 : LA PREVISION DE LA POP _{ACT}	105
FIGURE III. 20 : LA PREVISION DE LA T _{CR}	106
FIGURE III. 21 : LA PREVISION DU PIB	106
FIGURE III. 22 : DE LA PREVISION DU P _{ESS}	107
FIGURE III. 23 : LA PREVISION DE LA CONSOMMATION DU C _{GPLC}	108

CHAPITRE IV :

FIGURE IV. 1: RACINES INVERSES DES POLYNOMES AUTOREGRESSIFS DU MODELE VAR(4).....	118
FIGURE IV. 2 : RACINES INVERSES DES POLYNOMES AUTOREGRESSIFS DU MODELE VAR(3).....	119

FIGURE IV. 3 : RACINES INVERSES DES POLYNOMES AUTOREGRESSIFS DU MODELE VAR(1)	119
FIGURE IV. 4 : CORRELOGRAMME DES RESIDUS DE LA SERIE DDC_{GPLC}	120
FIGURE IV. 5 : CORRELOGRAMME DES RESIDUS DE LA SERIE DC_{ESS}	120
FIGURE IV. 6 : CORRELOGRAMME DES RESIDUS DE LA SERIE DDC_{GAS}	121
FIGURE IV. 7 : CORRELOGRAMME DES RESIDUS DE LA SERIE DDP_{ESS}	121
FIGURE IV. 8 : LA CAUSALITE ENTRE LES VARIABLES SELON GRANGER	124
FIGURE IV. 9 : EVOLUTION HISTORIQUE DES CONSOMMATIONS DE CARBURANTS AVEC LE PRIX DU GASOIL	125
FIGURE IV. 10 : EFFET D'UN CHOC AU NIVEAU DE LA CONSOMMATION DE L'ESSENCE.....	130
FIGURE IV. 11: EFFETS D'UN CHOC AU NIVEAU DE LA CONSOMMATION DU GPLC.....	130
FIGURE IV. 12 : EFFETS D'UN CHOC AU NIVEAU DE LA CONSOMMATION DU GASOIL	131
FIGURE IV. 13 : EFFETS D'UN CHOC AU NIVEAU DE PRIX DU GASOIL	131
FIGURE IV. 14 : L'ACCELERATION DE LA CROISSANCE DU PRIX DU GASOIL	132
FIGURE IV. 15 : EVOLUTION HISTORIQUE DE LA CONSOMMATION DU GPLC AVEC LE PRIX DU GASOIL	132
FIGURE IV. 16 : EVOLUTION HISTORIQUE DE LA CONSOMMATION DE L'ESSENCE AVEC LE PRIX DU GASOIL	133

ANNEXES :

ANNEXE 1 :

FIGURE 1. 1 : LA PARTIE ARRIERE D'UN VEHICULE DE MOTORISATION GPL CARBURANT	139
FIGURE 1. 2: INSTALLATION GPL C SUR UNE CITROËN XANTIA BREAK	141

ANNEXE 3 :

FIGURE 3. 1: L'ENVIRONNEMENT DE L'ESPACE DE TRAVAIL (WORKFILE) DE EViews 5	144
FIGURE 3. 2: CREATION D'UN WORKFILE	145
FIGURE 3. 3: CREATION DES SERIES CONSTANTE ET RESIDUELLE PAR DEFAUT.....	145
FIGURE 3. 4: CREATION D'UNE SERIE.....	146
FIGURE 3. 5: LA SAISIE DES DONNEES D'UNE SERIE.....	146
FIGURE 3. 6: GRAPHE D'UNE SERIE	147
FIGURE 3. 7: CORRELOGRAMME D'UNE SERIE.....	147
FIGURE 3. 8 : GENERATION D'UNE NOUVELLE SERIE	147
FIGURE 3. 9: ESTIMATION D'UN MODELE	148
FIGURE 3. 10 : TESTS STATISTIQUES	148
FIGURE 3. 11: ESTIMATION DU MODELE VAR	149
FIGURE 3. 12: TEST DE GRANGER.....	149
FIGURE 3. 13: VALIDATION DES PARAMETRES ET L'ANALYSE DYNAMIQUE D'UN VAR.....	149

ANNEXE 5 :

FIGURE 5. 1 : REPRESENTATION GRAPHIQUE DE LA SERIE T_{CR}	153
FIGURE 5. 2 : CORRELOGRAMME DE LA SERIE T_{CR}	153
FIGURE 5. 3 : CORRELOGRAMME DE LA SERIE T_{CR}	155
FIGURE 5. 4 : REPRESENTATION GRAPHIQUE DE LA SERIE RESIDUELLE, ACTUELLE ET ESTIMEE DE T_{CR}	157
FIGURE 5. 5 : CORRELOGRAMME DES RESIDUS DU MODELE ARIMA (1, 0, 5).....	158
FIGURE 5. 6 : HISTOGRAMME DE LA DISTRIBUTION DES RESIDUS DU MODELE ARIMA(1,0,5).....	158
FIGURE 5. 7 : REPRESENTATION GRAPHIQUE DE LA SERIE PIB	159
FIGURE 5. 8 : CORRELOGRAMME DE LA SERIE PIB	159
FIGURE 5. 9: REPRESENTATION GRAPHIQUE DE LA SERIE $DPIB$	161
FIGURE 5. 10 : REPRESENTATION GRAPHIQUE DE LA SERIE $DDPIB$	162
FIGURE 5. 11: CORRELOGRAMME DE LA SERIE $DPIB$	164
FIGURE 5. 12: REPRESENTATION GRAPHIQUE DE LA SERIE RESIDUELLE, ACTUELLE ET ESTIMEE DE PIB	167
FIGURE 5. 13 : CORRELOGRAMME DES RESIDUS DU MODELE ARIMA (2, 2, 4).....	167
FIGURE 5. 14 : DISTRIBUTION DES RESIDUS DU MODELE ARIMA (2, 2, 4).....	168
FIGURE 5. 15 : REPRESENTATION GRAPHIQUE DE LA SERIE PESS	169
FIGURE 5. 16 : CORRELOGRAMME DE LA SERIE PESS.....	169
FIGURE 5. 17: REPRESENTATION GRAPHIQUE DE LA SERIE DP_{ESS}	171
FIGURE 5. 18 : REPRESENTATION GRAPHIQUE DE LA SERIE DDP_{ESS}	172

FIGURE 5. 19 : CORRELOGRAMME DE LA SERIE DDP_{ESS}	174
FIGURE 5. 20 : REPRESENTATION GRAPHIQUE DE LA SERIE RESIDUELLE, ACTUELLE ET ESTIMEE DE P_{ESS}	178
FIGURE 5. 21 : CORRELOGRAMME DES RESIDUS DU MODELE ARIMA(1, 2, 8).....	178
FIGURE 5. 22 : HISTOGRAMME DE LA DISTRIBUTION DES RESIDUS DU MODELE ARIMA (1, 2, 8).....	179

ANNEXE 7 :

FIGURE 7. 2 : GRAPHIQUE RELATIF A LA NON CONVOITISE DU GPLC	193
FIGURE 7. 3 : GRAPHIQUE RELATIF AUX TAUX DE SUBSTITUTION D'UN CARBURANT PAR LE GPLC	193
FIGURE 7. 4 : GRAPHIQUE RELATIF A LA PERCEPTION DE LA POLLUTION DU GPLC	194
FIGURE 7. 5 : IMPACT DE L'AUGMENTATION DU PRIX DU GASOIL	194
FIGURE 7. 6 : IMPACT DE L'AUGMENTATION DU PRIX DE L'ESSENCE	194

LISTE DES TABLEAUX

CHAPITRE I :

TABLEAU I. 1 : PRIX A LA POMPE DES DIFFERENTS CARBURANTS EN ALGERIE (2007).....	20
---	----

CHAPITRE II :

TABLEAU II. 1 : TEST DE DURBIN WATSON.....	40
TABLEAU II. 2 : 1 ^{ERE} ETAPE DE LA REGRESSION PAS A PAS	54
TABLEAU II. 3 : 2EME ETAPE DE LA REGRESSION PAS A PAS	55
TABLEAU II. 4 : 3EME ETAPE DE LA REGRESSION PAS A PAS	56
TABLEAU II. 5 : 4 ^{EME} ETAPE DE LA REGRESSION PAS A PAS	57
TABLEAU II. 6 : 5 ^{EME} ETAPE DE LA REGRESSION PAS A PAS	58
TABLEAU II. 7 : 6 ^{EME} ETAPE DE LA REGRESSION PAS A PAS	59
TABLEAU II. 8 : 1 ^{ERE} ETAPE DE LA REGRESSION PAS A PAS POUR LE MODELE LOGARITHMIQUE	60
TABLEAU II. 9 : 2 ^{EME} ETAPE DE LA REGRESSION PAS A PAS POUR LE MODELE LOGARITHMIQUE.....	61
TABLEAU II. 10 : CRITERE DE CHOIX DU MEILLEUR MODELE.....	62
TABLEAU II. 11: ESTIMATION DES PARAMETRES DU MODELE.....	62

CHAPITRE III :

TABLEAU III. 1 : TEST DE DICKEY-FULLER SUR LA POP_{TC}	86
TABLEAU III. 2 : TEST DE DICKEY FULLER SUR $DPOP_{TC}$	88
TABLEAU III. 3 : TEST DE DICKEY FULLER SUR $DDPOP_{TC}$	89
TABLEAU III. 4 : ESTIMATION DU MODELE ARIMA (1, 2, 2) DE LA SERIE $DDPOP_{TC}$	91
TABLEAU III. 5 : ESTIMATION DU MODELE ARIMA (1, 2, 7) DE LA SERIE $DDPOP_{TC}$	91
TABLEAU III. 6 : ESTIMATION DU MODELE ARIMA (1, 2, 0) DE LA SERIE $DDPOP_{TC}$	92
TABLEAU III. 7 : CHOIX DU MEILLEUR MODELE DE LA SERIE $DDPOP_{TC}$	93
TABLEAU III. 8 : LA PREVISION DE LA POP_{TC}	96
TABLEAU III. 9 : TEST DE DICKEY FULLER SUR POP_{ACT}	97
TABLEAU III. 10 : TEST DE DICKEY FULLER SUR $DPOP_{ACT}$	99
TABLEAU III. 11: ESTIMATION DU MODELE ARIMA (7, 1, 7) DE LA VARIABLE $DPOP_{ACT}$	100
TABLEAU III. 12: ESTIMATION DU MODELE ARIMA(4, 1, 11) DE LA VARIABLE $DPOP_{ACT}$	101
TABLEAU III. 13 : ESTIMATION DU MODELE ARIMA (3, 1, 7) DE LA VARIABLE $DPOP_{ACT}$	101
TABLEAU III. 14 : CHOIX DU MEILLEUR MODELE POUR LA VARIABLE $DPOP_{ACT}$	102
TABLEAU III. 15 : PREVISION DE POP_{ACT}	105
TABLEAU III. 16 : PREVISION DU T_{CR}	106
TABLEAU III. 17 : PREVISION DU PIB	106
TABLEAU III. 18: LA PREVISION DU P_{ESS}	107
TABLEAU III. 19 : TABLEAU RECAPITULATIF DES PREVISIONS DES VARIABLES EXPLICATIVES	107
TABLEAU III. 20: LA PREVISION DE LA CONSOMMATION DU C_{GPLC}	108

CHAPITRE IV :

TABEAU IV. 1 : CHOIX DU RETARD P.....	118
TABEAU IV. 2 : TEST DE GRANGER	123
TABEAU IV. 3 : DECOMPOSITION DE LA VARIANCE DE LA VARIABLE DDC_{GPLC}	127
TABEAU IV. 4 : DECOMPOSITION DE LA VARIANCE DE LA VARIABLE DC_{ESS}	127
TABEAU IV. 5 : DECOMPOSITION DE LA VARIANCE DE LA VARIABLE DC_{GAS}	128
TABEAU IV. 6 : DECOMPOSITION DE LA VARIANCE DE LA VARIABLE DDP_{GAS}	129

ANNEXES :

ANNEXE 2 :

TABEAU 2. 1 : DONNEES HISTORIQUES DES DONNEES 1980-2006.....	143
--	-----

ANNEXE 5 :

TABEAU 5. 1 : TEST DE DICKEY- FULLER SUR T_{CR}	154
TABEAU 5. 2 : ESTIMATION DU MODELE ARIMA (1,0,4) DE T_{CR}	155
TABEAU 5. 3 : ESTIMATION DU MODELE ARIMA (1,0,5) DE T_{CR}	156
TABEAU 5. 4 : CHOIX DU MEILLEUR MODELE POUR T_{CR}	157
TABEAU 5. 5 : TEST DE DICKEY FULLER SUR PIB.....	160
TABEAU 5. 6 : TEST DE DICKEY- FULLER SUR DPIB	161
TABEAU 5. 7 : TEST DE DICKEY- FULLER SUR DDPIB	163
TABEAU 5. 8 : ESTIMATION DU MODELE ARIMA (2, 2, 2) DU PIB.....	164
TABEAU 5. 9 : ESTIMATION DU MODELE ARIMA(2,2,0) DU PIB	165
TABEAU 5. 10 : ESTIMATION DU MODELE (2, 2, 4) DU PIB.....	165
TABEAU 5. 11 : CHOIX DU MEILLEUR MODELE POUR LE PIB.....	166
TABEAU 5. 12 : TEST DE DICKEY FULLER SUR PESS	170
TABEAU 5. 13 : TEST DE DICKEY FULLER SUR DP_{ESS}	171
TABEAU 5. 14 : TEST DE DICKEY FULLER SUR DDP_{ESS}	173
TABEAU 5. 15 : ESTIMATION DU MODELE ARIMA(1,2,1) DU DDP_{ESS}	174
TABEAU 5. 16 : ESTIMATION DU MODELE ARIMA (1, 2, 8) DU DDP_{ESS}	175
TABEAU 5. 17 : ESTIMATION DU MODELE ARIMA(2,2,11) DU DDP_{ESS}	175
TABEAU 5. 18 : ESTIMATION DU MODELE ARIMA (9, 2, 3) DU DDP_{ESS}	176
TABEAU 5. 19 : CHOIX DU MEILLEUR MODELE DE DDP_{ESS}	177

ANNEXE 6 :

TABEAU 6. 1 : TEST DE DICKEY- FULLER SUR C_{GPLC}	180
TABEAU 6. 2 : TEST DE DICKEY- FULLER SUR DC_{GPLC}	181
TABEAU 6. 3 : TEST DE DICKEY- FULLER SUR DDC_{GPLC}	182
TABEAU 6. 4 : TEST DE DICKEY- FULLER SUR C_{GAS}	183
TABEAU 6. 5 : TEST DE DICKEY- FULLER SUR DC_{GAS}	184
TABEAU 6. 6 : TEST DE DICKEY- FULLER SUR DDC_{GAS}	185
TABEAU 6. 7 : TEST DE DICKEY- FULLER SUR C_{ESS}	186
TABEAU 6. 8 : TEST DE DICKEY- FULLER SUR DC_{ESS}	187
TABEAU 6. 9 : TEST DE DICKEY- FULLER SUR $PGAS$	188
TABEAU 6. 10 : TEST DE DICKEY- FULLER SUR DP_{GAS}	189
TABEAU 6. 11 : TEST DE DICKEY- FULLER SUR DDP_{GAS}	190

Introduction générale

L'un des sujets d'actualité le plus présent dans le monde est celui du réchauffement climatique. Ce thème a mobilisé un grand nombre de scientifiques qui lui ont consacré autant d'études que de recommandations. Parmi les causes de ce réchauffement qui sont citées, on évoque naturellement et souvent les émissions de dioxyde de carbone, dues à la consommation excessive de carburants, essence et diesel, mais aussi à un nombre en continue croissance de véhicules de toutes sortes à travers le monde. Si ces émissions sont fréquemment citées, c'est parce qu'elles sont vécues au quotidien par chacun de nous.

Pour répondre à cette situation de plus en plus alarmante et en raison également de l'envolée des prix des produits pétroliers ces dernières années, certains pays d'Amérique Latine et d'Asie, se sont attelés à promouvoir les carburants propres et les biocarburants, dont la demande devenait manifeste dans les pays développés qui doivent faire face à une opinion très exigeante en matière d'environnement. Cependant, de plus en plus de voix s'élèvent pour mettre en garde l'opinion mondiale sur les risques de dérèglements écologiques et économiques, dus à une exploitation extensive de ces produits qui se traduirait par une déforestation massive et l'orientation accrue de ces cultures vers la production de carburants, au détriment de l'alimentation.

Ainsi, la problématique des carburants propres de substitution à l'essence et au diesel reste entière et c'est dans ce cadre que s'inscrit notre étude qui se propose de prévoir la place du GPL carburant, en tant que carburant alternatif. Sachant que l'Algérie le produit dans la foulée du processus d'extraction des produits pétroliers, ses réserves très importantes, pourraient satisfaire la demande nationale si cette dernière devait connaître une augmentation significative.

Actuellement la consommation du GPLc est très faible et ne dépasse pas les trois millions de tonnes, le souci de la SH est de savoir comment elle va évoluer cette consommation dans les 5 prochaines années. Pour ce faire, l'utilisation des méthodes de prévisions fiables s'impose.

C'est dans ce cadre que s'inscrit ce présent travail dont l'objectif est de faire un état des lieux sur la consommation du GPLc, d'essayer de déterminer les facteurs qui l'influencent afin de la modéliser et la prévoir.

Notre document sera structuré de la manière suivante :

Dans le premier chapitre, nous donnerons un bref historique sur la genèse du GPL, en tant que carburant. Il vise à présenter ses caractéristiques essentielles qui font de lui un carburant qui pourrait s'avérer très avantageux au niveau national. Nous nous intéresserons ensuite au marché algérien des carburants, afin de définir la structure de l'offre et de la demande et ses tendances dans le but de comprendre le fonctionnement de ce marché.

Dans le second chapitre, nous présentons, en première partie, l'outil de modélisation qui est la régression multiple, en insistant sur son caractère adapté aux grands phénomènes économiques. Dans ce cadre, nous passerons en revue, d'une manière succincte les principes de base de la modélisation, fondée sur la régression multiple, en seconde partie, nous traitons la modélisation de la consommation du GPLc, basée sur la régression économétrique. Le but étant de concevoir un modèle explicatif de la consommation de ce dernier, en tenant compte des différents paramètres qui sont susceptibles de l'influencer.

Le troisième chapitre sera consacré à la prévision proprement dite de la consommation nationale du GPLc. Nous y développons les fondements théoriques et les principes de base de l'une des méthodes les plus pertinentes de prévision, à savoir la méthode de Box & Jenkins. Cette dernière sera mise en oeuvre pour prévoir les variables explicatives du modèle élaboré, afin d'aboutir à la prévision de la consommation du GPLc.

A prendre comme élément complémentaire à l'analyse du résultat de la prévision, nous avons élaboré un questionnaire destiné à établir les motivations des usagers de véhicules automobiles quant à leur préférence d'un type de carburant sur un autre.

Le quatrième et dernier chapitre, est consacré à l'étude de la relation qui existe entre la consommation du GPLc et un ensemble de variables préalablement choisies tel que : le prix du gasoil etc ... Il s'agira ensuite d'étudier les liaisons dynamiques à l'intérieur du groupe formé par ces variables, de chercher à identifier et à comprendre les sources de fluctuation de ces dernières.

Enfin, pour achever notre travail, nous ferons une conclusion générale concernant les différents chapitres.

Problématique

L'énergie doit contribuer à accroître le niveau de qualité de vie ainsi qu'à fournir des ressources nécessaires à un monde de plus en plus respectueux de l'environnement.

Dans la structure d'apport énergétique, le pétrole occupe une place privilégiée, un certain nombre de questionnement se posent sur l'épuisement des réserves, leur adaptation aux contraintes environnementales et aux exigences du marché.

La chaîne pétrolière est formée de plusieurs maillons, qui requièrent des investissements lourds et indispensables à la réalisation d'un objectif unique : fournir au consommateur les produits pétroliers qu'il demande.

Ces maillons se composent notamment de :

- L'exploration (recherche de nouveaux gisements) et l'extraction ;
- Le transport du pétrole vers les raffineries (pipe marin, pipe terrestre) ;
- Le raffinage (traitement du pétrole pour fabriquer les différents produits) : c'est l'ensemble des opérations visant à transformer le pétrole brut extrait du gisement en produits utilisables : combustibles, carburants, produits de base de l'industrie chimique...
- Le stockage et la distribution vers les consommateurs.

Parmi les produits pétroliers issus de cette chaîne on distingue les carburants, qui sont des combustibles alimentant les moteurs thermiques. Ils transforment l'énergie chimique contenue dans le carburant en une énergie mécanique. Valorisés notamment dans les transports, les carburants permettent de réaliser des déplacements, offrant aux consommateurs des services conséquents et reconnus.

Le parc de transport national a connu ces dernières années un essor remarquable, entraînant ainsi une très forte pression sur les consommations de carburants. Cette consommation est appelée à croître avec la croissance attendue et le développement du niveau de vie des citoyens.

Les capacités nationales de raffinage sont actuellement suffisantes pour satisfaire cette demande, mais risquent de ne plus l'être à moyen et long termes. Il est probable que l'Algérie devienne un pays importateur de produits pétroliers (notamment pour le diesel dont la demande affiche un taux très élevé) si les tendances actuelles persistent.

En effet, la demande en produits pétroliers, utilisés essentiellement dans le domaine des transports, a connu une augmentation de 4% en 2005 liée principalement au développement continu du parc automobile caractérisé par une tendance de plus en plus prononcée vers le Diesel, qui représente aujourd'hui 74% de la consommation des carburants.

Pour pallier à ce problème, la Sonatrach s'est engagée à développer sa capacité de raffinage pour répondre à la demande et satisfaire les exigences du marché national et international. A cet effet, plusieurs projets de renforcement des capacités de raffinage ont été lancés, (ex : raffinerie d'Adrar, projet de raffinerie à Tiaret...).

Cependant, une manière plus efficace de détendre le marché des carburants en Algérie serait entre autres, la promotion de la consommation du GPL carburant.

En effet, ce carburant présente plusieurs avantages dont les principaux sont :

- Les GPL ne sont pas uniquement les produits des opérations de raffinage, puisqu'une grande partie est produite en amont au niveau des gisements, ce qui permet d'avoir des coûts de production plus faibles que ceux des produits raffinés.
- Le potentiel dont dispose l'Algérie en GPL est très important.
- Le GPLc est un carburant propre, avec des taux d'émissions beaucoup plus faibles que ceux du gasoil ou des essences;
- La consommation du GPLc, permettrait de dégager des capacités d'exportation de produits pétroliers, qui pourraient être mieux valorisés sur le marché international.

Afin de prendre des décisions pour le développement du GPL carburant sur le marché national, il est important de comprendre la dynamique d'évolution de la demande sur ce produit. Pour cela il s'agira d'apporter des éléments de réponse aux questions suivantes :

- Quels sont les déterminants de cette consommation ?
- Comment peut-on modéliser et prévoir cette consommation ?

C'est dans ce cadre que s'inscrit la présente étude dans laquelle nous nous intéresserons à

« La modélisation et la prévision à moyen terme de la consommation nationale du GPL carburant 2007- 2011 »

Chapitre I : Etude de l'existant

Dans ce chapitre, nous présentons les caractéristiques essentielles du GPLc qui font de lui un carburant très avantageux, par la suite nous nous intéresserons au marché algérien des carburants avec l'évaluation de l'offre et de la demande afin de comprendre le fonctionnement de ce marché.

I. Le GPL carburant (GPLc) :

1. Historique et définition : [UFIP], [CFBP]

Le 24 décembre 1910, un ingénieur de la Riverside Oil Company a découvert les Gaz de pétrole liquéfié (GPL). En effet, H. Stukeman réussit pour la première fois à obtenir les 658 premiers litres de GPL liquide

Cet ingénieur avait entrepris ses travaux sous l'impulsion du Directeur Général (A.N. Kerr), de la raffinerie. Ce dernier se trouvait depuis un certain temps confronté au problème de l'évaporation des essences minérales (extraites du gaz naturel associé aux gisements de pétrole) qui étaient stockées dans des bacs à ciel ouvert.

Pour limiter les pertes, il eut l'idée de récupérer les vapeurs et d'essayer de les recondenser. Ainsi étaient nés les Gaz de Pétrole Liquéfiés (**GPL**).

Les "GPL" sont des hydrocarbures composés majoritairement de coupes en C₃ et C₄, le propane (C₃H₈) et le butane (C₄H₁₀), étant les principaux composants. Ces produits, gazeux dans les conditions normales de température et de pression, peuvent être facilement liquéfiés rendant ainsi leur transport plus aisé.

C'est en 1912 que fut expérimentée la première installation domestique fonctionnant au GPL.

Dans la même année, la carburation automobile au GPL fit ses premiers pas. Le pré-développement de cette invention a pris cependant une dizaine d'années.

En 1922, les statistiques officielles font état de 417 tonnes de GPL vendues dans l'ensemble des Etats-Unis. Les compagnies pétrolières commencent à s'intéresser au GPL et le commercialisent dès 1927. L'impulsion est alors donnée en 1930, 20 ans après leur découverte.

Dans le langage courant, les GPL sont devenus le GPL, même s'il s'agit de plusieurs gaz.

Le "GPL" signifie "Gaz de Pétrole Liquéfié". Ces gaz qui, par leurs utilisations, sont omniprésents dans notre vie quotidienne (bouteilles de gaz butane, GPLc, etc...).

Cependant, le GPLc a été créé pour répondre à une problématique qui n'a pas perdu de son actualité. Il s'agissait de concevoir un carburant propre, peu coûteux, facile à produire à partir de ressources naturelles disponibles. Le GPL répondrait dans une large mesure à ces critères.

Le GPLc est un mélange constitué d'environ **50 %** de butane et **50 %** de propane.

A l'heure des débats sur les effets des émissions dans l'atmosphère, dues à l'utilisation de différents types énergies, sur la vie de notre planète, le GPL carburant s'impose comme un carburant d'une performance technique acceptable dans les moteurs thermiques, économique et surtout écologique.

2. Sources d'obtention du GPL :

Le GPL est issu soit : [SH]

- Des unités de séparations (au niveau des gisements) ;
- Du raffinage du pétrole brut ;
- Du traitement du gaz naturel (la liquéfaction) ;

En Algérie, la plus grande partie produite du GPL est obtenue au niveau des unités de séparations (installées souvent à proximité des gisements). Par ailleurs, les GPL représentent entre 2 et 3 % de l'ensemble des produits obtenus, lors des opérations de raffinage et, selon sa provenance, une tonne de pétrole brut peut donner de 20 à 30 Kg de GPL. [SFC]

Nous pouvons illustrer cela par la *figure I.1*, qui représente la production nationale du GPL par source (de 1993 à 2002) (en milliers de tonnes).

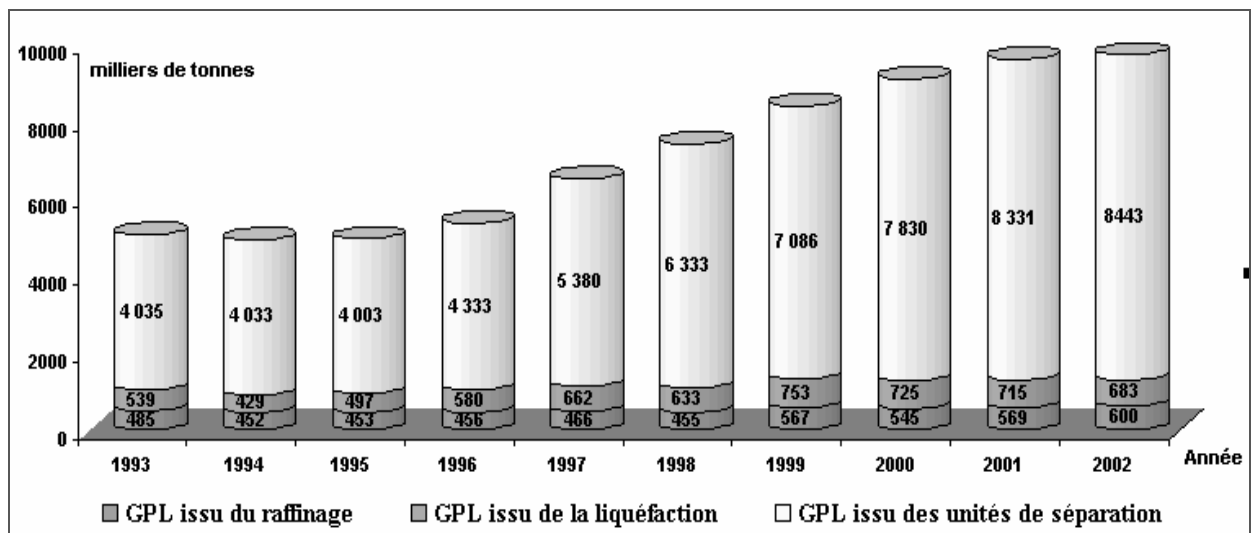


Figure I. 1 : Evolution de la production nationale du GPL par source [MEM]

On remarque bien que la production nationale du GPL est beaucoup plus importante au niveau des unités de séparation qu'au niveau du raffinage et de la liquéfaction.

3. Pourquoi le GPLc ?

Plusieurs arguments sont favorables au GPLc et le présentent comme un carburant avantageux ayant de bonnes performances techniques, économiques et environnementales :

3.1 Le GPLc est écologique : [PRIMAGAZ]

Le GPLc est un produit écologique, il représente le carburant alternatif disponible pour le grand public, le plus respectueux de l'environnement :

- Mélange de butane et de propane, le GPLc ne contient en effet ni plomb, ni benzène, ni soufre, dangereux pour l'environnement.
- La combustion du GPLc ne produit pas de particules et limite significativement les émissions de CO₂ (dioxyde de carbone), un des gaz responsables de l'effet de serre. Conséquence : les voitures roulant au GPLc ne sont pas interdites de circulation en cas de forte pollution.
- Comparé au diesel, le GPLc produit 20 fois moins de Nox (abréviation utilisée dans le domaine de pollution et qui regroupe deux molécules le NO et le NO₂, respectivement monoxyde et dioxyde d'Azote responsables des pluies acides et des pics de pollution urbaine mais surtout de maladies respiratoires graves).
- Ses caractéristiques intrinsèques permettent aux moteurs de répondre aux normes antipollution les plus sévères.
- Le GPLc apparaît donc comme une réponse adaptée tant pour les problèmes de qualité d'air que pour les émissions de gaz à effet de serre.

La *figure I.2* représente une étude menée par le groupe pétrolier TOTAL et qui compare le taux d'émission du CO₂ de chaque carburant : essence, gasoil et le GPL c. [TOT 04]

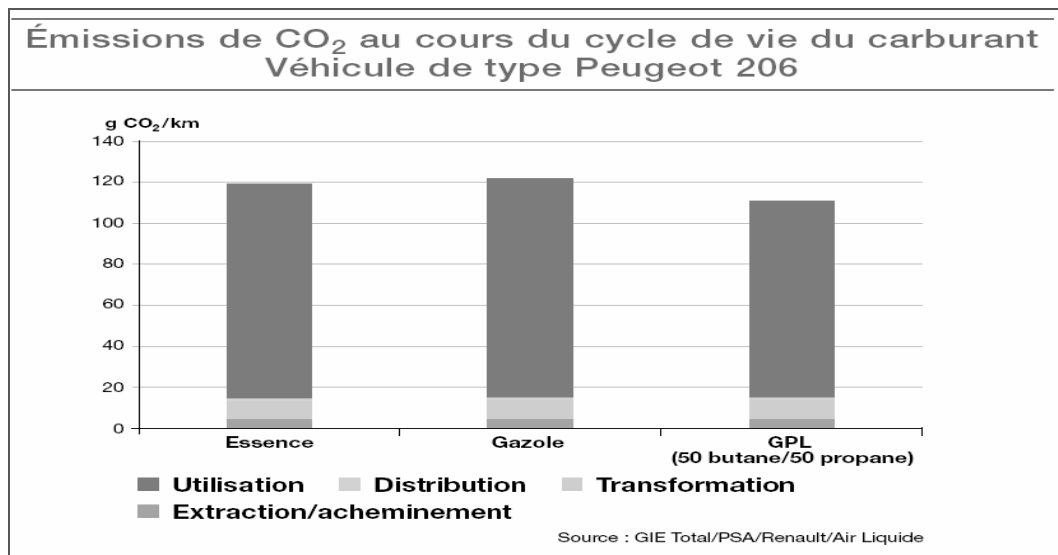


Figure I. 2 : Emission du CO₂ au cours du cycle de vie du carburant (gasoil, essence, GPLc)

3.2 Le GPLc est économique :

Actuellement, les prix des carburants sont administrés sur le marché national, à tous les niveaux de la chaîne pétrolière, allant du puits au consommateur final, à savoir : [NAFTAL]

- Le prix du pétrole brut à l'entrée de raffinerie ;
- Le prix des produits raffinés en sortie d'usine ;
- Le prix des produits aux grossistes ;
- Le prix des produits aux consommateurs finaux.

La *figure I.3* illustre l'administration des prix et des marges des carburants en Algérie:

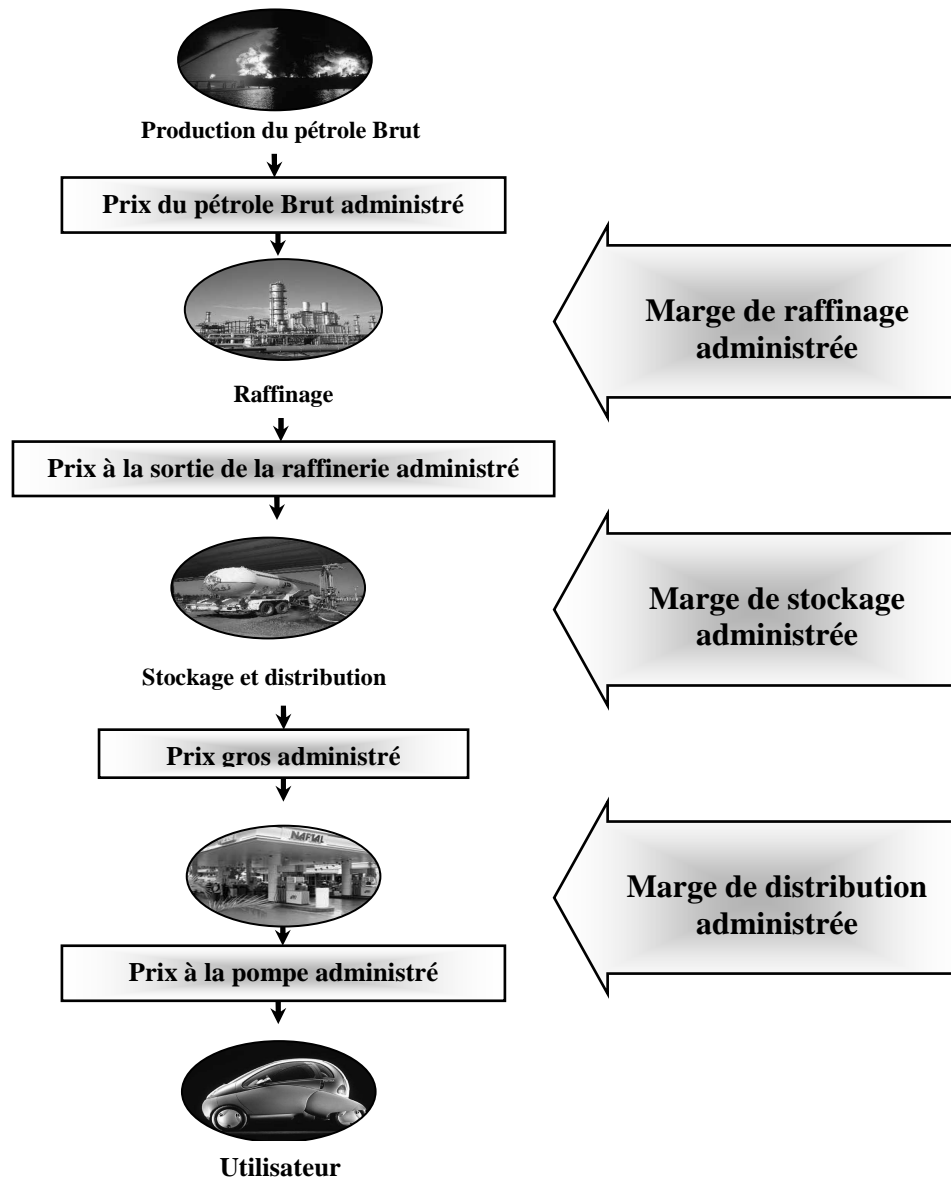


Figure I. 3 : Administration des prix et des marges des carburants

Comme les GPL ne sont pas seulement des produits du raffinage du pétrole et que la plus grande partie est produite plus en amont, au niveau des gisements, ceci permet d'avoir des coûts de production plus intéressants que ceux des produits raffinés, faisant ainsi du GPLc le carburant le moins cher au niveau du marché.

Le *tableau I.1* illustre le prix affiché à la pompe des différents carburants distribués en Algérie :

Tableau I. 1 : Prix à la pompe des différents carburants en Algérie (2007)

<i>Carburant</i>	<i>Gasoil</i>	<i>Essence super</i>	<i>Essence normale</i>	<i>Essence sans plomb</i>	<i>GPL c</i>
<i>Le prix (DA/l)</i>	13,7	23	21,20	22,60	9

On voit bien que le GPLc représente le carburant le moins cher au niveau du marché.

3.3 Le GPLc est performant : [PRIMAGAZ]

- Le GPLc ne produit pas de calamine (un composé de résidus de combustion, provenant essentiellement de l'huile), c'est un carburant propre et le moteur n'est pas sujet à l'encrassement que connaissent les moteurs classiques ;
- Les véhicules roulant au GPLc affichent des performances confortées par les qualités techniques qu'offre ce carburant : moteur plus silencieux, plus souple et une décélération en douceur grâce à l'homogénéité de la combustion ;
- Les caractéristiques du moteur GPLc rendent son fonctionnement moins bruyant. Il faut savoir également que la réduction des vibrations diminue sensiblement les nuisances sonores, améliorant ainsi le confort et la souplesse de la conduite ;
- Le remplissage du réservoir se fait sans odeur ;
- Le GPLc prolonge la durée de vie du lubrifiant et du moteur ; sa combustion ne laisse aucun dépôt ; les révisions peuvent être espacées à 30000 Km. Ainsi l'acquisition d'un véhicule GPLc peut être amortie très rapidement ;
- Les huiles n'étant pas altérées, elles conservent plus longtemps leurs qualités lubrifiantes ;
- Enfin, la plupart des véhicules étant « bi- carburant », présentant de ce fait deux réservoirs (un pour l'essence et un pour le GPLc), leur autonomie est augmentée de façon significative.

3.4 Le GPL c est disponible :

Les réserves du GPL dont dispose l'Algérie sont très importantes (8,8 Millions de tonnes par an [SH 05]), pouvant satisfaire la demande nationale en carburant.

Sa consommation permettrait de dégager des capacités d'exportation de produits pétroliers, qui sont nettement mieux valorisés sur le marché international.

Tous ces arguments font que la promotion de l'utilisation du GPLc, en tant que carburant, peut générer des avantages économiques conséquents.

4. Comment faire pour rouler en GPLc ?

4.1 Les modalités : [CFBP]

Deux modalités s'offrent pour rouler au GPL carburant :

- ***Acheter un véhicule neuf chez un constructeur :***

Ce véhicule aura été assemblé :

- soit en « ***1ère monte*** » (directement sur une chaîne de fabrication spécifique dans les usines du constructeur),
- soit en « ***1ère monte décalée*** » (transformé après la fin de sa fabrication par un installateur indépendant agréé, mais avant sa 1ère mise en circulation ; sa commercialisation est toujours assurée par le réseau habituel de vente du constructeur).

- **Faire transformer un véhicule essence déjà en circulation :**

Cette installation, dite « **2ème monte** », doit être réalisée par un installateur agréé.

4.2 La sécurité : [ANTAGAZ]

Malgré le phénomène qui colporte la peur de risques d'explosion non avérés, il y a lieu de savoir que :

- Depuis le 1er janvier 2000, les réservoirs des nouveaux véhicules, roulant au GPLc, doivent être équipés **d'une soupape de sécurité** obligatoire en Algérie. En cas d'incendie, survenant à l'arrêt ou après un choc, les risques d'explosion sont écartés par cette dernière.
- Par ailleurs, des évaluations chez les constructeurs comme chez les équipementiers ont été faites. Elles démontrent que les risques ne sont pas supérieurs à ceux encourus à bord d'un véhicule essence classique.
- En cas de forte collision, le réservoir, fabriqué en acier et d'une épaisseur de 2 à 3,5 mm (soit 6 fois plus épais qu'un réservoir de carburant traditionnel), ne se déchire pas (même après un choc à 50 km/h contre un tranchant en fonte, il ne cède pas). A titre d'information, il est en moyenne 20 fois plus résistant au percement qu'un réservoir traditionnel.
- Un entretien du véhicule équipé du GPLc est nécessaire, surtout en ce qui concerne les bougies et le filtre à air. Une révision sur tout le matériel consommable est nécessaire tous les trois ans.

Enfin, outre la soupape de sécurité qui régule la pression à l'intérieur du réservoir, les réservoirs GPLc doivent être équipés de 4 accessoires contribuant à la sécurité :

1. Un clapet anti-retour ;
2. Un limiteur de remplissage à 80 % de la capacité totale du réservoir ;
3. Un limiteur de débit qui permet un débit de carburant régulier ;
4. Une électrovanne qui stoppe le gaz lorsqu'on roule en mode "essence" et quand le moteur cale.

Le tout devant être construit en fonction de normes très strictes.

D'autres précisions sur les caractéristiques du GPL carburant, devant répondre en principe à certaines interrogations d'ordre pratique sur le fonctionnement des véhicules à bicarburation que chacun de nous peut se poser ont été réunies. (cf. annexe 1)

Après avoir défini les caractéristiques principales du GPLc et explicité les avantages de son utilisation, nous proposons dans ce qui suit une analyse de la structure du marché des carburants en Algérie, qui pourrait nous éclairer sur l'approche méthodologique à adopter en vue de la modélisation et la prévision de la consommation du GPLc. Cette analyse pourrait entre autres nous aider à distinguer les variables pertinentes à intégrer dans notre étude.

II. L'étude du marché:

L'étude de marché se définit comme une étude qui regroupe un ensemble d'outils et de techniques permettant de rechercher des données quantitatives et qualitatives sur un marché et de les analyser dans le but de bien comprendre le fonctionnement de ce dernier [KAL 02]

1. Etude de la demande :

L'étude de la demande a pour objet d'examiner l'évolution de la demande des carburants sur le marché algérien, ainsi que les facteurs qui l'expliquent tels que le parc auto, le prix des carburants et le revenu des ménages.

1.1 L'évolution de la demande des carburants en Algérie :

La figure I.4, représente l'évolution de la consommation nationale des carburants (De 1980 à 2006) :

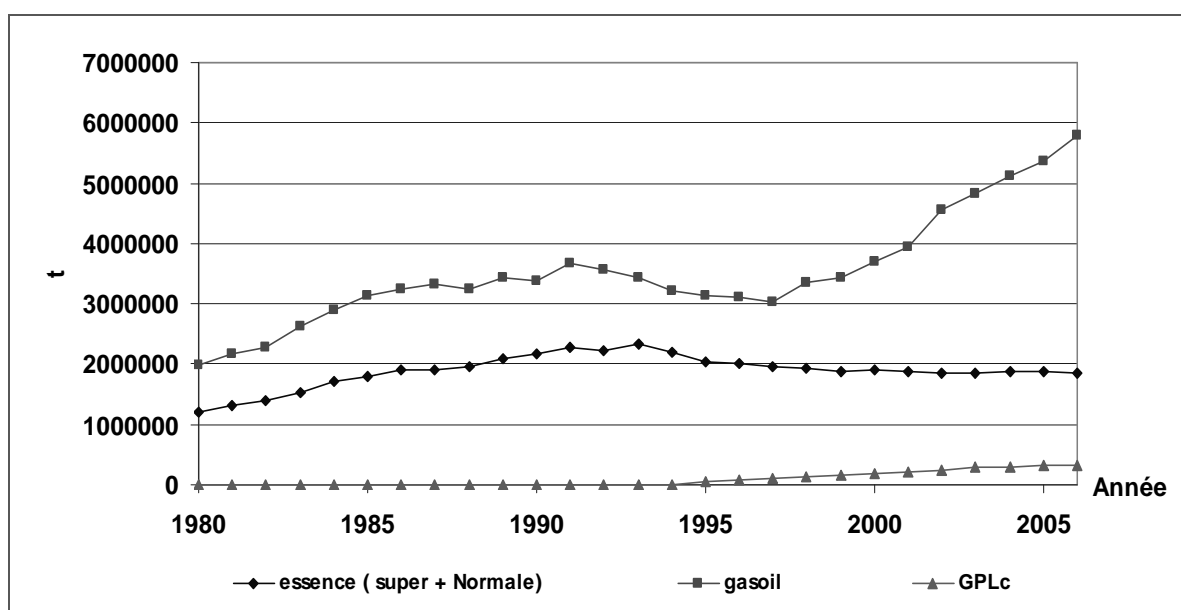


Figure I. 4 : Evolution de la consommation des carburants [MEM]

Après examen du graphe, nous pouvons identifier les périodes suivantes:

- De 1980 à 1991 : cette période se caractérise par une absence de consommation du GPLc et en même temps une croissance continue de la demande des autres carburants (essences + gasoil). En effet, on remarque que cette consommation a presque doublé, ceci étant dû surtout à la stabilité des prix et la croissance de la population active.

- De 1992 à 1997 : on remarque un début de consommation du GPLc, mais cette consommation reste très insignifiante. La consommation des autres carburants a sensiblement diminué, à hauteur de 14%, à cause de la crise économique et de l'augmentation des prix qui a engendré une diminution du pouvoir d'achat.
- De 1998-2006 : on remarque que durant cette période :
La consommation du GPLc a progressé pour atteindre plus de 0,3 Mt en 2006 mais elle reste nettement inférieure à la consommation des autres carburants. En effet :
 - La consommation de gasoil routier a presque doublé (80%), sur la même période, pour atteindre les 6Mt, sous l'influence des deux facteurs suivants:
 - Une multiplication par deux du taux de diésélisation du marché des véhicules particuliers en Algérie (7 % en 1985 à 14 % en 2003) ;
 - Le développement du trafic routier de marchandise par véhicule professionnel.
 - La demande en essence a légèrement diminué (2 Mt), ceci est dû à la diésélisation du parc automobile

1.2 L'évolution des prix des carburants en Algérie :

Le prix des carburants est sans doute le premier facteur influant le comportement de consommateurs.

La *figure I.5*, ci-dessous, représente l'évolution des prix des carburants en Algérie (de 1990 à 2006)

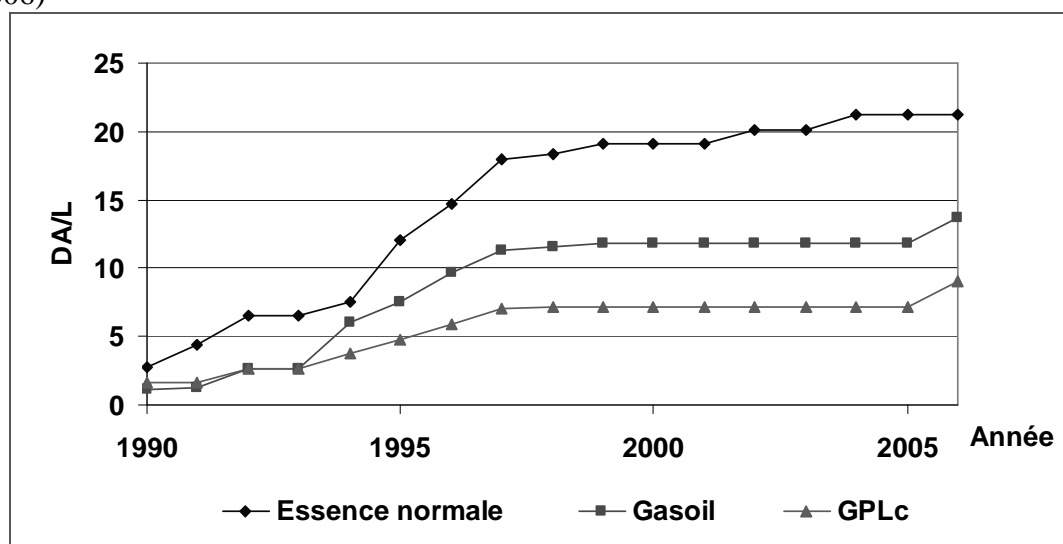


Figure I. 5 : Evolution des prix des carburant en Algérie
(Source: [Ministère de l'Energie et des Mines] [Les stations services])

Sur la base du graphe, on distingue les périodes suivantes :

- Jusqu'au début des années 90 les prix des différents carburants étaient très bas du fait de la subvention de l'Etat, destinée à encourager la consommation des carburants
- A partir de 1993, les prix des carburants ont connu une évolution rapide. En effet :

- Pour l'essence : le prix a fortement évolué allant de 7.00 DA/L en 1993 à 18.40 DA/L en 1998 soit une augmentation de 163 %
- pour le gasoil : le prix a aussi fortement évolué allant de 2.50 DA/l en 1993 à 11.50 DA/l en 1998 soit une augmentation de 360 % ;
- Pour le GPLc : le prix a évolué allant de 2.5 DA/L en 1993 à 7.2 DA/L en 1998, soit une augmentation de plus de 188 %. Néanmoins, il reste le carburant le moins cher.
- A partir de 1998, les prix ont continué leur progression mais de façon moins rapide et plutôt stable, en effet :
 - Le prix du gasoil s'est maintenu à 11,75 DA/L, ce qui a eu pour conséquence la multiplication par deux de la demande du gasoil entre 1998 et 2005 avant de connaître une hausse, en 2006, le plaçant à 13.7 DA/L.
 - Le prix de l'essence a continué à progresser jusqu'à atteindre le seuil de 21,20 DA/L en 2004, seuil auquel il se maintient à ce jour.
 - Quant au GPLc, son prix est resté stable à 7.2 DA/L jusqu'à 2005, avant d'atteindre 9DA/L en 2006, demeurant toujours le carburant le moins cher sur le marché national.

1.3 Evolution du parc automobile en Algérie :

Le deuxième facteur qui influence la consommation de carburants automobiles est le nombre total de véhicules en circulation (la taille du parc automobile). Aussi allons nous illustrer l'évolution du parc auto par type d'activité et par type de carburant utilisé. Mais avant cela, il conviendra d'analyser l'évolution du parc auto, dans sa totalité, telle que représentée par le graphe ci-dessous (*figure I.6*):

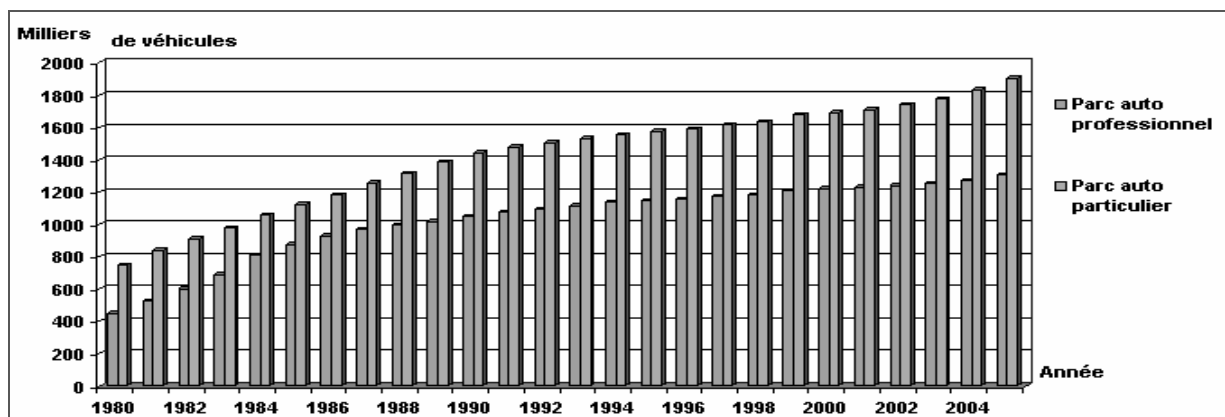


Figure I. 6 : Evolution du parc automobile total en Algérie [ONS]

D'après la *figure I.6*, on remarque que la taille du parc auto total est en progression continue d'année en année. Les figures qui suivent illustrent l'évolution du parc auto par type d'activité et par genre de carburant utilisé :

- ***L'Evolution du parc auto essence :***

Le graphique (*figure I.7*) suivant représente l'évolution du parc automobile des véhicules roulant à l'essence

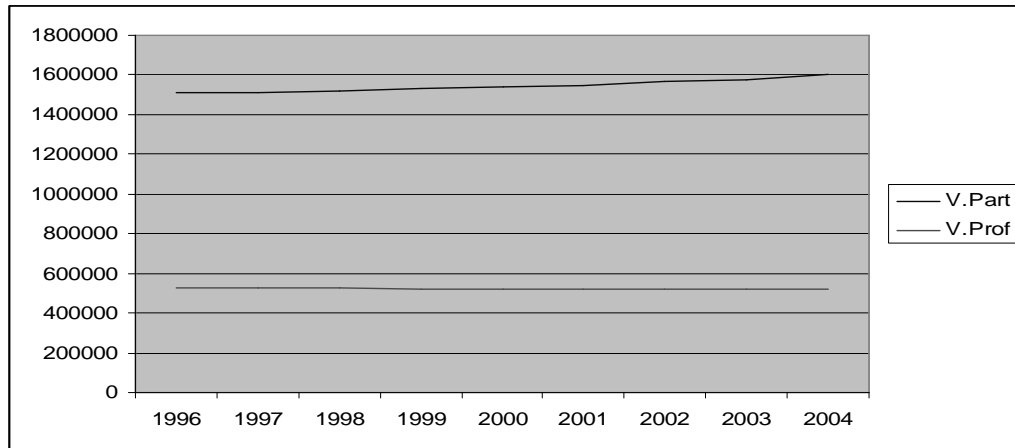


Figure I. 7 : Evolution du parc auto essence en Algérie.
[ONS]

Nous remarquons que :

- Le parc de véhicules professionnels reste constant, ceci s'explique par le fait que les professionnels préfèrent le gasoil pour des raisons économiques (un prix relativement faible) et techniques (la plupart des véhicules industriels et de transport roulent au diesel).
- Pour ce qui est des véhicules particuliers, le parc auto augmente légèrement.

Malgré l'augmentation de la taille du parc des véhicules à essence, la consommation baisse depuis 1991 et ce, en raison de l'augmentation des performances des moteurs qui consomment de moins en moins de carburant pour la même distance parcourue.

Remarque :

Il faut noter que les données, relatives à la taille du parc auto essence, incluent bien sûr les véhicules à essence, mais aussi les véhicules roulant en GPLc en raison de la bicarburation des moteurs.

- *Evolution du parc gasoil :*

La figure I.8 représente l'évolution du parc automobile des véhicules roulant au diesel :

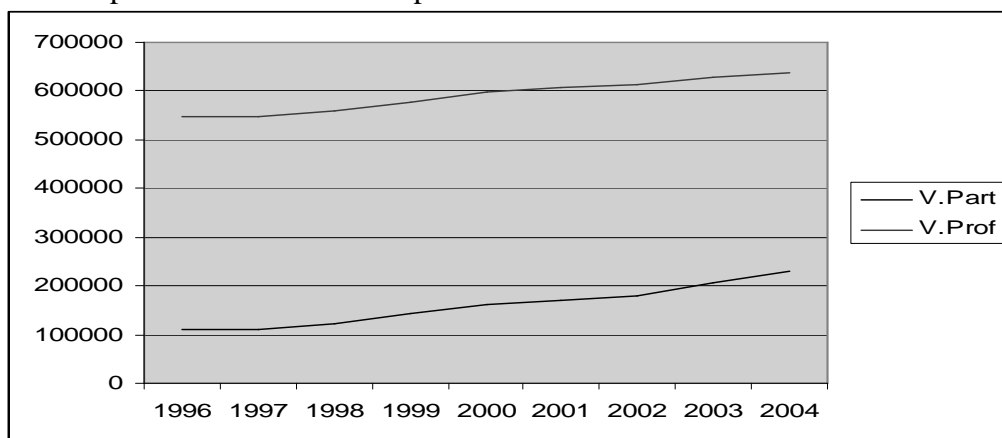


Figure I. 8 : Evolution du parc auto Diesel en Algérie.
[ONS]

L'examen du graphe fait apparaître une forte augmentation du marché diesel des voitures autant professionnelles que particulières. Ceci a fait que la consommation du gasoil routier a plus que doublé en Algérie. L'amorce et le développement de la consommation du diesel en Algérie sont dus essentiellement :

- A une importation de véhicules diésélisés très importante;
- Au prix du gasoil très faible par rapport à l'essence à cause des différentes taxations des deux carburants (13.7 DA pour le gasoil contre 23 DA pour l'essence super en 2006)
- A l'efficacité énergétique des moteurs diesel.

2. Etude de l'offre:

2.1 Etude de la production en Algérie:

A. La Production en Amont : Sonatrach et associés

L'activité en Amont a en charge la recherche, l'exploitation et la production des hydrocarbures. Ses missions sont principalement axées sur le développement des gisements découverts, l'amélioration du taux de récupération et la mise à jour des réserves.

Elle est conduite soit d'une manière autonome soit en association avec d'autres opérateurs pétroliers internationaux. Avec une production en association de plus de 232 millions de tep en 2005, la progression de la production se confirme (5% par rapport à l'exercice précédent) [SH 05]

L'activité Amont intègre dans sa stratégie opérationnelle les filiales qui lui sont rattachées telles que : [SH]

- ENAGEO (Entreprise Nationale de GEOphysique) ;
- GCB (Société Nationale de Génie Civil et Bâtiment) ;
- ENSP (Entreprise Nationale des Services aux Puits) ;
- ENTTP (Entreprise Nationale des Travaux aux Puits) ;
- ENAFOR (Entreprise Nationale de FORage) ;
- et ENGTP (Entreprise Nationale de Grands Travaux aux Puits).

Ces filiales participent au développement et à l'exploitation des hydrocarbures

B. La Production en aval : Naftec

La production aval est assurée par *Naftec*.

Naftec est une filiale à 100% du Groupe Sonatrach, à travers son Holding RCH.

Sa mission principale est le raffinage du pétrole et la production des :

- a. Carburants;
- b. Aromatiques;
- c. Lubrifiants;
- d. Bitumes.

La société dispose d'un siège situé à Alger, de structures fonctionnelles et de raffineries. Un tiers de sa production assure la couverture des besoins nationaux en produits raffinés

• **Les principales raffineries de Naftec :**

L'entreprise Naftec possède cinq raffineries réparties à travers le territoire national :

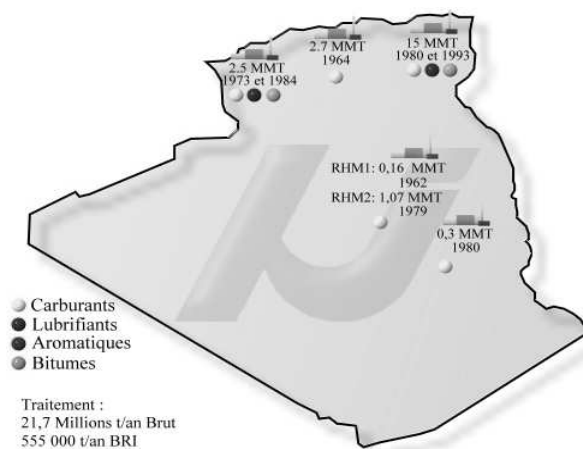


Figure I. 9 : La répartition des différentes raffineries en Algérie [Naftec]

<p><i>Raffinerie d'In Amenas (RIA)</i></p>	<p><i>Raffinerie d'Arzew (RAIZ)</i></p>
<p>Mise en service en 1980, d'une capacité de 300.000 tonnes et produisant des essences, gasoil et kérosène, cette raffinerie est à l'arrêt depuis 1986 pour causes techniques.</p>	<p>Raffinerie plus complexe, spécialisée dans la production des lubrifiants et des bitumes, elle approvisionne également les régions ouest et sud-ouest en carburants et GPL.</p>
<p><i>Raffinerie de Hassi Messaoud (RHM)</i></p>	<p><i>Raffinerie de Skikda (RAIK)</i></p>
<p>Constituée de deux installations :</p> <ul style="list-style-type: none"> • Une ancienne raffinerie (RHM1) mise en service en 1960 et assurant une production de 120.000 tonnes/An de carburants et de butane. • Une nouvelle raffinerie (RHM2) d'une capacité de traitement de pétrole brut de 1.116.500 tonnes/An mise en service en 1979 et permettant le ravitaillement d'une grande partie du Sud en gasoil, kérosène et essences. 	<p>Mise en service en 1980, elle est à connotation plus pétrochimique et orientée vers l'exportation pour près de 80% de sa production.</p> <p>02 unités de traitement de pétrole brut d'une capacité de 7.500.000 tonnes/an chacune pour la production des GPL et des carburants et 01 unité de production de bitumes routiers et oxydés d'une capacité de 145.000 tonnes/an.</p>
<p><i>Raffinerie d'Alger (RAIG)</i></p>	
<p>Mise en service en 1964 avec une capacité de traitement de pétrole brut de 2 700 000 tonnes par an, elle est orientée vers la production de carburants et GPL et dessert la région centre.</p>	

- **La répartition de la production par raffinerie :**

En 2005, la production nationale s'est élevée à 18,25 millions de tonnes, cette production se répartit comme suit :

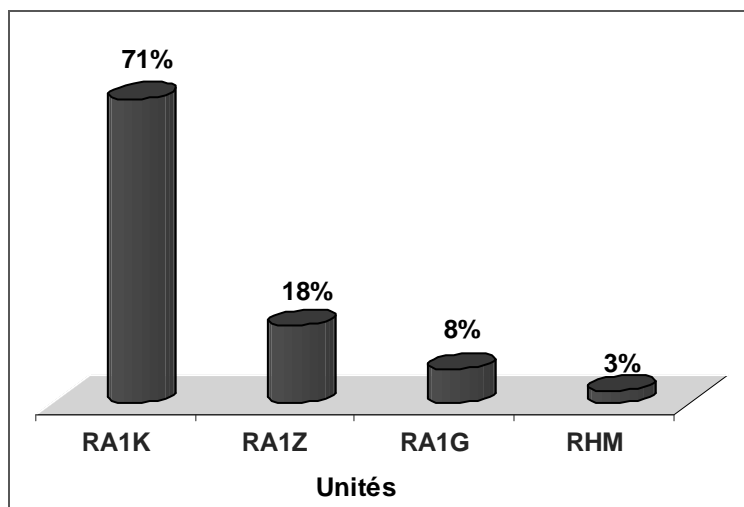


Figure I. 10 : Répartition de la production totale par raffinerie (2005) [SH 05]

D'après la *figure I.10*, nous remarquons que la raffinerie de Skikda produit à elle seule 71% des produits pétroliers.

- **La répartition par produit :**

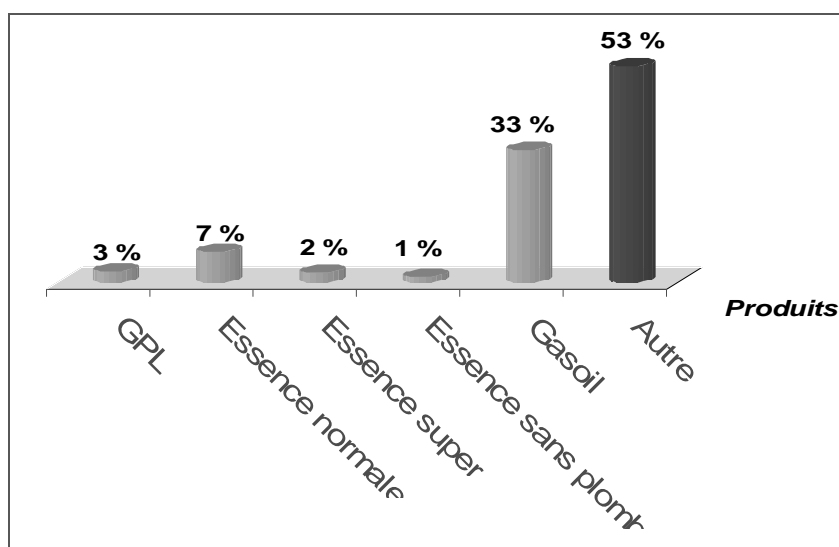


Figure I. 11 : Répartition de la production par produit (2005) [SH 05]

D'après la *figure I.11*, nous remarquons que le Gasoil représente 33% de la production des produits pétroliers raffinés alors que le GPL ne représente que 3%. En effet, la plus grande partie produite du GPL est issue au niveau des gisements.

2.2 Etude de la distribution et de la commercialisation des carburants en Algérie :

Société Nationale de commercialisation et de distribution de produits pétroliers, filiale 100 % de la Sonatrach, à travers son holding SVH, Naftal distribue et commercialise les produits pétroliers sur le marché National.

A. Produits distribués par Naftal :

Les produits distribués par Naftal sont classés en cinq catégories :

- les carburants;
- lubrifiants;
- bitumes;
- produits spéciaux;
- pneumatiques (importés).

La vente de carburants représente plus de 50 % du chiffre d'affaires de l'entreprise. Naftal assure la distribution de (06) types de carburants :

- L'essence Super ;
- L'essence Normale;
- Le Gasoil;
- Le GPLc;
- Le Kérosène;
- Le Fuel lourd.

Le carburant terre est en croissance de 5%. Ses ventes ont été dynamisées par la progression des ventes d'essence, notamment les ventes d'essence super et super sans plomb. L'essence normale est par contre en recul de 2%, suite à une forte diésélisation du parc automobile et à l'augmentation du parc équipé de pot catalytique, favorisant la consommation du super sans plomb. [Naftal]

B. Description de la chaîne de logistique des carburants en Algérie : [Naftal]

Sur le marché algérien, Naftal occupe une grande partie de l'activité de la logistique des carburants. Son objectif principal consiste à rendre disponible sur l'ensemble du territoire national les carburants récupérés auprès des raffineries. Naftal dispose d'un maillage du territoire, avec différentes « strates », qui peut être décrit ainsi :

- Des dépôts primaires directement alimentés par les raffineries, disposant d'une position centrale dans leur zone de couverture ;
- Des dépôts secondaires alimentés par les dépôts primaires dont ils forment un maillage de la zone de couverture ;
- Des stations service alimentées par les dépôts secondaires, qui constituent un maillage plus fin autour des dépôts primaires et secondaires ;

C. Le réseau de stations services de Naftal :

Naftal se charge surtout de ravitailler directement les stations service, et fait un appel important à la sous-traitance pour le transport terrestre. Le réseau national de distribution de carburants est constitué de plus 1 847 points de vente, dont 60% appartiennent à des opérateurs privés. [MEM]

Les stations services du réseau de distribution de Naftal peuvent être réparties en cinq catégories : [Naftal]

- Les stations du réseau propre, liées contractuellement à Naftal ;
- Les stations en gestion directe appartiennent à Naftal qui en assure elle-même la gestion;
- Les stations en gestion libre appartiennent à Naftal et leur gestion est confiée à un gérant.
- Les revendeurs ordinaires (ces stations pourraient être reclassées en stations privées)
- Les stations du réseau privé, sans lien avec Naftal mais qui s'approvisionnent auprès d'elle.

D. Le développement du GPL carburant : [Naftal]

Suite à une phase d'étude et d'expérimentation entamée en 1977, la décision d'introduire le GPL carburant est intervenue en 1983 avec l'adoption de la bicarburant et la mise en place de la réglementation liée aux conditions d'utilisation et de distribution du GPLc. La **Sonatrach** contribue à travers sa filiale **Naftal** à développer le GPL carburant en substitution des carburants traditionnels

Cette action a actuellement atteint sa maturité du fait de :

- La mise en place à travers tout le territoire d'installateurs de kits GPL sur véhicules ;
- La disponibilité d'un réseau important de stations service distribuant le GPLc, le nombre total de stations service est passé de 130 stations en 1996 à près de 300 stations service en 2003 ;
- L'existence d'un parc national de plus en plus grandissant converti au GPLc et qui s'élèverait aujourd'hui à 120 000 véhicules ;
- Le prix du GPLc à la pompe très attractif.

Il est important de préciser que :

- A l'horizon 2011, il est prévu l'installation de 8000 Kits GPL à raison de 2000 kits par an.
- Des subventions sont accordées aux installateurs par le biais du Fonds National de maîtrise de l'Energie (FNME), pour la conversion des véhicules essence à ce carburant alternatif. Cette aide est de 20 à 25 % pour chaque kit GPL, en fonction du prix moyen du kit GPL sur le marché national.
- Le coût du kit GPL varie entre 25000 DA et 100000 DA selon la nature de la carburant et l'ancienneté du véhicule mais aussi l'origine du kit;
- Le niveau de droits de douane du kit importé est de 5 % ;
- Le GPLc est disponible dans 343 stations services (2005) ;
- Pour renforcer :

- La lutte contre les faux installateurs,
- Le respect de la réglementation administrative et technique en vigueur et les normes de sécurité (soupape de sécurité) et des équipements
- La formation technique et la mise à niveau professionnelle ;

Les installateurs sont invités à souscrire à la charte qualité des professionnels et au guide technique afin de garantir : la sélection des meilleurs équipements sur le marché, le service offert et le service après vente et assurer les meilleures garanties de fiabilité et de longévité au moteur.

Conclusion :

L'analyse du marché de carburants a permis d'expliciter l'évolution de la consommation d'essence, de gasoil et du GPLc en Algérie, ainsi que les principales caractéristiques de l'offre et des circuits de distribution de ces produits.

Il ressort de cette première analyse que la consommation de GPLc est naturellement influencée par des **facteurs de demande** tels que le parc automobile, les prix des différents carburants, mais aussi par des **facteurs liés à l'offre** tels que les réseaux de distribution de GPLc, la disponibilité d'un réseau fiable d'installation et de maintenance de kits GPLc et les politiques de subventions et de promotions de ce carburant.

Cependant, notre approche dans ce travail s'intéresse plutôt aux facteurs de demande ; il s'agit d'identifier les variables pertinentes et apprécier leur pouvoir explicatif dans la consommation de GPLc.

L'intérêt de cette approche est donc d'apprécier le potentiel d'évolution de la consommation de GPLc par rapport aux tendances d'évolution des variables de demande. Une fois apprécié, ce potentiel pourrait servir à argumenter une politique d'investissement en réseaux de distribution GPLc (offre) ou de subventions en réponse à ce potentiel. On s'inscrit alors dans une logique tirée par la demande.

Chapitre II:
La modélisation de la consommation
du GPLc

Avant d'entrer dans le vif de notre sujet qui est celui de la modélisation de la consommation du GPLc, nous nous proposons d'exposer quelques principes de base qui régissent le processus de modélisation. Son aboutissement est en fait l'élaboration d'un modèle qui se définit comme la représentation simplifiée d'un phénomène en vue de le décrire, le comprendre, l'expliquer et enfin le prévoir.

Le modèle est donc l'outil incontournable utilisé lorsque nous voulons à appréhender un phénomène à travers ses multiples facettes. Dans cette démarche, des hypothèses sont émises et des relations sont explicitées avec pour finalité la représentation des traits les plus marquants d'une réalité que le modélisateur cherche à styliser. [FOU 04]

A titre illustratif, l'économétrie représente le principal outil d'analyse quantitative, utilisé par les économistes dans divers domaines d'application comme la macroéconomie ou la finance. Elle permet de vérifier l'existence de certaines relations entre des phénomènes économiques à travers divers modèles. Ces derniers consistent en une présentation formalisée de ces phénomènes sous forme d'équations, dont les variables sont des grandeurs économiques.

Dans ce cadre, **la régression multiple** représente la méthode la plus utilisée dans la modélisation des phénomènes économiques. Elle se définit comme étant une méthode statistique largement utilisée pour modéliser la relation existant entre la variable qu'on cherche à expliquer et les variables jugées susceptibles de l'expliquer (dites variables explicatives) [KAL 02]

Cette méthode de modélisation est employée, dans le cas où l'on s'intéresse directement à la quantification de l'effet de chacune des variables explicatives sur la variable qu'on cherche à expliquer, ou bien dans celui où l'on souhaite disposer d'un modèle en vue de la prédiction des valeurs de la variable dépendante à partir des valeurs des variables explicatives.

En d'autres termes, la régression multiple peut être utilisée aux fins de trouver la meilleure équation linéaire de prévision (modèle) et en évaluer la précision et la signification, estimer la contribution relative de deux ou plusieurs variables explicatives sur la variation d'une variable à expliquer et enfin juger de l'importance relative de plusieurs variables explicatives sur une variable dépendante [JON 99]

La régression multiple s'impose donc comme une méthode très efficace pour la modélisation des phénomènes économiques. A ce titre, plusieurs travaux, basés sur cette méthode, appliquée à différents domaines, dont celui de l'énergie, ont été menés à travers le monde. Parmi ces travaux, on peut citer ceux de F. LANTZ [LAN 03] qui s'est intéressé à la modélisation économétrique des marchés du pétrole et des produits pétroliers en France. Il s'est aussi intéressé à la modélisation basée sur une régression multiple de la demande d'énergie dans le monde [LAN 05].

On peut également évoquer le travail de Alves.D. Bueno [ALV 03] qui s'est intéressé à la modélisation de la demande de gaz au Brésil, à court et moyen terme. Il a aussi basé son modèle sur la régression multiple.

Ayant introduit brièvement le processus de modélisation et mis en relief la méthode de la régression multiple, comme outil privilégié de modélisation des phénomènes économiques, nous nous proposons d'appliquer cette méthode en vue d'atteindre l'objectif du présent travail qui vise dans un premier temps, la modélisation de la consommation nationale du GPLc. A cet effet, notre travail sera scindé en deux sections : la première consacrée à la présentation succincte des principes de base de la modélisation, fondée sur la régression multiple et la seconde, dédiée à la modélisation de la consommation du GPLc.

Section I : Outil de modélisation: la régression multiple

I. Présentation du modèle : [BOU 02]

Un modèle de régression multiple s'écrit de la manière suivante :

$$y_t = a_0 + a_1x_{1t} + a_2x_{2t} + \dots + a_kx_{kt} + \varepsilon_t; \quad t = \overline{1, n}$$

Avec :

y_t : La variable à expliquer à la date t ,

x_{1t} : La variable explicative 1 à la date t ,

x_{2t} : La variable explicative 2 à la date t ,

⋮

x_{kt} : La variable explicative k à la date t ,

$a_0, a_2, a_3, \dots, a_k$: Les paramètres du modèle

ε_t = erreur de spécification (différence entre le modèle vrai et le modèle spécifié)

n = nombre d'observations

L'écriture précédente du modèle est d'un maniement peu pratique. Afin d'en alléger l'écriture et de faciliter l'expression de certains résultats, on a recours aux notations matricielles.

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}; \quad a = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix}; \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}; \quad X = \begin{pmatrix} 1 & x_{11} & x_{21} & \dots & x_{k1} \\ 1 & x_{12} & \dots & \dots & x_{k2} \\ \vdots & & & & \vdots \\ 1 & x_{1n} & x_{2n} & \dots & x_{kn} \end{pmatrix}$$

Nous avons $Y = Xa + \varepsilon$

II. Rôle du terme aléatoire :

Le terme ε_t mesure la différence entre les valeurs réellement observées Y_t et les valeurs qui auraient été observées si la relation spécifiée avait été rigoureusement exacte.

Le terme regroupe donc trois erreurs :

- Une erreur de spécification, c'est-à-dire le fait que la seule variable explicative n'est pas suffisante pour rendre compte de la totalité du phénomène expliqué ;
- Une erreur de fluctuation d'échantillonnage : les observations changent d'un échantillon à l'autre et donc les estimations sont légèrement différentes ;
- Une erreur de mesure : les données ne représentent pas exactement le phénomène.

Le modèle tel qu'il vient d'être spécifié n'est qu'une caricature de la réalité. En effet, ne retenir que k variables explicatives pour expliquer la variable exogène est à l'évidence insuffisant puisqu'il existe une multitude d'autres facteurs susceptibles de l'expliquer. C'est pourquoi, nous ajoutons un terme ε_t qui synthétise l'ensemble de ces informations non explicitées dans le modèle.

III. Estimation des paramètres du modèle :

1. Hypothèses : [BOU 02]

H1 : le modèle est linéaire en x_{it} (ou n'importe quelle transformation de x_{it})

H2 : les valeurs x_{it} sont observées sans erreurs (x_{it} non aléatoire) ;

H3 : $E(\varepsilon_t) = 0$: l'espérance mathématique de l'erreur est nulle : en moyenne le modèle est bien spécifié et donc l'erreur moyenne est nulle ;

H4 : $E(\varepsilon_t^2) = a$: la variance de l'erreur est constante : le risque d'amplitude de l'erreur est le même quelle que soit la période (hypothèse d'homoscédasticité) ;

H5 : $E(\varepsilon_t, \varepsilon_s) = 0$ si $t \neq s$: les erreurs sont non corrélées (ou encore indépendantes) : une erreur à l'instant 't' n'a pas d'influence sur les erreurs suivantes (absence d'autocorrélation) ;

H6 : $Cov(X_t, \varepsilon_t) = 0$: l'erreur est indépendante de la variable explicative.

2. Formulation des estimateurs: [BOU 02] [AIK 03]

En traçant un graphique des couples de données, liant la variable à expliquer et les variables explicatives, nous obtenons un nuage de points que nous pouvons ajuster à l'aide d'une droite.

L'estimateur des coefficients a_i est obtenu en minimisant la distance au carré entre chaque observation et la droite, d'où le nom d'estimateur des Moindres Carrés Ordinaires (MCO).

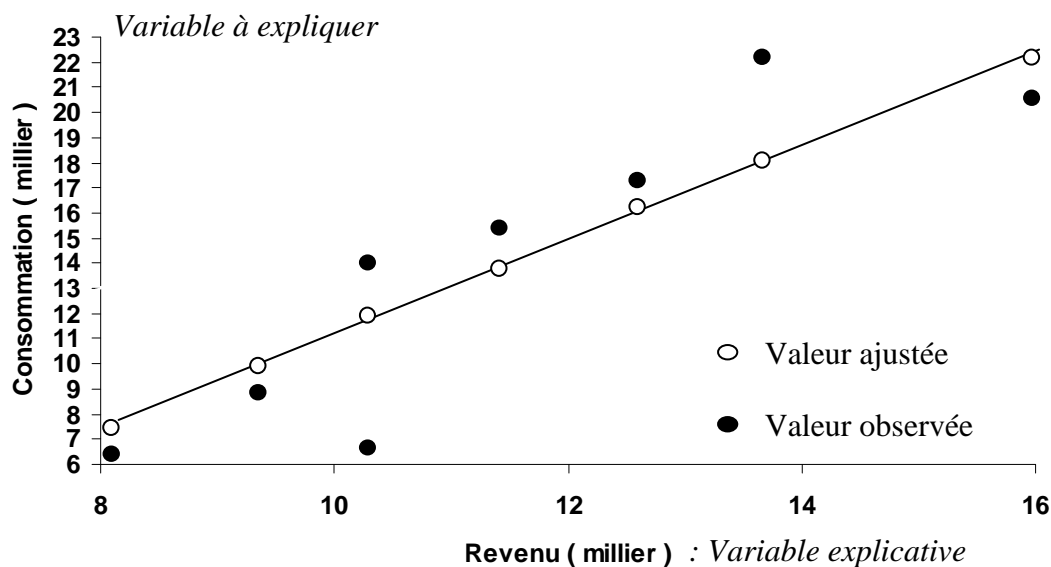


Figure II. 1 : Ajustement d'un nuage de points par une droite

La résolution analytique est la suivante : $Min \sum_{t=1}^T \varepsilon_t^2$

On a : $Min \sum_t \varepsilon_t^2 = Min \varepsilon' \varepsilon = Min (Y - X\hat{a})'(Y - X\hat{a}) = Min S$

Avec : $\varepsilon' = \text{transposé du vecteur } \varepsilon$

Pour minimiser cette fonction par rapport à a , nous différencions S par rapport à a , ce qui donne : $-2X'Y + 2X'X\hat{a} = 0 \rightarrow \hat{a} = (X'X)^{-1} X'Y$

Cette solution est réalisable, si la matrice carrée $X'X$ de dimension $(k+1, k+1)$ est inversible. Le modèle estimé s'écrit :

$$y_t = \hat{a}_0 + \hat{a}_1 x_{1t} + \hat{a}_2 x_{2t} + \dots + \hat{a}_k x_{kt} + e_t$$

Avec $e_t = Y_t - \hat{Y}_t$ Où e_t représente l'écart entre la valeur observée de la variable à expliquer et sa valeur estimée (ajustée)

IV. Equation d'analyse de la variance :

L'équation fondamentale d'analyse de la variance s'écrit : [BOU 02]

$$\sum_t (Y_t - \bar{Y})^2 = \sum_t (\hat{Y}_t - \bar{Y})^2 + \sum_t e_t^2$$

$$SCT = SCE + SCR$$

Avec :

SCT : la Somme des Carrés Totale (la variabilité totale).

SCE : la Somme des Carrés Expliquée (la variabilité expliquée)

SCR : la Somme des Carrés Résiduelle (la variabilité des résidus)

Cette équation va nous permettre de juger de la qualité de l'ajustement d'un modèle.

En effet, plus la variance expliquée est proche de la variance totale, meilleur est l'ajustement du nuage de points par la droite des moindres carrés.

À cet effet, Il est important de calculer :

1. Le coefficient de détermination R^2 : [BOU 06]

Il indique la qualité de l'ajustement. En effet, le coefficient de détermination R^2 reflète l'écart entre les réalisations et les estimations du modèle, ce qui permet d'évaluer l'efficacité du modèle c'est à dire la possibilité de représenter le phénomène à une erreur moyenne égale à $(1-R^2)$

$$R^2 = \frac{\sum_t (\hat{y}_t - \bar{y})^2}{\sum_t (y_t - \bar{y})^2} = 1 - \frac{\sum_t e_t^2}{\sum_t (y_t - \bar{y})^2} = \frac{SCE}{SCT}$$

R^2 = La variance expliquée par la régression / variance de Y .

- Une valeur de R^2 proche de 1, montre que l'ajustement est bon.
- Une valeur de R^2 proche de 0, montre que l'ajustement est mauvais.
- Une valeur de $R^2 = 1$ montre que l'ajustement est parfait.

2. Le coefficient de corrélation ρ : [BOU 06]

ρ est le coefficient de corrélation entre la série Y_t et \hat{Y}_t . Il représente la valeur maximale du coefficient de corrélation linéaire simple entre les coordonnées de Y et celles de tout vecteur

de la forme $X \hat{a}$. Il est donné par :
$$\rho = \frac{Cov(Y, \hat{Y}_t)}{\sigma_Y \sigma_{\hat{Y}_t}}$$

V. Test de spécification de l'ajustement :

La détermination des paramètres a_i , R^2 , ρ n'est pas suffisante pour affirmer que la liaison entre X et Y est significative. En effet il est indispensable de vérifier:

- La significativité globale du modèle (le F- Test);
- La significativité des variables explicatives du modèle (le T- test);
- La blancheur et la normalité des résidus ;
- Et la stabilité du modèle.

Pour ce faire, nous allons présenter les différents tests d'hypothèses qui permettent de vérifier ceci :

1. Test global de Fisher : (le F- Test) [BOU 02]

Ce test permet de voir la relation entre Y et les variables $x_1, x_2, x_3, \dots, x_k$ dans leur ensemble, c'est-à-dire si l'ensemble des variables explicatives a une influence sur la variable à expliquer. Ce test peut être formulé de la manière suivante : existe-t-il au moins une variable explicative significative ? Soit le test d'hypothèse :

$H_0 : a_0 = a_1 = a_2 = \dots = a_k = 0$

$H_1 : \text{il existe au moins un des coefficients qui est non nul.}$

La statistique utilisée est :
$$F^* = \frac{SCE/k}{SCR/(n-k-1)} = \frac{R^2/k}{(1-R^2)/(n-k-1)} \text{ pour } i = \overline{1, k}$$

- Si : $F^* \geq F_{(1, n-k-1)}^\alpha$, alors l'hypothèse H_0 est rejetée, le modèle est globalement explicatif.
- Sinon l'hypothèse H_0 est acceptée, ce qui signifie qu'il n'existe aucune relation linéaire significative entre la variable à expliquer et les variables explicatives.

2. Test de significativité des variables : (le T- Test) [BOU 02]

Pour une droite de régression : $y_t = a_0 + a_1 x_{1t} + a_2 x_{2t} + \dots + a_k x_{kt} + \varepsilon_t ; t = \overline{1, n}$

Afin de vérifier la significativité des variables x_{it} , c'est-à-dire si ces variables expliquent significativement le modèle, on propose le test d'hypothèse suivant:

Soit A , le vecteur de composantes $(a_0, a_1, a_2, \dots, a_k)$.

Les hypothèses sont formulées par :

$H_0 : a_i = 0$ (la variable x_{it} n'explique pas le modèle)

$H_1 : a_i \neq 0$ (la variable x_{it} explique le modèle).

La statistique utilisée est celle de Student.

Pour un seuil ($\alpha = 0.05$), la règle de décision est alors la suivante :

- Si $t_{\hat{\alpha}_i}^* = \frac{\hat{\alpha}_i}{\sigma_{\hat{\alpha}_i}} > t_{n-2}^{\alpha/2}$, alors l'hypothèse H_0 est rejetée et la variable x_{it} est alors contributive à l'explication de la variable y_t .
- Sinon l'hypothèse H_0 est acceptée et la variable x_{it} n'est pas contributive à l'explication de la variable y_t .

3. Test sur les résidus

L'analyse des résidus représente une phase très importante pour la validation du modèle.

Il s'agira avant tout, de vérifier que les résidus forment bien un bruit blanc, c'est-à-dire que :

- Le risque d'amplitude de l'erreur est le même quelle que soit la période.
- Les erreurs sont non corrélées (ou encore indépendantes) : une erreur à l'instant 't' n'a pas d'influence sur les erreurs suivantes.

Par la suite, il s'agira de vérifier que les résidus suivent une loi normale. En effet, pour pouvoir calculer les intervalles de confiance prévisionnels et aussi pour pouvoir effectuer les tests de Student, les résidus doivent suivre une loi normale.

3.1 Test de blancheur des résidus :

Un bruit blanc ($\varepsilon_t, t \in T$) est une suite de variables aléatoires non corrélées de moyenne nulle et de variance finie constante. En effet, un processus bruit blanc vérifie les propriétés suivantes : [UPM 04]

$$\begin{cases} E(\varepsilon_t) = 0 \\ Var(\varepsilon_t) = E(\varepsilon_t^2) = \sigma_\varepsilon^2 \end{cases} ; \forall t \in T$$

Par conséquent, sa fonction d'auto covariance est donnée par :

$$\gamma(h) = Cov(X_t, X_{t+h}) = E(\varepsilon_t, \varepsilon_{t+h}) = \begin{cases} 0 & h \neq 0 \\ \sigma_\varepsilon^2 & h = 0 \end{cases}$$

Si de plus : $\varepsilon_t \rightarrow N(0, \sigma^2)$ alors ε_t est un bruit blanc gaussien.

Il existe plusieurs tests qui permettent de vérifier la blancheur des résidus, parmi lesquels on a :

A. Test de Durbin- Waton: [BOU 06]

Le test de Durbin- Watson permet de détecter une autocorrélation des résidus. Le test d'hypothèse s'écrit comme suit :

$H_0 : \rho = 0$ (absence de corrélation).

$H_1 : \rho \neq 0$ (présence de corrélation).

La statistique de Durbin- Watson est donnée par: $DW = \frac{\sum_{t=2}^n (\hat{\epsilon}_t - \hat{\epsilon}_{t-1})^2}{\sum_{t=1}^n \hat{\epsilon}_t^2}$ avec $0 < DW < 4$

Afin de tester l'hypothèse H_0 , Durbin- Watson ont tabulé les valeurs critiques de DW , au seuil 5%, en fonction de la taille de l'échantillon (n) et le nombre de variables explicatives (k).

La lecture de la table permet de déterminer deux valeurs critiques : d_1 et d_2 comprises entre 0 et 2.

Selon la position du DW empirique, entre ces deux valeurs, nous pouvons conclure que si :

Tableau II. 1: Test de Durbin Watson

$0 < DW < d_1$	H_0 EST REJETEE	$\rho \geq 0$
$d_1 < DW < d_2$	<i>On ne peut pas conclure</i>	
$d_1 < DW < 4 - d_2$	H_0 est acceptée	$\rho = 0$
$4 - d_2 < DW < 4 - d_1$	<i>On ne peut pas conclure</i>	
$4 - d_1 < DW < 4$	H_0 est rejetée	$\rho < 0$

B. Tests de portemanteau : [BOU 02] [JON 99]

Outre le test de Durbin Waston, il existe d'autres tests qui permettent d'identifier les processus de bruit blanc. Nous devons donc identifier $Cov(Y_t, Y_{t-k}) = 0$ ou encore $\rho_k = 0, \forall k$. En effet, un processus de bruit blanc implique que $\rho_1 = \rho_2 = \dots = \rho_h = 0$

Les statistiques Q et Q' de Box- Pierce et de Ljung- Box permettent de tester cette hypothèse.

Soit les hypothèses :

$H_0 : \rho_1 = \rho_2 = \dots = \rho_h = 0$ (absence d'autocorrélation)

$H_1 : \text{il existe } j \text{ tel que } : \rho_j \neq 0$ (présence d'autocorrélation).

➤ **Test de Box Pierce (1970) :**

Pour effectuer ce test, on a recours à la statistique de Quenouille Q (due à Box Pierce) qui

est donnée par: $Q = T \sum_{h=1}^H \hat{\rho}_h^2(\hat{\epsilon}_t)$

Avec :

H : nombre de retards.

ρ_k : autocorrélation empirique d'ordre k .

T : nombre d'observations.

La statistique Q est distribuée de manière asymptotique comme un χ^2 à h degré de liberté
 Les règles du test sont :

- Si $Q \leq \chi^2(h)$ au seuil $(1 - \alpha)$, on accepte H_0 , alors les résidus forment un bruit blanc et donc le modèle est valide.
- Sinon on rejette H_0 , et les résidus ne forment pas un bruit blanc et donc le modèle n'est pas valide

➤ **Test de Ljung-Box :**

Ce test est préféré au test de Box- Pierce, lorsque l'échantillon est de petite taille. La distribution de la statistique du test de Ljung- Box, notée Q' est en effet plus proche de celle de Khi deux χ^2 en petit échantillon que ne l'est celle du test de Box- pierce.

Dérivée de la première statistique Q , la statistique Q' de Ljung- Box s'écrit :

$$Q^* = T(T + 2) \sum_{h=1}^H \frac{\hat{\rho}_h^2(\hat{\epsilon}_t)}{T - h}$$

Avec :

T : taille de l'échantillon.

H : le nombre de retard

$\hat{\rho}_k$: Autocorrélation empirique d'ordre k .

Cette statistique est aussi distribuée selon un χ^2 à h degré de liberté et dont les règles de décisions sont identiques au précédent.

Ces tests sont appelés par les anglo-saxons « portemanteau test » soit littéralement test : « fourre- tout »

3.2 Test de normalité des résidus : [BOU 02] [BOU 98]

Le test de Jarque & Bera (1984) peut s'appliquer pour tester la normalité des résidus.

Ce dernier est fondé sur la notion Skewness (moment d'ordre 3, l'asymétrie de la distribution) et sur la notion de Kurtosis (moment d'ordre 4, l'aplatissement – épaisseur des queues de distribution).

Soit le moment empirique d'ordre k du processus : $\mu_k = E[\hat{\epsilon}_t - E(\hat{\epsilon}_t)]^k = \frac{1}{n} \sum_{t=1}^n (\hat{\epsilon}_t - \bar{\epsilon}_t)^k$

Où n : nombre d'observations,

a. Test de Skewness : [BOU 02]

Le Skewness est une mesure de l'asymétrie de la distribution de la série autour de sa moyenne. Le coefficient du Skewness est défini par : $(S_k)^{1/2} = \frac{\mu_3}{\mu_2^{3/2}}$

Lorsque le nombre d'observations est grand ($n \geq 30$), c'est-à-dire :

$$n \rightarrow \infty : (S_k)^{1/2} \sim N(0, \sqrt{\frac{6}{n}})$$

Remarque :

- Le Skewness d'une distribution symétrique est nulle telle que la distribution normale.
- Le Skewness positif signifie que la distribution a une queue allongée vers la droite.
- Le Skewness négatif signifie que la distribution a une queue allongée vers la gauche.

b. Test de Kurtosis : [BOU 02]

Le Kurtosis mesure le caractère pointu ou plat de la distribution de la série. Le coefficient du

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

Lorsque le nombre d'observations est grand, c'est-à-dire ($n \geq 30$) :

$$n \rightarrow \infty : K_u \sim N\left(3, \sqrt{\frac{24}{n}}\right)$$

Remarque :

- Le Kurtosis de la distribution normale est 3.
- Si le Kurtosis est supérieur à 3, la distribution est plutôt pointue relativement à la normale.
- Si le Kurtosis est inférieur à 3, la distribution est plutôt aplatie relativement à la normale

On construit alors les statistiques centrées réduites correspondantes que l'on compare à 1.96 qui représente la valeur de la loi normale au seuil de 5%

$$v_1 = \frac{|S_k^{1/2}|}{\sqrt{\frac{6}{N}}} \sim N(0,1) ; v_2 = \frac{|k_u - 3|}{\sqrt{\frac{24}{N}}} \sim N(0,1)$$

- Si $v_1 \leq 1.96$ et $v_2 \leq 1.96$, on accepte l'hypothèse de symétrie et l'hypothèse de queue de distributions plates et les résidus sont donc gaussiens.
- Dans le cas contraire, l'hypothèse de normalité est rejetée.

c. Test de Jarque et Bera : [UPM 04]

Le test de Jarque-Bera est une statistique, qui permet d'examiner si la série est normalement distribuée. La statistique mesure la différence du Skewness et du Kurtosis de la série avec ceux de la distribution normale. La statistique est calculée comme suit :

$$JB = \frac{N}{6} S_k + \frac{N}{24} (k_u - 3)^2 \sim \chi^2(2) \quad \text{Où } S_k: \text{ le Skewness ; } k_u : \text{ le Kurtosis.}$$

Sous l'hypothèse nulle d'une distribution normale, la statistique de Jarque-Bera suit asymptotiquement une loi de χ^2 ($1 - \alpha$) avec 2 degrés de liberté ($\alpha = 5\%$)

Si : $JB > \chi_{1-\alpha}^2(2)$ alors on rejette l'hypothèse de normalité des résidus au seuil $\alpha = 5\%$.

4. Test de stabilité du modèle : Le test de CUSUM : [BOU 02]

Le test de CUSUM (CUMulative SUM) est fondé sur la dynamique de l'erreur de prévision. Il permet de détecter les instabilités structurelles des équations de régression au cours du temps.

L'idée générale de ce test est d'étudier l'évolution au cours du temps de l'erreur de prévision normalisée. Il est fondé sur la somme cumulée des résidus récursifs. On appelle résidus récursifs cette succession d'erreur de prévision calculée en $t-1$ pour t .

Le résidu récursif w_t au temps t est donc défini comme le rapport entre :

- l'écart entre la prévision calculée en $t-1$ pour t et la réalisation en t ,
- et une variable d'échelle, calculée pour cette même période t ,

Soit :

$$w_t = \frac{e_t = y_t - \hat{a}_t}{S_e} = \frac{y_t - x'_t \hat{a}_{t-1}}{\sqrt{(1 + x'_t (X'_{t-1} X_{t-1})^{-1} x_t)}}$$

Avec : $t = K+2, K+3, K+4, \dots, n$ ($K = k+1$ est le nombre total de paramètres estimés du modèle). Ce résidu récursif suit donc une loi normale $N(0, \sigma^2)$.

A partir du résidu récursif w_t , on calcule la statistique CUSUM : $w_t = \frac{n-K}{SCR} \sum_{j=K+2}^t w_j$

Avec : $t = K+2, \dots, n$ ($K = k+1$ est le nombre total de paramètres estimés du modèle) et SCR la somme des carrés des résidus, calculée à partir du modèle à n observations.

Si les coefficients du modèle sont stables au cours du temps, alors les résidus récursifs doivent rester dans l'intervalle défini par les deux droites : $[K, \pm \alpha \sqrt{n-K}]$ et $[n, \pm 3\alpha \sqrt{n-K}]$

Avec $\alpha = 1,143 ; 0,948$ et $0,850$ respectivement pour des seuils de confiance 1%, 5% et 10%.

Dans le cas contraire, le modèle est instable.

VI. La prévision : [BOU 02]

Lorsque les coefficients du modèle ont été estimés, il est possible de calculer une prévision à un horizon h . Le problème consiste à déterminer la valeur qui doit être attribuée à la variable endogène (variable à expliquer) lorsque nous connaissons les valeurs des variables exogènes (variables explicatives).

Le modèle général estimé est le suivant : $Y_t = \hat{a}_0 + \hat{a}_1 x_{1t} + \hat{a}_2 x_{2t} + \dots + \hat{a}_k x_{kt} + e_t$

La prévision pour la donnée $t+h$ est la suivante :

$$\hat{Y}_{t+h} = \hat{a}_0 + \hat{a}_1 x_{1t+h} + \hat{a}_2 x_{2t+h} + \hat{a}_3 x_{3t+h} + \dots + \hat{a}_k x_{kt+h}$$

L'erreur de prévision est donnée par : $e_{t+h} = Y_{t+h} - \hat{Y}_{t+h}$

Après avoir présenté les règles principales d'estimation des paramètres d'un modèle de régression à partir des différentes hypothèses ainsi que les règles de sa validation (T-test, F-test...), nous allons procéder à la résolution du problème de « la modélisation de la consommation nationale du GPLc » sur la base des règles énoncées.

Section II : Modélisation de la consommation du GPLc par la régression multiple

Dans cette partie, nous proposons une approche de modélisation de la consommation du GPLc, basée sur la régression économétrique, telle que présentée précédemment. Ce type d'approche va nous permettre de tester la pertinence des variables jugées à priori explicatives dans la consommation de ce carburant.

L'identification de ces variables repose sur une segmentation et une analyse de la demande des carburants en Algérie afin de mieux comprendre sa structure. Cette segmentation vient en complément à l'analyse du marché effectuée dans le premier chapitre.

Notre travail va donc concerner, tour à tour, la segmentation de la demande du GPLc, la définition des variables susceptibles d'expliquer la consommation du GPLc et enfin l'estimation de la fonction de la demande par le biais d'une régression multiple à plusieurs variables.

I. Segmentation de la demande du carburant :

Comme pour toute étude de modélisation de la fonction « Demande », nous commençons par la segmentation de cette dernière pour mieux comprendre sa structure et choisir les meilleurs facteurs explicatifs. La segmentation se définit comme étant une technique de découpage du marché consistant à isoler des groupes de consommateurs ayant des comportements ou des habitudes d'achat homogènes. La segmentation permet d'adapter les actions marketing à chacun des groupes ainsi constitués [KAL 02]

La segmentation de la demande des carburants se fait selon les deux critères suivants:

- ❖ Le type d'utilisation
- ❖ Le type de clients.

1. Segmentation selon le type d'utilisation :

Le demande des carburants peut être décomposée sur la base du type d'utilisation en deux segments bien différents : la demande routière et la demande non routière.

- **La demande routière :**

Où les carburants sont utilisés pour la consommation énergétique des véhicules. On peut distinguer trois sous segments différents:

- **Les véhicules particuliers**
- Le transport en commun
- Les véhicules de transport de marchandises.

- **La demande non routière :**

Où les carburants ne sont pas utilisés pour la consommation énergétique des véhicules. C'est la demande qui utilise le carburant hors route. On peut distinguer également trois sous segments différents :

- Dans l'industrie où le carburant est utilisé comme un additif ;
- Dans l'agriculture, la construction de bâtiments, les usines.

On peut schématiser cette segmentation comme suit :

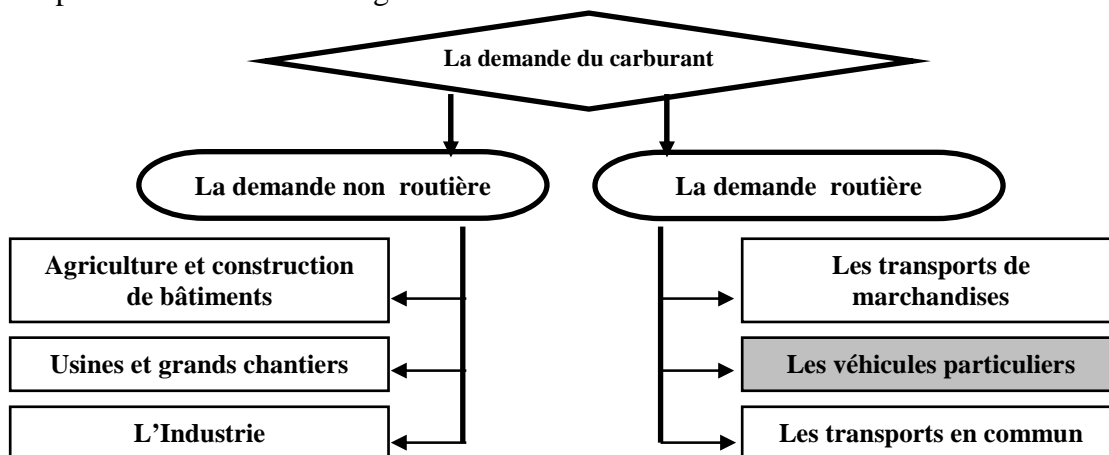


Figure II. 2 : Segmentation selon le type d'utilisation

2. Segmentation selon le type de clients :

Une deuxième segmentation peut se faire sur la base du type de client (consommateur). Il existe deux types de consommateurs : les particuliers et les professionnels sur lesquels on distingue deux segments de la demande :

- *La demande des particuliers :*

Représentée par les automobilistes qui possèdent en général des véhicules de petite taille (véhicules de tourisme).

- *La demande des professionnels :*

Elle intègre également les activités de :

- Transport de voyageurs ;
- Véhicules lourds (Transport de marchandises) ;
- Grands chantiers, Agriculture, Industrie.

On peut schématiser cette segmentation comme suit:

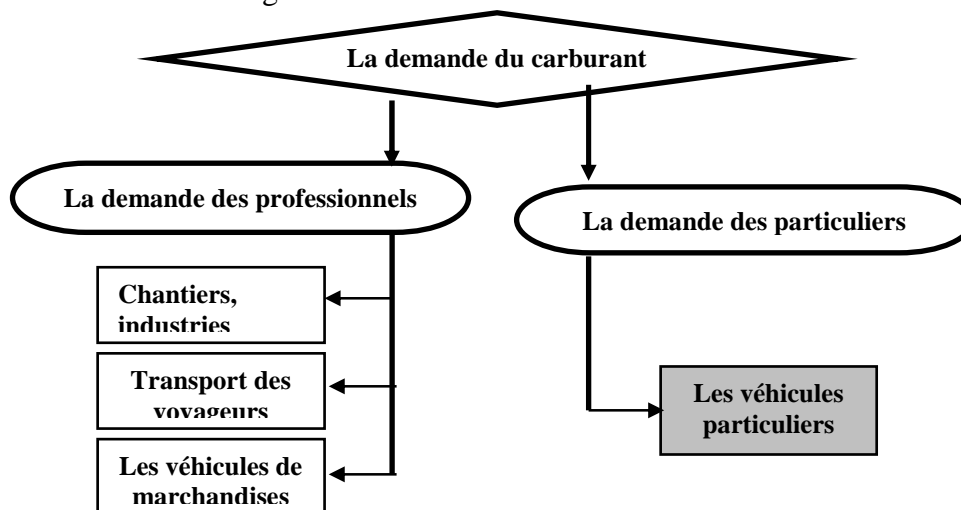


Figure II. 3 : Segmentation selon le type de clients

Notre étude va cibler essentiellement la demande des véhicules particuliers (légers), afin d'en évaluer le potentiel de conversion au GPLc. Ce choix est justifié par la différence de comportement entre les deux types de consommateurs (particuliers et professionnels) par rapport au prix de carburants. En effet, la demande des particuliers est plus sensible au changement de prix que la demande professionnelle, car pour les particuliers la consommation du carburant est un besoin final. Cependant pour les professionnels la consommation de carburants n'est qu'une charge d'exploitation et toute augmentation des prix des carburants est reportée sur le prix du produit ou du service final.

Après avoir précisé le segment qui délimite notre travail, il convient de déterminer les variables susceptibles d'expliquer la consommation du GPLc par le type de consommateur retenu. Ces variables doivent être groupées en familles de variables ayant un impact sur cette consommation tels que : le prix du carburant, la croissance économique, la démographie, les revenus de la population, les déplacements de la population, etc...

II. Les variables explicatives de la consommation du GPLc :

Comme l'objectif est de réaliser un modèle explicatif de la consommation du GPLc, en considérant les différents paramètres qui sont susceptibles de l'influencer, et d'après la segmentation et l'étude de marché des carburants faites nous avons été amenés à sélectionner plusieurs facteurs qui pourraient avoir une influence sur la demande de ce carburant. Les variables retenues sont citées ci-dessous :

1. Le prix : (dinars par litre)

On peut supposer que le prix est le premier facteur ayant un impact direct sur la consommation. Car on sait que plus un produit est coûteux, moins sa consommation devient accessible. Nous distinguons :

Le prix du GPLc : (P_{GPLc})

L'augmentation ou la réduction du prix du GPLc pourrait provoquer respectivement une diminution ou une augmentation de la consommation de ce carburant. En effet, face à l'augmentation du prix, les consommateurs auraient tendance à choisir entre la réduction de leurs déplacements journaliers ou l'utilisation d'autres carburants de substitution, ce qui impliquerait dans les deux cas une diminution de la consommation du GPLc.

Le prix du gasoil : (P_{gas})

L'augmentation ou la réduction du prix du Gasoil pourrait provoquer une augmentation ou une diminution (respectivement) de la consommation du GPLc.

En effet, face à l'augmentation du prix du gasoil, les usagers auront tendance à rechercher des carburants plus attractifs en matière de coûts, entre autres, le GPLc.

Le prix de l'essence (normale) : (P_{ess})

L'augmentation ou la réduction du prix de l'essence provoquerait une croissance ou une diminution (respectivement) de la consommation du GPLc. En effet, face à l'augmentation du prix de l'essence, les usagers auront tendance à rechercher le carburant le plus attractif sur le plan coût, en l'occurrence, le GPLc.

2. La taille du parc automobile : (Parc)

La consommation du carburant varie en proportion directe avec la taille du parc automobile. En effet, la taille du parc automobile implique que le potentiel de conversion est d'autant plus important que le parc est important, donc l'augmentation de la taille du parc conduirait à une augmentation de la consommation.

3. Le taux de croissance de la taille du parc automobile : (T_{cr})

On peut supposer aussi que la consommation du carburant varie proportionnellement avec le taux de croissance du parc automobile. En effet, le taux de croissance du parc implique que les nouveaux véhicules seraient plus susceptibles d'être convertis en GPLc que les anciens.

4. La population : (en millions) (POP)

La taille de la population a un effet positif sur la demande du carburant. Il est logique de supposer que l'augmentation de ce facteur, influera sur la quantité consommée du carburant, vu que plus la taille de la population est grande, plus le nombre d'acquéreurs potentiels de véhicules est élevé et plus la probabilité d'avoir des consommateurs GPLc augmente. La taille de la population représente un très bon indicateur de la consommation potentielle du carburant.

5. La population active: (en millions) (POP_{act})

Les déplacements de la population active auront tendance à pousser à la hausse la consommation du carburant. Aussi, l'augmentation de la taille de la population active augmenterait ses déplacements, ce qui conduirait à une consommation plus importante du carburant.

6. Le PIB : (en millions de Dinars)

Le PIB est la mesure du revenu provenant de la production et son augmentation est l'un des indicateurs clés de la croissance économique et le développement de cette dernière impliquerait une augmentation du PIB. C'est donc le volume d'argent circulant dans le pays. Cet indicateur économique expliquerait la demande de carburant, car le transport représente un segment important dans l'activité économique du pays. Aussi, plus le taux annuel de croissance de cet indicateur de richesse est grand, meilleur est le pouvoir d'achat des citoyens et meilleures seront les opportunités d'achat de nouveaux véhicules.

7. La masse salariale : (en millions de dinars) (Ms)

Elle caractérise la capacité des ménages à demander le carburant.

8. La distance : (Dis)

La distance a une influence sur la consommation du carburant. En effet, plus la distance parcourue par les consommateurs est grande, plus la consommation du carburant est grande.

9. La taille de la population utilisant le Transport en Commun (en millions) (POP_{tc})

Les citoyens pourraient, dans des circonstances particulières, être amenés à délaissé leur propre véhicule au profit du transport en commun ce qui représente un facteur de réduction de la consommation du carburant.

Remarque:

Par manque de données, la distance moyenne parcourue ne sera pas prise en considération dans notre étude.

Ayant choisi les variables explicatives, nous aborderons les étapes de la formulation économique, suivie de la formulation économétrique de la consommation du GPLc, le but étant de modéliser cette dernière.

III. Formulation économique:

Dans notre étude, il s'agit de définir la relation qui existe entre les variables explicatives (indépendantes) retenues et la variable à expliquer (dépendante) qui est la consommation du GPLc. Cette relation est symbolisée par la fonction f . Elle peut s'écrire comme suit :

$$C_{GPLc}(t) = f[\text{var}_1(t), \text{var}_2(t), \dots, \text{var}_n(t)]$$

Avec :

C : Variable dépendante

$\text{var}_1(t) + \text{var}_2(t) + \dots + \text{var}_n(t)$: Variables explicatives (indépendantes)

Nous avons donc : $C_{GPLc}(t) = f(P_{GPLc}, P_{gas}, P_{ess}, Ms, Tcr, Pop, Pib, Pop_{act}, Pop_{tc}, Parc)(t) + \varepsilon_t$

Où :

$C_{GPLc}(t)$: La consommation du carburant à l'année t ;

P_{GPLc} : Le prix du GPLc à l'année t ;

P_{gas} : Le prix du gasoil à l'année t ;

P_{ess} : Le prix de l'essence à l'année t ;

Ms : La masse salariale de l'année t ;

Tcr : Le taux de croissance du parc automobile à l'année t ;

Pop : La taille de la population à l'année t ;

Pib : Le taux de croissance économique à l'année t ;

Pop_{act} : La taille de la population active à l'année t .

Pop_{tc} : La taille de la population qui utilise le transport en commun à l'année t .

$Parc$: la taille du parc automobile à l'année t ;

ε_t : L'erreur statistique qui représente la différence entre l'observation et la valeur estimée par le modèle à l'année t .

IV. Formulation économétrique :

L'estimation de la consommation du GPLc, en ses variables explicatives, sera conduite selon la méthode de la **régression multiple**, présentée précédemment.

1. Les modèles testés :

Pour trouver le modèle susceptible d'expliquer au mieux la consommation du GPLc, il est nécessaire de tester plusieurs variantes.

Nous avons pour notre part, analysé 3 spécifications possibles du modèle de régression :

- Le modèle linéaire simple ;
- Le modèle linéaire logarithmique;
- Le modèle à effet de parc.

1.1 Le modèle linéaire simple : [JON 99]

Le modèle statistique le plus simple à utiliser et à comprendre est le modèle de régression linéaire. C'est le modèle qui permet de mesurer l'influence d'un ensemble de variables quantitatives sur une autre variable quantitative. Dans notre cas le modèle s'écrit comme suit:

$$C_{\text{GPL}_c}(t) = \beta_0 + \beta_1 P_{\text{gpl}_t} + \beta_2 P_{\text{gas}_t} + \beta_3 P_{\text{ess}_t} + \beta_4 Ms_t + \beta_5 Tcr_t + \beta_6 Parc + \beta_7 Pop_t + \beta_8 PIB_t + \beta_9 Pop_{act_t} + \beta_{10} Pop_{tc_t} + \varepsilon_t$$

Où : les B_i représentent les paramètres du modèle (coefficients de régression).

Lorsque la relation entre les variables est supposée linéaire, chaque paramètre est interprété comme la dérivée partielle de la variable dépendante par rapport à la variable explicative concernée. Chaque paramètre mesure donc les variations de la variable dépendante suite à une augmentation ou une réduction de la variable explicative concernée, les autres variables explicatives restant inchangées.

1.2 Le modèle logarithmique : [JON 99]

Les principes de la régression linéaire se généralisent à la régression logarithmique. Dans ce dernier type de régression, on utilise les log des quantités comme des variables. Ainsi, les quantités n'interviennent plus directement, mais à travers leur logarithme.

Dans notre cas, le modèle s'écrit comme suit :

$$\text{Log } C_{\text{GPL}_c}(t) = \beta_0 + \beta_1 \log P_{\text{gpl}_t} + \beta_2 \log P_{\text{gas}_t} + \beta_3 \log P_{\text{ess}_t} + \beta_4 \log Ms_t + \beta_5 \log Tcr_t + \beta_6 \log Parc + \beta_7 \log Pop_t + \beta_8 \log PIB_t + \beta_9 \log Pop_{act_t} + \beta_{10} \log Pop_{tc_t} + \varepsilon_t$$

La spécification logarithmique permet, entre autres, d'interpréter les coefficients en termes d'élasticité.

1.3 Le modèle à effet de parc : [LAN 03]

Dans ce type de modèle, nous introduisons dans le modèle logarithmique le nombre de véhicules à essence, le nombre de véhicules Diesel, et le nombre de véhicules GPLc, ainsi que la part de ces derniers dans le parc automobile total. L'objectif étant de tester la pertinence de chaque type de véhicules dans l'explication de la consommation du GPLc.

Cette approche se traduit par :

$$\log C_{\text{GPLc}}(t) = \beta_0 + \beta_1 \log P_{\text{gpl}_t} + \beta_2 \log P_{\text{gas}_t} + \beta_3 \log P_{\text{ess}_t} + \beta_4 \log Ms_t + \beta_5 \log Tcr_t + \beta_6 \log Pop_t + \beta_7 \log Pib_t + \beta_8 \log Pop_{\text{act}_t} + \beta_9 \log Pop_{\text{tc}_t} + \beta_{10} \log V_t + \beta_{11} \log (V_t / \text{parc}_t) + \beta_{12} \log X_t + \beta_{13} \log (X_t / \text{Parc}_t) + \beta_{14} \log W_t + \beta_{15} \log (W_t / \text{Parc}_t) + \varepsilon_t$$

Avec :

V_t : Le nombre de véhicules essence ;

W_t : Le nombre de véhicules Diesel ;

X_t : Le nombre de véhicules GPLc ;

Parc_t : le nombre total de véhicules (la taille du parc automobile).

Remarque : Vu que les données, relatives, au nombre de véhicules essences, diesel et GPLc ne nous étaient pas disponibles, seules les deux premières variantes seront testées.

Comme nous disposons, pour chaque modèle, d'un nombre assez élevé de variables susceptibles d'être des variables explicatives de la consommation du GPLc (le nombre est égal à 10), il est plus que nécessaire d'utiliser une des méthodes qui permettent de trouver un modèle avec le minimum de variables exogènes et un maximum d'explication à un risque prés. En effet, la sélection des entrées est une étape essentielle pour toute conception de modèle. Cette sélection va nous permettre de réduire le nombre de variables explicatives de telle sorte que le modèle final ne contienne que les variables les plus pertinentes pour la modélisation.

Pour ce faire, il existe plusieurs méthodes suggérées par la théorie. Il s'agit: [BOU 02]

- Des méthodes de combinaison de toutes les régressions ;
- De la procédure d'élimination progressive des variables « Backward Elimination »
- De la méthode de sélection progressive « Forward Regression »
- De la méthode de régression pas à pas « Stepwise Regression »

➤ *Méthodes de combinaison de toutes les régressions*

La méthode de combinaison de toutes les régressions, représente la méthode la plus simple dans son exposé. En effet, il s'agit d'estimer toutes les combinaisons de régressions possibles ($2^k - 1$ possibilités avec $k =$ nombre de variables explicatives candidates) et l'on compare toutes les possibilités dans le but de conserver la meilleure, selon un critère bien défini. La limite d'utilisation de cette méthode est liée au nombre initial de variables explicatives candidates. Une telle approche, certes optimale, est de mise en œuvre très

lourde, car sa complexité croît exponentiellement avec le nombre de variables. C'est pour cette raison que nous n'allons pas la retenir pour déterminer le modèle optimal.

➤ *La procédure d'élimination progressive*

Elle consiste, sur le modèle complet, à éliminer de proche en proche (c'est à dire en ré-estimant l'équation après chaque élimination) la variable explicative la moins significative. A cette fin, on compare tous les sous modèles au modèle complet et on conserve le meilleur selon un critère donné. Cette procédure n'est utilisable que s'il y a absence de multi colinéarité entre toutes les variables explicatives, ce qui n'est pas toujours le cas (Le terme de multi- colinéarité est employé dans le cas d'un modèle incorporant des séries explicatives qui sont très liées entre elles). Aussi n'avons-nous pas retenu cette méthode.

➤ *La méthode de sélection progressive*

On commence par estimer le modèle le plus simple et on le compare aux autres ($k-1$) modèles simples à une variable explicative (k étant le nombre initial de variables explicatives). On choisit alors le meilleur, suivant un critère donné, et l'on réitère la procédure jusqu'à ce que l'ajout d'une nouvelle variable n'améliore plus la qualité du modèle.

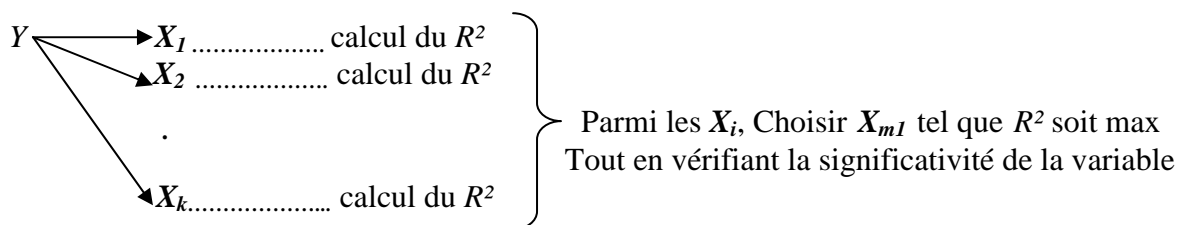
➤ *La méthode de régression pas à pas*

Elle est identique à la précédente, sauf qu'après avoir incorporé une nouvelle variable explicative, nous examinons la pertinence des variables explicatives préalablement sélectionnées en vue d'éliminer du modèle celles qui sont non pertinentes. Ainsi, nous utilisons, dans notre application, « la régression pas à pas » pour chacun des modèles cités ci-dessus. Cette régression va nous permettre de trouver le modèle optimal pour le modèle simple et le modèle optimal pour le modèle logarithmique.

Nous pouvons résumer les étapes de la « régression pas à pas » comme suit :

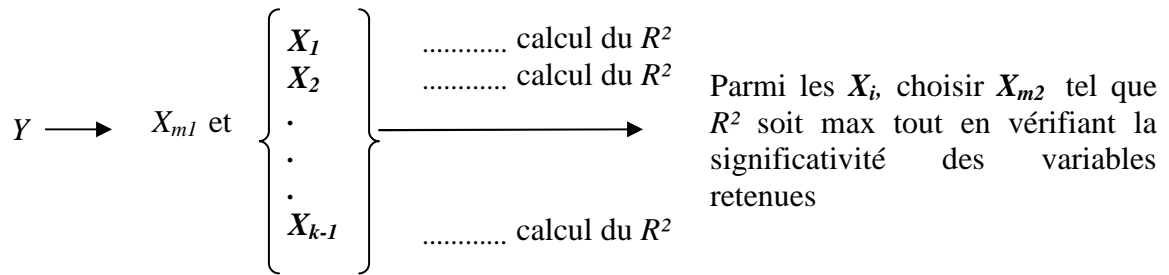
1ere étape : Modèle à 2 variables

Dans cette étape, il s'agit de faire k régressions simples et choisir la variable X_{m1} dont le taux de réalité R^2 est le plus élevé :



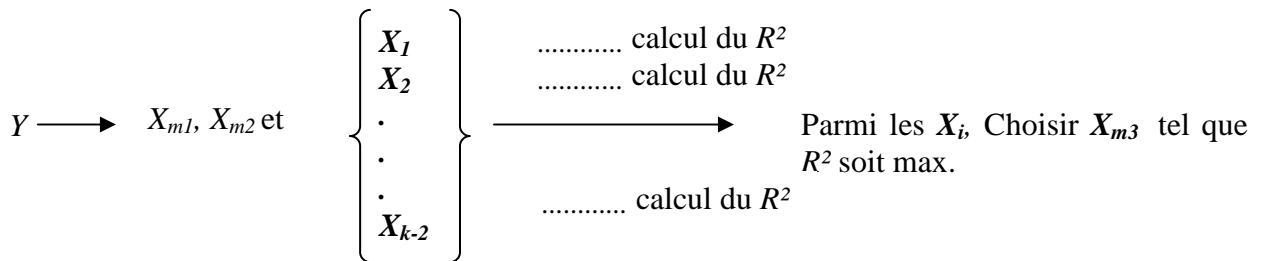
2eme étape : Modèle à 3 variables

Une fois la 1^{ère} variable sélectionnée, il s'agira de faire $(k - 1)$ régressions simples et de choisir la variable X_{m2} dont le taux de réalité R^2 est le plus élevé comme suit :



3eme étape : Modèle à 4 variables

Faire $(k - 2)$ régressions simples et choisir la variable X_{m3} dont le taux de réalité R^2 est le plus élevé comme suit :



Répéter la procédure pour le choix de la 4ème variable X_{m4} , la 5ème variable X_{m5} etc....
 La sélection s'arrête lorsque l'ajout d'une nouvelle variable n'améliore plus le R^2 .

2. Critère de choix de la meilleure variante :

Une fois la régression pas à pas appliquée, le modèle optimal étant trouvé pour chaque variante (pour le modèle simple et pour le modèle logarithmique), le choix de la meilleure variante se fera sur la base du meilleur taux de réalité (R^2) tout en vérifiant :

- La significativité globale du modèle ;
- La significativité des variables explicatives ;
- La non corrélation des résidus ;
- L'absence du multi colinéarité entre les variables ;
- La stabilité du modèle et de ses coefficients ;

On peut illustrer ceci par la *figure II.4* qui résume la procédure de sélection de la meilleure variante :

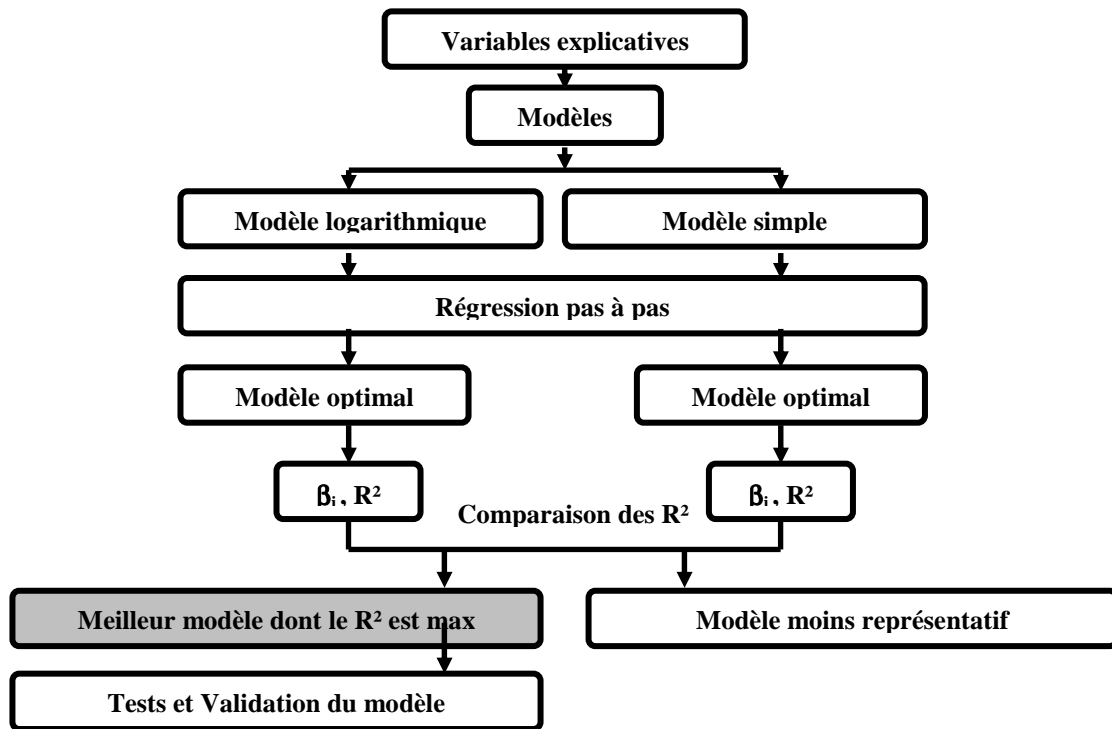


Figure II. 4 : Schéma simplifié de la procédure de la sélection du meilleur modèle

V. Application :

Les données :

La collecte des données des différentes variables nécessaires à la modélisation de la consommation du GPLc s'est faite auprès du Ministère de l'Énergie et des Mines, de NAFTA, de l'Office Nationale des Statistiques (ONS), de l'Agence nationale pour la Promotion et la Rationalisation de l'Utilisation de l'Énergie (APRUE) et des stations de services.

Les données sont représentées sous forme de chroniques annuelles avec 27 observations (1980-2006) (cf. annexe 2)

L'outil informatique utilisé:

Il existe plusieurs logiciels statistiques qui peuvent être utilisés : *SPSS*, *STATISTICA*, *EVIIEWS*, etc. Le logiciel *EVIIEWS* est choisi dans notre étude pour son efficacité et sa pertinence (cf. annexe 3).

Aussi, ce logiciel est utilisé actuellement, comme principal outil informatique, dans de nombreuses études de modélisation et de prévision, ce qui nous a rendu sa compréhension et son utilisation plus aisées.

Nous proposons dans ce qui suit d'appliquer la régression pas à pas sur chaque variante, à savoir : le modèle simple et le modèle logarithmique.

1. Estimation des paramètres du modèle simple :

Nous nous proposons d'appliquer la régression pas à pas sur le modèle simple. Ce dernier est donné par :

$$C_{\text{Gplc}}(t) = \beta_0 + \beta_1 P_{\text{gpl } t} + \beta_2 P_{\text{gas } t} + \beta_3 P_{\text{ess } t} + \beta_4 \text{Ms } t + \beta_5 \text{Tcr } t + \beta_6 \text{Parc } t + \beta_7 \text{Pop } t + \beta_8 \text{Pib } t + \beta_9 \text{Pop}_{\text{act } t} + \beta_{10} \text{Pop}_{\text{tc } t} + \varepsilon_t$$

Les étapes de la régression pas à pas sont résumées dans les tableaux suivants :

➤ **1ère étape :**

Tableau II. 2 : 1^{ère} étape de la régression pas à pas

Modèle	Variables explicatives	R ²	Modèle retenu (R ² max)	T- test (significativité des variables)	Variables retenues
Modèle 1 :	P _{GPLc}	0.7809			
Modèle 2 :	P _{gas}	0.7767			
Modèle 3 :	P _{ess}	0.8478			
Modèle 4 :	POP	0.7352	Modèle 9 :		
Modèle 5 :	POP _{act}	0.8721	$C_{\text{GPLc } t} = \beta_0 +$	<u>Pour PIB :</u> t- Statistic (=20.17919)	PIB
Modèle 6 :	POP _{tc}	0.3869	$\beta_1 \text{PIB}_t + \varepsilon_t$	≥ t- tabulée (= 2.060)	
Modèle 7 :	T _{cr}	0.2147			
Modèle 8	Parc	0.5556			
Modèle 9 :	PIB	0.9421			
Modèle 10	Ms	0.9278			

• **Choix du modèle :**

D'après les 10 régressions effectuées, on remarque que la meilleure régression est celle qui correspond à la variable explicative PIB, car c'est celle qui présente le meilleur taux de réalité R². Le modèle s'écrit comme suit : $C_{\text{GPLc } t} = \beta_0 + \beta_1 \text{PIB}_t + \varepsilon_t$ avec $R^2 = 0.9421$

• **Le T- Test :**

D'après le T- Test donné dans le tableau II.2, la variable PIB est significativement différente de zéro. La 1ère variable retenue est donc la variable « PIB »

Nous passons à la deuxième étape pour la sélection de la 2ème variable explicative

➤ 2ème étape :

Tableau II. 3 : 2ème étape de la régression pas à pas

Modèle	Variables explicatives	R ²	Modèle retenu (R ² max)	T- test (significativité des variables)	Variables retenues
Modèle 1	PIB, P _{gpl}	0.9651	Modèle 6: $C_{Gplc_t} = \beta_0 + \beta_1 PIB_t + \beta_2 POP_{tc_t} + \varepsilon_t$	Pour PIB : t- Statistic(= 26.64) ≥ t- tabulée (=2.064) Pour POP _{tc} : t- Statistic(= 6.72) ≥ t- tabulée (=2.064)	PIB POP _{tc}
Modèle 2	PIB, P _{gas}	0.9650			
Modèle 3	PIB, P _{ess}	0.9537			
Modèle 4	PIB, POP	0.9614			
Modèle 5	PIB, POP _{act}	0.9507			
Modèle 6	PIB, POP_{tc}	0.9799			
Modèle 7	PIB, T _{cr}	0.9644			
Modèle 8	PIB, Parc	0.9574			
Modèle 9	PIB, MS	0.9422			

• **Choix du modèle:**

D'après les 9 régressions effectuées, on remarque que la meilleure régression est celle qui correspond aux variables explicatives : PIB et POP_{tc}, car c'est celle qui présente le meilleur taux de réalité R², taux supérieur à celui du modèle retenu à l'étape précédente.

Le modèle s'écrit comme suit :

$$C_{GPLc_t} = \beta_0 + \beta_1 PIB_t + \beta_2 POP_{tc_t} + \varepsilon_t \text{ avec } R^2 = 0.9799$$

• **Le T- Test :**

D'après le T- Test (le tableau II.3), les variables PIB et POP_{tc} sont significativement différentes de zéro. Dans cette étape, les variables retenues sont donc : PIB, et POP_{tc}

➤ 3ème étape :

Tableau II. 4 : 3ème étape de la régression pas à pas

Modèle	Variables explicatives	R ²	Modèle retenu (R ² max)	T- test (significativité des variables)	Variables retenues	
Modèle 1 :	PIB, POP _{tc} , P _{Gplc}	0.9800	_Modèle 6: $C_{gplc_t} = \beta_0 + \beta_1 PIB_t + \beta_2 POP_{tc_t} + \beta_3 T_{cr_t} + \varepsilon_t$	<u>Pour PIB :</u> t- Statistic (= 31.51) ≥ t- tabulée (=2.069)	PIB POP _{tc} , T _{cr}	
Modèle 2 :	PIB, POP _{tc} , P _{gas}	0.9799				
Modèle 3 :	PIB, POP _{tc} , P _{ess}	0.9807				
Modèle 4 :	PIB, POP _{tc} , POP	0.9828				
Modèle 5 :	PIB, POP _{tc} , POP _{act}	0.9808				
Modèle 6 :	PIB, POP_{tc}, T_{cr}	0.9862				<u>Pour POP_{tc} :</u> t- Statistic(5.95) ≥ t- tabulée (=2.069)
Modèle 7	PIB, POP _{tc} , Parc	0.9840				<u>Pour T_{cr} :</u> t- Statistic(= 3.23) ≥ t- tabulée (=2.069)
Modèle 8 :	PIB, POP _{tc} , MS	0.9822				

• **Choix du modèle :**

Nous remarquons là aussi, que la meilleure régression est celle qui correspond aux variables explicatives : PIB, POP_{tc} et T_{cr} car c'est celle qui présente le meilleur taux de réalité R², ce dernier est supérieur au taux de réalité du modèle retenu à l'étape précédente.

Le modèle s'écrit comme suit :

$$C_{GPLc_t} = \beta_0 + \beta_1 PIB_t + \beta_2 POP_{tc_t} + \beta_3 T_{cr_t} + \varepsilon_t \text{ avec : } R^2 = 0.9862$$

• **Le T- Test :**

Les variables PIB, POP_{tc} et T_{cr} sont significativement différentes de zéro. Ceci est illustré par le T- Test donné dans le tableau II.4. Les variables retenues sont donc : PIB, POP_{tc} et T_{cr}.

➤ 4ème étape :

Tableau II. 5 : 4^{ème} étape de la régression pas à pas

Modèle	Variables explicatives	R ²	Modèle retenu (R ² max)	T- test (significativité des variables)	Variables retenues	
Modèle 1 :	PIB, POP _{tc} , T _{cr} , P _{gpl}	0.9863	Modèle 5 : $C_{gplc_t} = \beta_0 + \beta_1 PIB_t + \beta_2 POP_{tc_t} + \beta_3 T_{cr_t} + \beta_4 POP_{act_t} + \varepsilon_t$	<u>Pour PIB:</u> t- Statistic(4.33) ≥ t- tabulée (=2.074)	PIB POP _{tc} , T _{cr} POP _{act}	
Modèle 2 :	PIB, POP _{tc} , T _{cr} , P _{gas}	0.9862				
Modèle 3 :	PIB, POP _{tc} , T _{cr} , P _{ess}	0.9877				
Modèle 4 :	PIB, POP _{tc} , T _{cr} , POP	0.9878				
Modèle5 :	PIB, POP_{tc}, T_{cr}, POP_{act}	0.9897				<u>Pour POP_{tc} :</u> t- Statistic(= 6.90) ≥ t- tabulée (=2.074)
Modèle 6 :	PIB, POP _{tc} , T _{cr} , PARC	0.9868				<u>Pour T_{cr} :</u> t- Statistic(= 4.38) ≥ t- tabulée (=2.074)
Modèle 7 :	PIB, POP _{tc} , T _{cr} , MS	0.9892				<u>Pour POP_{act} :</u> t- Statistic(= 2.75) ≥ t- tabulée (=2.074)

• **Choix du modèle:**

La meilleure régression est celle qui correspond aux variables explicatives: PIB, POP_{tc}, T_{cr} et POP_{act}, car c'est celle qui présente le meilleur taux de réalité R², taux supérieur au R² du modèle retenu à l'étape précédente.

Le modèle s'écrit comme suit :

$$C_{GPLc_t} = \beta_0 + \beta_1 PIB_t + \beta_2 POP_{tc_t} + \beta_3 T_{cr_t} + \beta_4 POP_{act_t} + \varepsilon_t \quad \text{avec : } R^2 = 0.9897$$

• **Le T- Test :**

D'après le T- Test qui apparaît dans le tableau II.5, les variables PIB, POP_{tc}, T_{cr} et POP_{act} sont significativement différentes de zéro. Les variables retenues sont donc : PIB, POP_{tc}, T_{cr}, et POP_{act}.

➤ 5ème étape :

Tableau II. 6 : 5^{ème} étape de la régression pas à pas

Modèle	Variabes explicatives	R ²	Modèle retenu (R ² max)	T- test (significativité des variables)	Variable s retenues
Modèle 1 :	PIB, POP _{tc} , T _{cr} , POP _{act} , P _{gpl}	0.9898	Modèle 3 : $C_{gplc_t} = \beta_0 + \beta_1 PIB_t + \beta_2 POP_{tc_t} + \beta_3 T_{cr_t} + \beta_4 POP_{act_t} + \beta_5 P_{ess_t} + \varepsilon_t$	<u>Pour PIB:</u> t- Statistic(= 2.84) ≥ t- tabulée (=2.080)	PIB POP _{tc} , T _{cr} POP _{act} P _{ess}
Modèle 2 :	PIB, POP _{tc} , T _{cr} , POP _{act} , P _{gas}	0.9898		<u>Pour POP_{tc} :</u> t- Statistic(= 7.14) ≥ t- tabulée (=2.080)	
Modèle3 :	PIB, POP_{tc}, T_{cr}, POP_{act}, P_{ess}	0.9915		<u>Pour T_{cr} :</u> t- Statistic(= 4.92) ≥ t- tabulée (=2.080)	
Modèle 4 :	PIB, POP _{tc} , T _{cr} , POP _{act} , POP	0.9904		<u>Pour POP_{act} :</u> t- Statistic(3.05) ≥ t- tabulée (=2.080)	
Modèle 5 :	PIB, POP _{tc} , T _{cr} , POP _{act} PARC	0.9911		<u>Pour P_{ess} :</u> t- Statistic(= 2.098) ≥ t- tabulée (=2.080)	
Modèle 6 :	PIB, POP _{tc} , T _{cr} , POP _{act} , MS	0.9898			

• **Choix du modèle:**

La meilleure régression est celle qui correspond aux variables explicatives : PIB, POP_{tc}, T_{cr}, POP_{act} et P_{ess}, car c'est celle qui présente le meilleur taux de réalité R², supérieur au R² du modèle retenu à l'étape précédente.

Le modèle s'écrit comme suit :

$$C_{GPLc_t} = \beta_0 + \beta_1 PIB_t + \beta_2 POP_{tc_t} + \beta_3 T_{cr_t} + \beta_4 POP_{act_t} + \beta_5 P_{ess_t} + \varepsilon_t \quad \text{avec :}$$

$$R^2 = 0.9915$$

• **Le T- Test :**

Les variables PIB, POP_{tc}, T_{cr}, POP_{act} et P_{ess} sont significativement différentes de zéro. Ceci est illustré par le T- Test donné dans le tableau II.6. Les variables retenues sont donc : PIB, POP_{tc}, T_{cr}, POP_{act} et P_{ess}.

➤ 6ème étape :

Tableau II. 7 : 6^{ème} étape de la régression pas à pas

Modèle	Variables explicatives	R ²	T- test (significativité des variables)	Variables retenues
Modèle1:	PIB, POP _{tc} , T _{cr} , POP _{act} , P _{ess} , P _{gpl}	0.9926	<u>Pour P_{gpl}</u> : t- Statistic(= 1.59) ≤ t- tabulée (=2.086)	PIB POP _{tc} , T _{cr} POP _{act} P _{ess}
Modèle2:	PIB, POP _{tc} , T _{cr} , POP _{act} , P _{ess} , P _{gas}	0.9927	<u>Pour P_{gas}</u> : t- Statistic(= 1.80) ≤ t- tabulée (=2.086)	
Modèle3:	PIB, POP _{tc} , T _{cr} , POP _{act} , P _{ess} , POP	0.9925	<u>Pour POP</u> : t- Statistic(= 1.35) ≤ t- tabulée (=2.086)	
Modèle4 :	PIB, POP _{tc} , T _{cr} , POP _{act} , P _{ess} , Parc	0.9921	<u>Pour Parc:</u> t- Statistic(= 1.807) ≤ t- tabulée (=2.086)	
Modèle5 :	PIB, POP _{tc} , T _{cr} , POP _{act} , P _{ess} , MS	0.9915	<u>Pour MS:</u> t- Statistic(= 0.26) ≤ t- tabulée (=2.086)	

D'après les 5 régressions effectuées, la nouvelle variable explicative intégrée dans chaque modèle n'est jamais significative. En effet, d'après le T- Test, illustré dans le tableau II.7, les variables P_{gpl}, P_{gas}, POP, Parc, Ms sont non significativement différentes de zéro. Nous remarquons que pour chaque variable, nous avons: t- Statistic ≤ t- tabulée. La procédure est donc arrêtée.

En conclusion, les variables retenues pour expliquer au mieux la consommation nationale du GPLc sont :

- Le PIB ;
- La taille de la population utilisant le transport en commun (POP_{tc}) ;
- Le taux de croissance du parc automobile (T_{cr}) ;
- La taille de la population active (POP_{act}).
- Et le prix de l'essence (P_{ess})

Le modèle s'écrit comme suit :

$$C_{GPLct} = \beta_0 + \beta_1 \text{PIB}_t + \beta_2 \text{POP}_{tc t} + \beta_3 \text{Tcr}_t + \beta_4 \text{POP}_{act t} + \beta_5 \text{P}_{ess t} + \varepsilon_t \quad \text{avec :}$$

$$R^2 = 0.9915$$

2. Estimation des paramètres du modèle logarithmique :

Le modèle logarithmique s'écrit comme suit :

$$\text{Log } C_{\text{GPL}t} = \beta_0 + \beta_1 \log P_{\text{gpl}t} + \beta_2 \log P_{\text{Gas}t} + \beta_3 \log P_{\text{ess}t} + \beta_4 \log \text{Ms}_t + \beta_5 \log \text{Tcr}_t + \beta_6 \log \text{Parc}_t + \beta_7 \log \text{Pop}_t + \beta_8 \log \text{PIB}_t + \beta_9 \log \text{POPact}_t + \beta_{10} \log \text{POPt}_t + \varepsilon_t$$

L'application de la « régression pas à pas » sur le modèle logarithmique conduit aux étapes suivantes :

➤ *1ère étape :*

Tableau II. 8 : 1^{ère} étape de la régression pas à pas pour le modèle logarithmique

Modèle	Variables explicatives	R ²	Modèle retenu (R ² max)	T- test (significativité des variables)	Variables retenues
Modèle 1 :	Pgpl	0.7979			
Modèle 2 :	Pgas	0.7678			
Modèle 3 :	Pess	0.9139			
Modèle 4 :	POP	0.9079	Modèle 9 :		
Modèle 5 :	POPact	0.9024	$\text{Log } (C_{\text{gpl}t}) = \beta_0 +$	<u>Pour PIB :</u>	PIB
Modèle 6 :	POPt	0.2174	$\beta_1 \log (\text{PIB}_t) + \varepsilon_t$	t- Statistic (=13.49)	
Modèle 7 :	Tcr	0.061		≥	
Modèle 8 :	Parc	0.8295		t- tabulée (= 2.060)	
Modèle 9 :	PIB	0.9238			
Modèle 10	MS	0.8828			

• *Choix du modèle :*

D'après les 10 régressions effectuées, on remarque que la meilleure régression est celle qui correspond à la variable explicative PIB, car c'est celle qui présente le meilleur taux de réalité R². Le modèle s'écrit comme suit :

$$\text{Log } (C_{\text{GPL}t}) = \beta_0 + \beta_1 \log \text{PIB}_t + \varepsilon_t \text{ avec : } R^2 = 0.9238$$

• *Le T- Test :*

D'après le T- Test donné dans le tableau II.8, la variable PIB est significativement différente de zéro, donc la 1^{ère} variable retenue est la variable « PIB ».

➤ 2^{ème} étape :

Tableau II. 9 : 2^{ème} étape de la régression pas à pas pour le modèle logarithmique

Modèle	Variables explicatives	R ²	T- test (significativité des variables)	Variables retenues
Modèle 1 :	PIB, P _{gpl}	0.9591	<u>Pour P_{gpl} :</u> t- Statistic(= 0.43) ≤ t- tabulée (=2.064)	PIB
Modèle 2 :	PIB, P _{gas}	0.9607	<u>Pour P_{gas} :</u> t- Statistic(= 1.02) ≤ t- tabulée (=2.064)	
Modèle 3 :	PIB, P _{ess}	0.9676	<u>Pour P_{ess} :</u> t- Statistic(= 1.18) ≤ t- tabulée (=2.064)	
Modèle 4 :	PIB, POP	0.9588	<u>Pour POP :</u> t- Statistic(= 0.78) ≤ t- tabulée (=2.064)	
Modèle 5 :	PIB, POP _{act}	0.9587	<u>Pour POP_{act} :</u> t- Statistic(= 0.64) ≤ t- tabulée (=2.064)	
Modèle 6 :	PIB, POP _{tc}	0.9594	<u>Pour POP_{tc} :</u> t- Statistic(= 0.98) ≤ t- tabulée (=2.064)	
Modèle 7 :	PIB, T _{cr}	0.9609	<u>Pour T_{cr} :</u> t- Statistic(= 1.17) ≤ t- tabulée (=2.064)	
Modèle 8 :	PIB, Parc	0.9588	<u>Pour Parc :</u> t- Statistic(= 0.24) ≤ t- tabulée (=2.064)	
Modèle 9 :	PIB, MS	0.9624	<u>Pour Ms :</u> t- Statistic(= 1.32) ≤ t- tabulée (=2.064)	

D'après les neuf régressions effectuées, la nouvelle variable explicative intégrée dans chaque modèle n'est jamais significative. En effet, le T- Test, illustré dans le tableau II.9, indique que les variables : P_{gpl}, P_{gas}, P_{ess}, POP, POP_{act}, POP_{tc}, T_{cr}, Parc, Ms sont non significativement différentes de zéro. En effet, pour chaque variable: t- Statistic ≤ t- tabulée. La procédure est donc arrêtée.

En conclusion, la seule variable retenue pour expliquer au mieux la consommation nationale du GPLc est le PIB. Le modèle s'écrit comme suit :

$$\text{Log}(C_{\text{GPLc}}) = \beta_0 + \beta_1 \log(\text{PIB}_t) + \varepsilon_t \quad \text{avec : } R^2 = 0.9238$$

3. Choix du modèle:

Le modèle simple et le modèle logarithmique étant construits, le choix du meilleur modèle se fera sur la base du taux de réalité (R^2_{\max}). Le modèle qui a le meilleur R^2 est celui qui représente au mieux la réalité.

Ainsi nous avons :

Tableau II. 10 : Critère de choix du meilleur modèle

Modèle	R^2	\bar{R}^2
Modèle simple	0.9915	0.9895
Modèle logarithmique	0.9238	0.9187

Nous remarquons bien que le taux de réalité R^2 du modèle simple est supérieur au taux de réalité R^2 du modèle logarithmique, donc le meilleur modèle est bien le modèle simple avec un $R^2 = 0.9915$. Ce dernier s'écrit comme suit :

$$C_{GPLc_t} = \beta_0 + \beta_1 \text{PIB}_t + \beta_2 \text{POP}_{tc_t} + \beta_3 \text{Tcr}_t + \beta_4 \text{POP}_{act_t} + \beta_5 \text{P}_{ess_t} + \varepsilon_t$$

4. Tests et validation du modèle simple:

Après avoir déterminé le modèle qui explique au mieux la consommation du GPLc, il est essentiel vérifier dans une seconde étape :

- La significativité globale du modèle;
- La significativité des paramètres estimés du modèle;
- La non corrélation des résidus ;
- L'absence de multi colinéarité entre les variables.
- La stabilité du modèle.

La régression du modèle simple, fournie par *EVIEWS* est donnée par le tableau II.11 :

Tableau II. 11: Estimation des paramètres du modèle

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-342126.1	61713.94	-5.543741	0.0000
PIB	0.068137	0.023935	2.846778	0.0097
POPTC	0.000706	0.000100	7.045937	0.0000
TCR	571136.7	114834.1	4.973580	0.0001
POPACT	0.027423	0.008995	3.048634	0.0061
PESS	3497.009	1678.042	2.083982	0.0496
R-squared	0.991522	Mean dependent var		89543.30
Adjusted R-squared	0.989504	S.D. dependent var		117292.6
S.E. of regression	12016.63	Akaike info criterion		21.81910
Sum squared resid	3.03E+09	Schwarz criterion		22.10706
Log likelihood	-288.5579	F-statistic		491.2262
Durbin-Watson stat	1.427023	Prob(F-statistic)		0.000000

4.1 La significativité globale du modèle :

Nous allons nous interroger sur la significativité globale du modèle de la régression obtenu, c'est-à-dire vérifier si l'ensemble des variables explicatives a une influence sur la variable à expliquer. Pour ce faire, il faudra comparer la F- statistique, fournie directement par *EViews*, avec la valeur lue dans la table de Fisher $F(1, n-k-1)$, où k représente le nombre de variables explicatives (dans notre cas $k=5$) et n le nombre d'observations ($n=27$)
Autrement dit, il faudrait comparer la F- statistique à $F(1, 21)$ lue dans la table de Fisher. On obtient :
 $F \text{ statistique} = 491.22 \geq F(1, 21) = 4.32$, ce qui permet de conclure que le modèle est globalement significatif, ceci est confirmé avec la probabilité de nullité (= 0.0000) qui est inférieure à 0.05

4.2 La significativité des paramètres :

Les paramètres du modèle estimé sont tous significativement différents de zéro (la régression pas à pas fait que nous ne retenons que les variables significatives et que les variables non significatives sont rejetées)

4.3 Test sur les résidus :

La représentation graphique de la série résiduelle actuelle et estimée, fournie directement par *EViews* est la suivante :

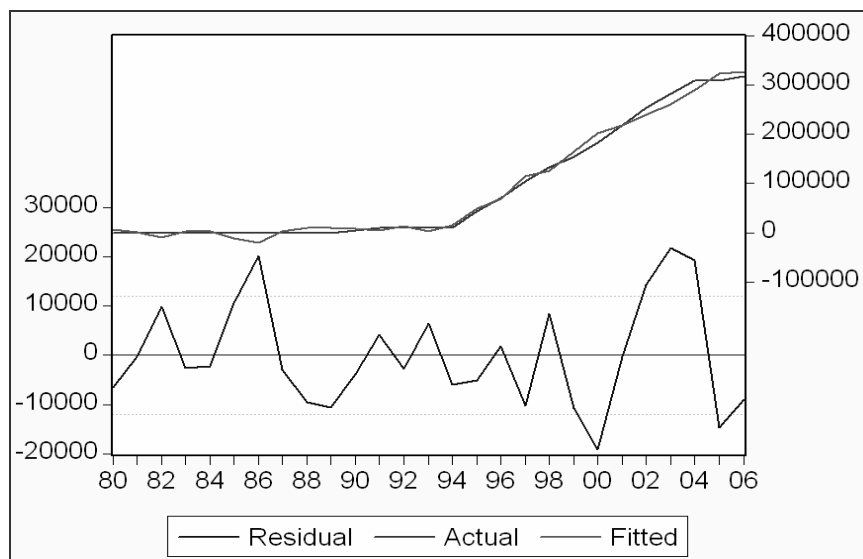


Figure II. 5: Représentation graphique de la série résiduelle, actuelle et estimée

A partir de cette représentation graphique, nous constatons que le graphe de la série estimée est semblable à celui de la série réelle, à un risque près, donc le modèle estimé explique bien à 99,15% la variation totale de la variable dépendante ($R^2=0.9915$).

- **Test de Blancher des résidus**

Nous allons appliquer le test de Durbin- Waston afin de vérifier la blancheur des résidus.

La valeur DW de Durbin Waston, fournie par *EVIIEWS* est égale à 1.4270. Les valeurs critiques, lues dans la table de Durbin- Waston au seuil 0.05, sont: $d_1 = 1.08$ et $d_2 = 1.76$.

On remarque bien que :

$$d_1 < DW < d_2 \quad (1.01 < 1.4270 < 1.86).$$

On ne peut donc rien conclure sur la blancheur des résidus.

Nous utilisons alors la statistique de Ljung- Box Q pour vérifier si les résidus forment un bruit blanc. Afin d'appliquer ce test, nous commençons d'abord par examiner le corrélogramme des résidus :

Le corrélogramme des résidus, fourni directement par *EVIIEWS*, est représenté par la *figure II.6*

Figure II. 6 : Corrélogramme des résidus

Date: 05/24/07 Time: 00:14 Sample: 1980 2006 Included observations: 27					
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.223	0.223	1.4981	0.221
		2 -0.109	-0.167	1.8691	0.393
		3 -0.211	-0.157	3.3161	0.345
		4 0.006	0.085	3.3173	0.506
		5 0.090	0.029	3.6041	0.608
		6 -0.045	-0.113	3.6806	0.720
		7 0.056	0.139	3.8039	0.802
		8 0.014	-0.024	3.8120	0.874
		9 0.010	-0.018	3.8164	0.923
		10 -0.008	0.039	3.8195	0.955
		11 0.022	0.027	3.8422	0.974
		12 -0.014	-0.061	3.8526	0.986

Le logiciel *EVIIEWS* fournit les résultats des fonctions d'autocorrélation simple (colonne AC) et d'autocorrélation partielle (colonne PAC), ainsi que leur corrélogramme associé.

Les bornes de l'intervalle de confiance sont stylisées par des pointillés horizontaux. Ainsi tous les termes qui sont à l'intérieur de cet intervalle sont significativement nuls et ceux qui sortent de ce dernier sont significativement différents de zéro, au seuil 5%.

L'analyse du corrélogramme des résidus (*figure II.6* montre que tous les termes sont à l'intérieur de l'intervalle de confiance, nous pouvons donc conclure que les résidus forment un bruit blanc. Ceci est confirmé par la statistique de Ljung- Box (Q - stat). En effet, cette statistique est inférieure à la valeur théorique de $\chi^2(h)$ quelque soit le retard h , et particulièrement pour $h= 12$, on a :

Q - stat (12) = 3.8526 inférieure à $\chi^2(12) = 21.03$, lue au seuil 0.05, donc les résidus forment bien un bruit blanc.

- **Test de normalité sur les résidus :**

L’histogramme de la distribution des résidus, les valeurs empiriques de Skewness, de Kurtosis ainsi que la statistique de Jarque- Bera, fournis directement par *EViews*, sont donnés ci-après :

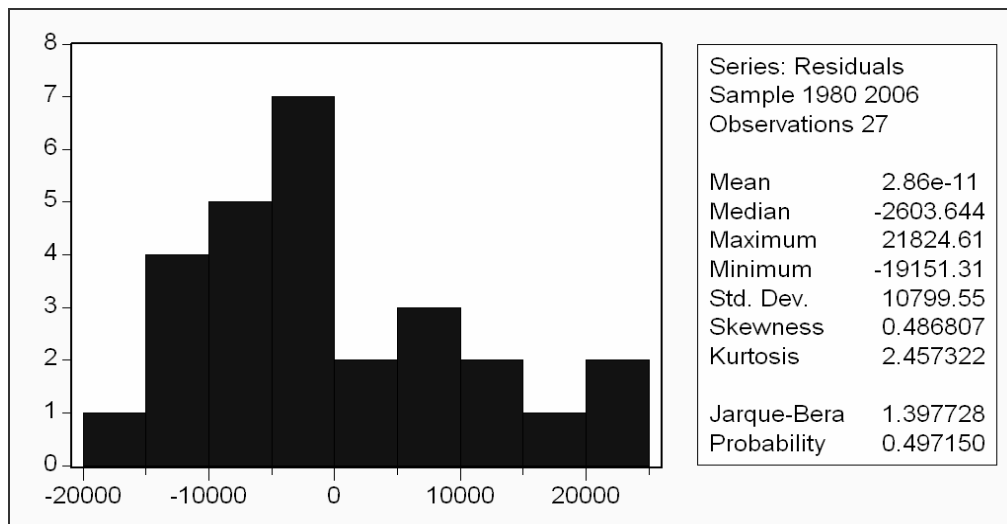


Figure II. 7 : Test de normalité des résidus

Le coefficient de Skewness (asymétrie) = 0.486807

Le coefficient de Kurtosis (aplatissement) = 2.457322

Sur cette base, nous construisons les statistiques centrées réduites correspondantes que l’on compare à 1.96 qui représente la valeur de la loi normale, au seuil 5%, comme suit :

$$v_1 = (\sqrt{(0,486807)} / \sqrt{6/27}) = 1.4800 \text{ inférieur à } 1.96$$

$$v_2 = (3 - 2.457322) / (\sqrt{(24/27)}) = 0.5755 \text{ inférieur à } 1.96$$

L’hypothèse de normalité est donc acceptée. Ceci est confirmé par la statistique de Jarque-Berra, donnée par : JB = 1.397728. Cette statistique est inférieure à la statistique de χ^2 (2) (5.991) lue au seuil 0.05.

Nous pouvons conclure alors que les résidus forment bien un bruit blanc gaussien.

4.4 Test d’existence de multi colinéarité entre les variables :

Le terme de multi- colinéarité est employé dans le cas d’un modèle incorporant des séries explicatives qui sont liées entre elles. [BOU 02]

En effet, lors de la modélisation des phénomènes économiques, les séries explicatives sont toujours plus ou moins liées entre elles

Une multi- colinéarité entre les variables explicatives peut avoir les conséquences suivantes :

- Augmentation de la variance estimée de certains coefficients lorsque la colinéarité entre les variables explicatives augmente ;
- Instabilité des estimateurs des coefficients des moindres carrés, des faibles fluctuations concernant les données entraînent des fortes variations des valeurs estimées des coefficients ;
- En cas de multi- colinéarité parfaite, l'estimation des coefficients est alors impossible et leur variance est infinie.

Afin de détecter l'existence d'une multi- colinéarité entre les variables explicatives, nous allons appliquer le test de Klein. Ce test est fondé sur la comparaison du coefficient de détermination R^2 calculé sur le modèle estimé avec les coefficients de corrélation simple $r_{i,j}$ entre les variables explicatives (pour $i \neq j$) [BOU 02]

Test de Klein :

La matrice des coefficients de corrélation entre les variables explicatives fournie par *EVIEWS* est comme suit :

	PIB	POP _{tc}	Tcr	POP _{act}	P _{ess}
PIB	1	-0.76	-0.60	0.98	0.97
POP _{tc}	-0.76	1	0.66	-0.79	-0.84
Tcr	-0.60	0.66	1	-0.71	-0.64
POP _{act}	0.98	-0.79	-0.71	1	0.96
P _{ess}	0.97	-0.84	-0.64	0.96	1

On compare ces coefficients de corrélation simple entre les variables explicatives avec le taux de réalité du modèle retenu : $R^2 = 0.9915$.

$$\begin{aligned}
 r^2_{\text{PIB, pop}_{tc}} &= 0.5776 \leq 0.9915 & r^2_{\text{POP}_{tc}, \text{POP}_{act}} &= 0.6241 \leq 0.9915 \\
 r^2_{\text{pib, tcr}} &= 0.36 \leq 0.9915 & r^2_{\text{POP}_{tc}, \text{P}_{ess}} &= 0.7056 \leq 0.9915 \\
 r^2_{\text{pib, POP}_{act}} &= 0.9604 \leq 0.9915 & r^2_{\text{Tcr, POP}_{act}} &= 0.5041 \leq 0.9915 \\
 r^2_{\text{pib, P}_{ess}} &= 0.9409 \leq 0.9915 & r^2_{\text{Tcr, P}_{ess}} &= 0.4096 \leq 0.9915 \\
 r^2_{\text{POP}_{tc}, \text{tcr}} &= 0.4356 \leq 0.9915 & r^2_{\text{POP}_{act}, \text{P}_{ess}} &= 0.9216 \leq 0.9915
 \end{aligned}$$

On remarque bien que les coefficients de corrélation simple entre les variables explicatives sont tous inférieurs au coefficient de corrélation multiple du modèle explicatif, on en déduit donc une absence de multi colinéarité.

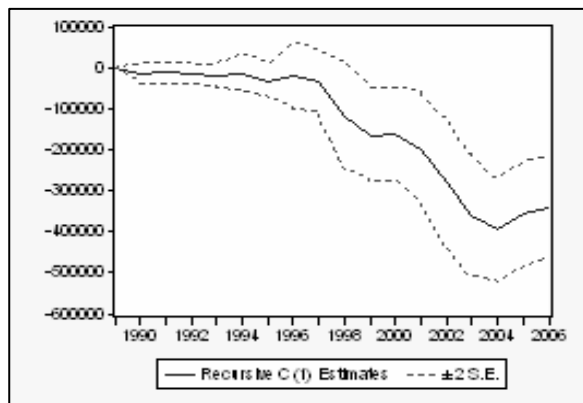
4.5 Test de stabilité :

Après avoir testé : la significativité globale du modèle, la significativité et l'absence de multi colinéarité entre ses variables explicatives ainsi que la blancheur des résidus, il convient de tester la stabilité de ce modèle ainsi que la stabilité de ses coefficients.

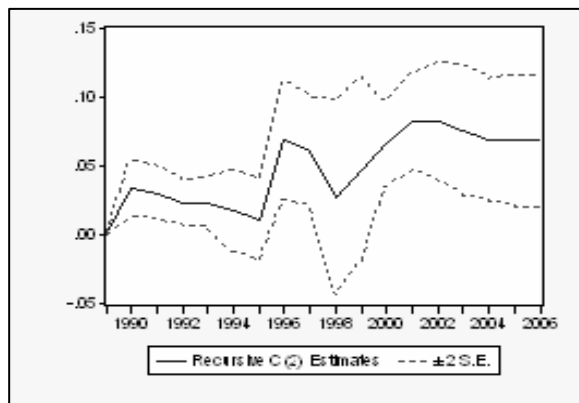
- **Stabilité des coefficients :**

Un simple examen graphique de l'évolution des coefficients du modèle assortis de leurs intervalles de confiance à \pm deux écarts types permet de déterminer des changements structurel. En effet :

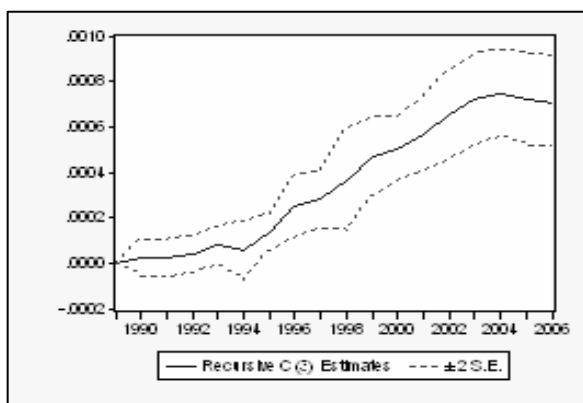
Les graphiques suivants illustrent l'évolution des 6 coefficients estimés :



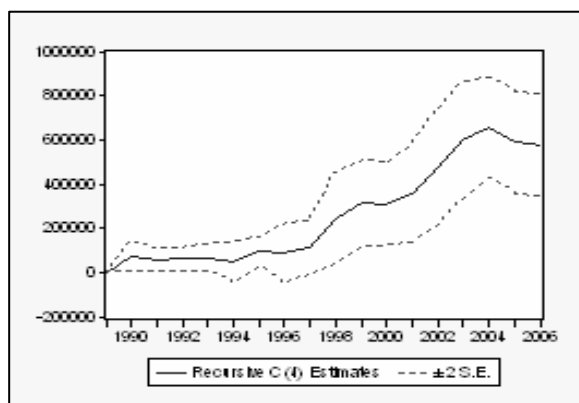
Evolution de la constante



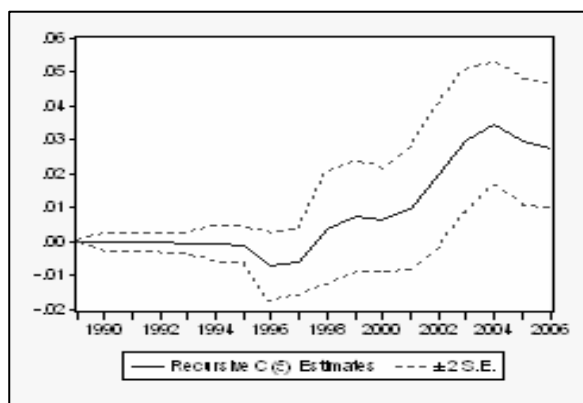
Evolution du coefficient correspondant à PIB



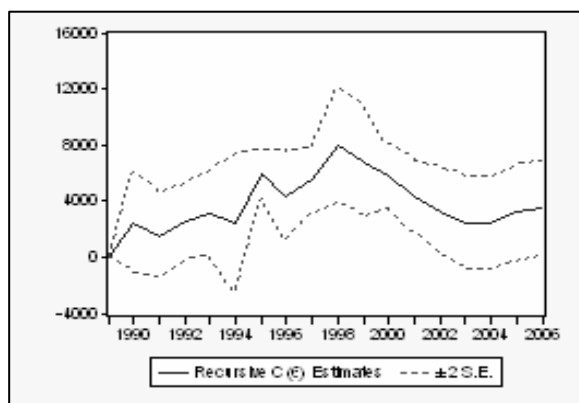
Evolution du coefficient correspondant à POP_t



Evolution du coefficient correspondant à T_{cr}



Evolution du coefficient correspondant à POP_{act}



Evolution du coefficient correspondant à P_{ess}

Figure II. 8 : Evolution des coefficients de régression

Nous constatons que les six coefficients restent toujours à l'intérieur de leur intervalle de confiance, nous rejetons donc l'hypothèse d'un changement structurel.

Nous passons à présent au test de stabilité du modèle de régression.

- **Test sur la stabilité du modèle (Test de *CUSUM*):**

Le test de *CUSUM* (CUmulative SUM) permet de détecter les instabilités structurelles des équations de régression et d'étudier la stabilité du modèle estimé au cours du temps.

Le test de *CUSUM* peut être effectué directement avec le logiciel *EViews*, le résultat est donné par le graphique suivant:

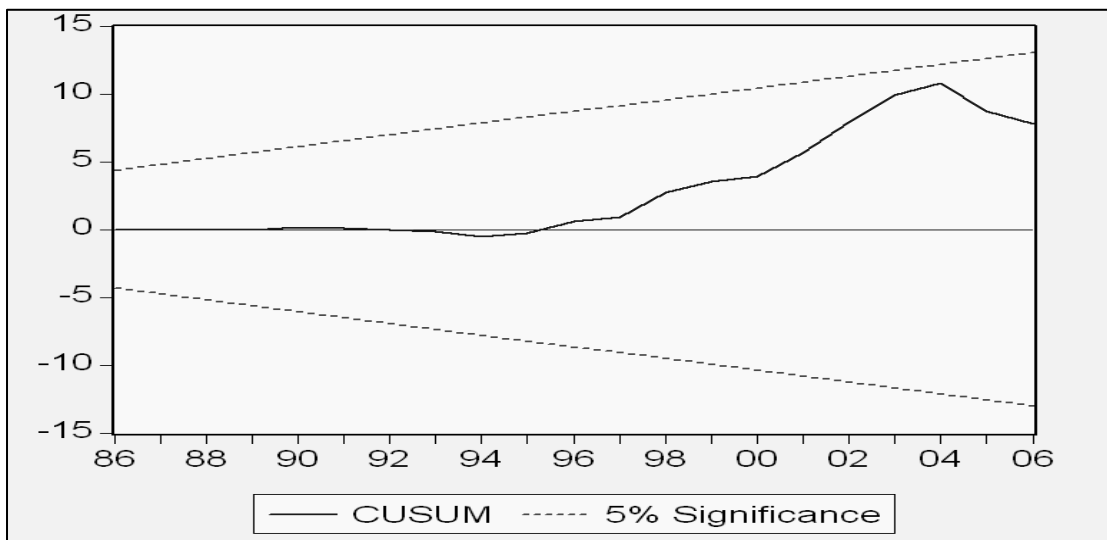


Figure II. 9 : Test de *CUSUM* pour un seuil de + ou - 5%

La statistique *CUSUM* reste toujours dans son intervalle de confiance, là encore nous rejetons l'hypothèse d'un changement structurel.

Conclusion :

Le modèle estimé est globalement significatif, stable, ses paramètres sont tous significativement différents de zéro et les résidus forment un bruit blanc gaussien, le modèle est donc validé. Ce modèle explique bien l'évolution de la consommation du GPLc, avec un taux de réalité égal à 99.15%.

5. Interprétation des résultats:

Après avoir construit le modèle, il conviendra d'interpréter l'influence de chaque variable explicative retenue sur la consommation du GPLc.

5.1 Influence du PIB sur la consommation du GPLc :

Le coefficient, correspondant à la variable explicative PIB (le Produit Intérieur Brut), est positif (= + **0.06813749497**), ceci signifie qu'une augmentation du PIB, aura un effet positif sur la consommation du GPLc. En effet, rappelons que le PIB est la mesure du revenu provenant de la production et son augmentation est l'un des indicateurs clés de la croissance économique. En effet, le PIB est un indicateur de richesse, plus il est grand, meilleur est le pouvoir d'achat des citoyens et meilleures seront les opportunités d'achat de nouveaux véhicules ce qui favorisera la consommation du carburant.

5.2 Influence de la population utilisant le Transport en Commun (T.C) sur la Consommation du GPLc :

Le coefficient correspondant à la variable explicative POP_{tc} (la population utilisant le Transport en Commun) est positif (= + **0.0007059843269**), ce qui indique qu'une croissance de cette population aura un effet positif sur la consommation du GPLc. Cet effet s'expliquerait par le fait qu'une croissance de la population utilisant le TC implique systématiquement une augmentation du nombre de personnes utilisant les *TAXIS*, qui représentent 30% de la totalité des services de transport en commun (Source : Ministère des Transports).

Comme la tendance pour la plupart des *TAXIS* est à l'utilisation du GPLc, comme carburant, du fait de son faible coût par rapport aux autres carburants, ceci entraînera une augmentation de sa consommation.

5.3 Influence du taux de croissance du parc automobile (T_{cr}) sur la consommation du GPLc :

Le coefficient correspondant à la variable explicative T_{cr} (le taux de croissance du parc automobile) est positif (= + **571136.6743**), ceci signifie qu'une croissance de ce taux aura un effet positif sur la consommation du GPLc. En d'autres termes, une augmentation du taux de croissance du parc automobile implique que les nouveaux véhicules seraient plus susceptibles d'être convertis en GPLc que les anciens, ce qui ne pourra que renforcer la consommation du GPLc.

5.4 Influence de la taille de la population active sur la consommation du GPLc :

Le coefficient correspondant à la variable explicative POP_{act} (la taille de la population active) est positif (= + **0.02742276076**), ceci signifie qu'une croissance de cette population aura un effet positif sur la consommation du GPLc. Nous avons déjà évoqué précédemment que l'augmentation de la taille de la population active s'accompagnerait de celle de ses déplacements et partant d'une consommation accrue de carburant.

Par ailleurs, ce même phénomène pourrait s'expliquer par le fait que l'acquisition et l'utilisation d'un véhicule est conditionnée par le pouvoir d'achat détenu par la population active d'un marché de consommation.

5.5 Influence du prix de l'essence (P_{ess}) sur la consommation du GPLc :

Le coefficient correspondant à la variable explicative P_{ess} (le prix de l'essence) est positif (= + **3497.008951**), ce qui signifie qu'une augmentation du prix de l'essence aura un effet positif sur la consommation du GPLc. Cet effet s'explique par le fait que, s'agissant du même moteur et grâce à la technologie de la bicarburation, une augmentation du prix de l'essence aurait tendance à pousser les usagers à lui substituer le GPLc, d'où le renforcement de la consommation de ce dernier.

Conclusion :

Il ressort de ce chapitre que la consommation du GPLc est naturellement influencée par les facteurs suivants: le PIB, la taille de la population utilisant le transport en commun, le taux de croissance de la taille du parc automobile, la taille de la population active ainsi que par le prix de l'essence. Cette influence se manifeste positivement sur la consommation du GPLc, étant donné l'effet positif de chacun de ces facteurs sur cette consommation. On remarquera que le prix du GPLc ne ressort pas dans notre modèle comme étant une variable explicative importante pour la consommation des GPLc, car d'autres facteurs clés, non quantifiables, influençant le choix du consommateur, ont un impact significatif sur cette consommation. Parmi ces facteurs, on pourra évoquer l'aspect sécurisant des véhicules utilisant le GPLc, l'aspect réglementaire régissant le marché, le coût de la transformation, le coût d'un véhicule neuf fonctionnant au GPLc et les tendances du marché de consommation.

Une fois le modèle construit, il s'agira de faire des prévisions sur la variable endogène qui est la consommation du GPLc. Pour ce faire, on commencera par faire des prévisions sur les variables explicatives définies ci-dessus, les prévisions obtenues seront alors injectées dans l'équation qui régit le modèle obtenu.

***Chapitre III : La prévision de la
consommation du GPLc***

La prise de décision représente un moment crucial dans la vie des entreprises. Aussi, une attention de plus en plus importante est-elle accordée à l'amélioration de son processus. L'un des aspects fondamentaux de cette amélioration consiste à distinguer les différents éléments qui composent le processus de prise de décision et d'agir sur chacun d'eux afin de ramener au plus bas les risques d'erreur. Dans ce cadre, la prévision est identifiée comme l'un des éléments clés sur lequel repose ce processus.

Pour répondre aux diverses situations où des prévisions sont nécessaires, un certain nombre de méthodes ou techniques a été développé. Nous pouvons identifier des classes de méthodes de prévision suivant la *figure III.1*.

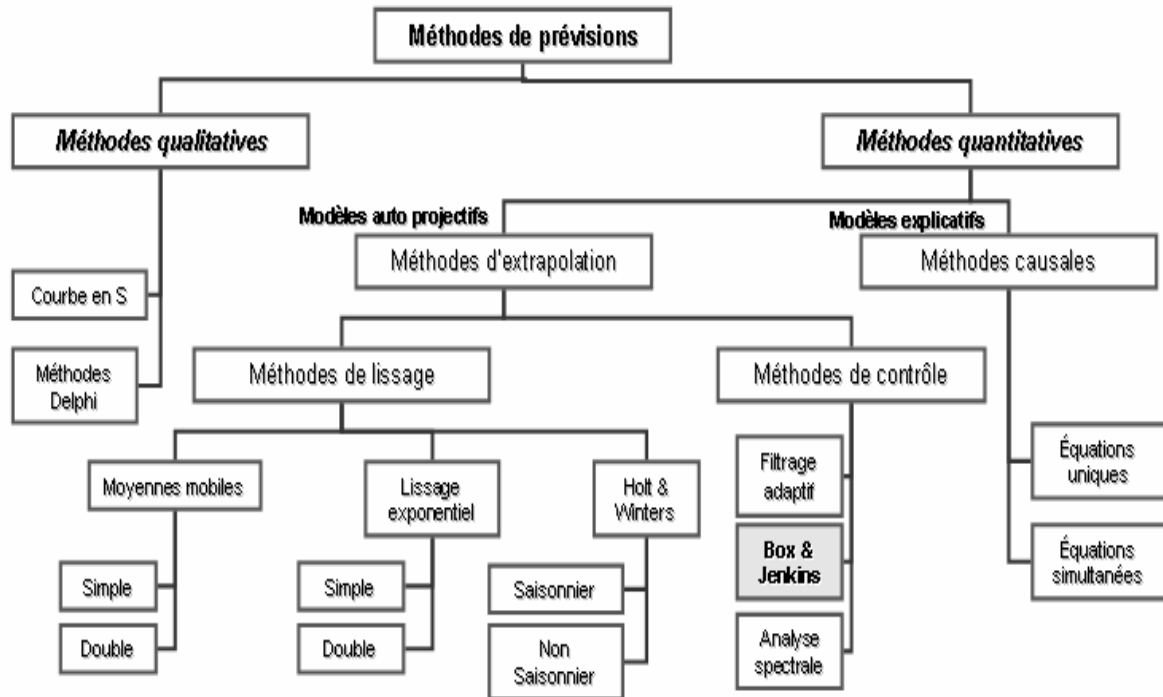


Figure III. 1 : Les méthodes de prévisions [BOU 02]

Le choix d'une méthode de prévision dépend de plusieurs facteurs. Ce choix doit s'opérer en cherchant à minimiser un coût, mais il faut également tenir compte du type de la chronique et de la finalité du traitement qui jouent souvent le rôle de contraintes : [CHA 03]

- Le coût de traitement qui lui-même dépend de deux facteurs : d'une part l'algorithme de calcul inhérent à la technique de prévision choisie (de ce point de vue, certaines techniques impliquent l'usage d'ordinateurs puissants) et, d'autre part, la longueur de l'historique qu'il faut conserver, ce qui se traduit par un coût de stockage de données et de traitement plus ou moins important.
- Le type de la série chronologique : plusieurs paramètres sont à prendre en compte dans le choix d'une technique : longueur de l'historique, existence ou non d'une saisonnalité, etc...
- La finalité du traitement : certaines techniques de prévision ne sont appropriées que pour le court terme et fournissent des résultats catastrophiques si l'on tente de les utiliser pour le moyen ou le long terme.

Cependant, le type de prévision rencontré le plus souvent dans la pratique se présente de la manière suivante:

Etant donné une série chronologique d'observations relevées sur plusieurs périodes, on voudrait établir une prévision pour une certaine période future. Toutes les méthodes destinées à ce type de situation admettent qu'il existe une loi fondamentale sous-jacente représentée par les données historiques qui, en outre, manifestent certains comportements aléatoires. L'activité centrale de la méthode de prévision consiste donc à isoler le mieux possible cette loi fondamentale, afin de l'utiliser comme base de projection dans le futur [DAU 05a]

Les situations concrètes sont beaucoup plus compliquées, car la loi résulte à la fois d'une tendance, d'un facteur saisonnier, d'un facteur cyclique ainsi que de fluctuations aléatoires. Dans ce genre de situations, on a besoin d'une méthode de prévision bien plus englobante.

La méthode de prévisions de Box- Jenkins est particulièrement bien adaptée au traitement des séries chronologiques complexes, et en général, aux situations dans lesquelles la loi de base n'est pas immédiatement apparente [CHA 03]

La méthode de Box- Jenkins présente un avantage certain face aux autres méthodes quantitatives, utilisées pour la prévision, grâce aux facteurs suivants : [BOU 02]

- Elle est applicable lorsque la série présente une saisonnalité et/ ou une tendance.
- Elle permet de déterminer un modèle adéquat dans la classe des modèles *ARIMA* Ces derniers permettent de représenter la plupart des processus stationnaires et se prêtent tout particulièrement aux prévisions à moyen terme.
- Le coût d'élaboration de ces modèles est relativement faible. De nos jours plusieurs logiciels prennent en charge ce type de modèles.

Tenant compte de ces facteurs, notre choix s'est porté sur la méthodologie de Box- Jenkins pour le calcul de la prévision de la consommation nationale du GPLc. Aussi, nous allons, articuler le présent chapitre autour de trois sections.

La première servira à l'énoncé des définitions de base et des différents éléments d'analyse des séries chronologiques qui nous permettront de comprendre la méthodologie de Box & Jenkins, dont l'objectif est de systématiser la démarche d'analyse des séries temporelles. La deuxième sera consacrée à la description d'une manière générale de la méthodologie de Box and Jenkins. Enfin, la troisième section prendra en charge le calcul, proprement dit, des prévisions de chacune des variables explicatives, retenues dans le chapitre précédent, en appliquant les étapes de cette méthodologie. Ceci fait, les prévisions retenues pour ces variables seront injectées dans le modèle explicatif obtenu précédemment.

.

Section I : Eléments d'analyse des séries chronologiques

L'étude des séries chronologiques ou temporelles correspond à l'analyse statistique d'observations régulièrement espacées dans le temps, dans le but de représenter des phénomènes aléatoires qui évoluent dans le temps. Le modèle obtenu sera par la suite utilisé, selon les objectifs recherchés comme la prévision ou le contrôle.

Sur ce fait, cette section est consacrée à une présentation sommaire des techniques d'analyse des séries chronologiques. Nous allons d'abord donner quelques définitions de base qui nous permettraient par la suite de présenter les différentes classes de modèle susceptibles de modéliser une série chronologique (AR, MA, ARMA). Nous étudierons par la suite les caractéristiques statistiques en terme de stationnarité des séries temporelles en présentant les différents tests, le test de Dickey- Fuller entre autres.

I. Définitions et concepts de base :

1. Série chronologique : [BOU 06]

Une série chronologique ou encore chronique est un ensemble d'observations d'un processus aléatoire $\{X_t, t \in T\}$ se réalisant en un instant spécifié $t \in T$.

- Une série chronologique est dite discrète si l'ensemble des instants d'observations est discret (dénombrable)
- Une série chronologique est dite continue si l'ensemble des instants d'observations est continu (non dénombrable)

2. Stationnarité : [UPM 04] [JON 99]

Avant le traitement d'une série chronologique, il convient d'en étudier les caractéristiques stochastiques. Si ses caractéristiques, c'est-à-dire son espérance et sa variance se trouvent modifiées dans le temps, la série chronologique est considérée comme non stationnaire.

Dans le cas d'un processus stochastique invariant, la série temporelle est alors stationnaire.

- *Processus faiblement stationnaire (stationnarité 2ème ordre) :*

Un processus $\{X_t, t \in T\}$ est dit stationnaire au second ordre, ou stationnaire au sens faible, ou stationnaire d'ordre deux si les trois conditions suivantes sont satisfaites :

- $E(X_t^2) < \infty, t \in T,$
- $E(X_t) = m,$ indépendant de $t.$ ($t \in T$)
- $(t, h) \in T^2, Cov(X_t, X_{t+h}) = E[(X_{t+h} - m)(X_t - m)] = \gamma(h)$ indépendant de $t.$

- ❖ La première condition $E(X_t^2) < \infty$ garantit tout simplement l'existence (ou la convergence des moments d'ordre deux.)
- ❖ La seconde condition $E(X_t) = m$ porte sur les moments d'ordre un et signifie tout simplement que les variables aléatoires X_t doivent avoir la même espérance quelle que soit la date t . Autrement dit, l'espérance du processus X_t doit être indépendante du temps.
- ❖ Enfin, la troisième condition, $\gamma(h)$ indépendant de t , porte sur les moments d'ordre deux résumés par la fonction d'autocovariance. Cette condition implique que ces moments doivent être indépendants de la date considérée et ne doivent dépendre

uniquement que de l'ordre des retards. Autrement dit la fonction d'autocovariance du processus X_t doit être indépendante du temps.

3. Fonction d'autocorrélation simple et fonction d'autocorrélation partielle

3.1 Fonction d'autocorrélation simple : [BOU 02] [JON 99]

La fonction d'autocorrélation d'un processus $\{X_t, t \in T\}$ de moyenne $E(X_t) = m$, noté ρ_k

Où ρ_k est définie par : $\rho(k) = \rho_k = \frac{\gamma(k)}{\gamma(0)}$ quelque soit $k \in T$

Elle mesure la corrélation de la série avec elle-même décalée de k périodes :

On a :

- $\rho(k) \in [-1, 1]$ et $\rho(0) = 1$
- $\gamma(k)$ désigne la fonction d'autocovariance tel que : $\gamma(k) = E[(X_t - m)(X_{t-k} - m)]$
- $\gamma(0)$ désigne la fonction de variance.
- Les fonctions $\rho(k)$ et $\gamma(k)$ sont symétriques : $\rho(k) = \rho(-k)$ et $\gamma(k) = \gamma(-k)$

Remarque :

La représentation graphique de $\rho(k)$ est appelée : **corrélogramme**.

3.2 Fonction d'autocorrélation partielle : [BOU 02]

Nous pouvons définir la fonction d'autocorrélation partielle (*FAP*) de retard k comme étant le coefficient de corrélation partielle entre Y_t et Y_{t-k} , c'est-à-dire comme étant la corrélation entre Y_t et Y_{t-k} , l'influence des autres variables décalées de k périodes ($Y_{t-1}, Y_{t-2}, \dots, Y_{t-k+1}$) ayant été retirée.

3.3 Les opérateurs linéaires : [BOU 02]

A. L'opérateur Retard :

On aura souvent à considérer une variable en fonction de son passé. Il est donc commode de définir un opérateur qui transforme une variable X_t en sa valeur passée. C'est l'opérateur retard désigné par la lettre B et tel que: $B(X_t) = X_{t-1}$ et $B^k(X_t) = X_{t-k}$

B. L'opérateur d'avance (forward) :

$$FX_t = X_{t+1} ; F^n X_t = X_{t+n}$$

C. L'opérateur de différence ordinaire :

L'opérateur de différence ordinaire noté ∇ associé au processus $\{X_t, t \in T\}$ tel que: $\forall t \in T$

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

Et par construction, nous obtiendrons l'opérateur de la $d^{\text{ième}}$ différence noté ∇^d tel que :

$$\nabla^d X_t = (1 - B)^d X_t$$

D. L'opérateur de différence saisonnier :

L'opérateur de différence saisonnier d'ordre s , noté : ∇_s associé au processus $\{X_t, t \in T\}$ tel que $\forall t \in T : \nabla_s X_t = (1 - B^s) X_t$

Et par construction, nous obtiendrons l'opérateur de la $d^{\text{ième}}$ différence d'ordre s , noté : ∇_s^d

Tel que : $\forall t \in T : \nabla_s^d X_t = (1 - B^s)^d X_t$

II. Modélisation des séries chronologiques :

L'objectif de la modélisation est de construire des modèles permettant de décrire le comportement d'une chronique, et de ce fait résoudre les problèmes liés à la prévision.

Nous allons présenter une famille de processus aléatoires qui sont censés recouvrir une gamme très large d'évolution possible de séries chronologiques.

1. Modèle autorégressif AR (p) :

Formulation : [BOU 02] [JON 99]

Dans le processus autorégressif d'ordre p , l'observation Y_t est générée par une moyenne pondérée des observations passées jusqu'à la $p^{\text{ième}}$ période sous la forme suivante :

$$AR(1) : X_t = \Phi_1 X_{t-1} + \varepsilon_t$$

$$AR(2) : X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \varepsilon_t$$

⋮

$$AR(p) : X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \dots + \Phi_p X_{t-p} + \varepsilon_t$$

Donc, un processus autorégressif d'ordre p s'écrit sous la forme suivante :

$$X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \dots + \Phi_p X_{t-p} + \varepsilon_t$$

Ou encore : $X_t - \Phi_1 X_{t-1} - \Phi_2 X_{t-2} - \dots - \Phi_p X_{t-p} = \varepsilon_t$

Où :

- Les Φ_i sont des réels positifs ou négatifs à déterminer.
- ε_t est un bruit blanc de variance σ^2 .

Remarque :

Ce modèle apparaît comme un modèle de régression multiple où l'on explique les valeurs de la série chronologique aux instants : $t-1, t-2, \dots, t-p$, et c'est la raison pour laquelle nous l'appelons autorégressif d'ordre p .

Notion de causalité (stationnarité) : [BOU 02]

Définition :

Un modèle de série chronologique de la forme :

$$X_t = g(X_{t-1}, X_{t-2}, \dots, X_{t-p}; \varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q})$$

(Où ε_t est un bruit blanc) est dit causal si et seulement si on peut examiner le processus stochastique X_t sous forme d'une combinaison linéaire (finie ou infinie) convergente du présent et du passé d'un bruit blanc ε_t

Théorème :

Le modèle autorégressif est causal si et seulement si les racines de la fonction caractéristique $\Phi(B) = 0$ sont en valeur absolue supérieures à 1, c'est-à-dire : $|z| > 1$.

Notion d'inversibilité : [BOU 02]

Définition :

Un modèle de série chronologique de la forme :

$$X_t = g(X_{t-1}, X_{t-2}, \dots, X_{t-p}; \varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q})$$

(Où ε_t est un bruit blanc) est dit inversible si et seulement si on peut examiner le processus stochastique sous la forme d'une combinaison linéaire (finie ou infinie) convergente, du présent et du passé de la variable aléatoire X_t .

D'après la définition d'inversibilité, un modèle autorégressif d'ordre fini AR (p) est toujours inversible.

2. Modèle MA (q) (moving average : moyenne mobile) :

Formulation : [BOU 02], [JON 99]

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

Un processus stationnaire $\{X_t, t \in T\}$ satisfait une représentation moyenne mobile d'ordre q, noté MA (q), s'il est solution de l'équation aux différences stochastiques suivante :

$$X_t = \varepsilon_t - \sum_{j=1}^q \theta_j \varepsilon_{t-j} \quad \text{Où : } \begin{array}{l} \text{les } \theta_j \text{ sont des réels positifs ou négatifs à déterminer} \\ \varepsilon_t \text{ est un bruit blanc de variance } \sigma^2 \end{array}$$

En introduisant l'opérateur au retard on obtient :

$$X_t = \Theta(B)\varepsilon_t \quad \text{où } \Theta(B) = -\sum_{j=0}^q \theta_j B^j ; \theta_0 = -1 ; \theta_j \in \mathfrak{R}, \forall j \in \mathfrak{R} \text{ et } \forall j < q ; \theta_q \in \mathfrak{R}^*$$

Avec :

$\Theta(B)$: Le polynôme de moyenne mobile de degré q (q représente le plus grand retard utilisé)

Dans ce processus, tout comme dans le modèle auto régressif AR, les aléas sont supposés être engendrés par un processus de type bruit blanc.

Nous pouvons interpréter le modèle MA comme étant représentatif d'une série chronologique fluctuant autour de sa moyenne de manière aléatoire d'où le terme 'moyenne mobile'.

Il est à noter qu'il y a équivalence entre un processus MA (1) et un processus AR d'ordre p infini : **MA (1) = AR (∞).**

Notion de causalité : [BOU 02]

D'après la définition de causalité, un modèle moyenne mobile d'ordre fini est toujours causal car c'est une combinaison linéaire finie du processus stationnaire $\{\varepsilon_t, t \in T\}$

Notion d'inversibilité : [BOU 02]

Le modèle de moyenne mobile MA (q) : $X_t = \varepsilon_t - \sum_{j=1}^q \theta_j \varepsilon_{t-j}$ est inversible si et seulement si les racines de l'équation : $\Theta(B) = 0$ sont en valeur absolue supérieures à 1, c'est-à-dire : $|z| > 1$.

3. Modèle mixte ARMA (mélange du processus AR et MA) :

3.1 Formulation : [BOU 02], [FIO 99]

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

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

Où encore :

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

Avec :

- $\Phi_p \neq 0, \theta_q \neq 0$.
- Les polynômes Φ et Θ d'ordre respectif p et q n'ont pas de racines communes
- ε_t est à bruit blanc.

Donc on peut noter AR (p) par ARMA (p, 0) et MA (q) par ARMA (0, q)

3.2 Conditions de causalité : [BOU 02]

Le modèle ARMA (p, q) défini par : $\Phi(B) X_t = \Theta(B) \varepsilon_t$ est causal si et seulement si les racines du polynôme Φ sont de modules strictement supérieures à 1 c'est-à-dire : $|z| > 1$

3.3 Condition d'inversibilité : [BOU 02]

Le modèle ARMA (p, q) défini par : $\Phi(B) X_t = \Theta(B) \varepsilon_t$ est inversible si et seulement si les racines du polynôme Θ sont en valeur absolue supérieures à 1 c'est-à-dire $|z| > 1$

III. Séries non stationnaires :

Les séries économiques présentent souvent des réalisations non stationnaires. Elle admettent une moyenne et une variance qui varient au cours du temps. La non stationnarité peut être détectée graphiquement par la présence d'une *tendance*, d'une *saisonnalité*, ou d'une modification dans la structure de la série, c'est pourquoi, on analysera les graphes de la fonction d'autocorrélation.

Mais cela reste insuffisant, et pour avoir une certitude, il est nécessaire d'appliquer les tests pour confirmer ou infirmer les déductions de l'analyse des graphes de la série et des corrélogrammes.

La plupart des résultats et des méthodes utilisées dans l'analyse des séries temporelles repose sur la notion de stationnarité du second ordre, ce qui nous mène à appliquer à la chronique non stationnaire certaines transformations (différence ordinaire, différence saisonnière, ...) que nous allons présenter ci-dessous :

1. Analyse de la tendance :

La non stationnarité d'un processus aléatoire dépend du moment du premier ordre (espérance mathématique) et de celui du second ordre (variance et covariance du processus), qui doivent être indépendants du temps. Dans le cas contraire, le processus aléatoire est dit non stationnaire. [BOU 06]

On analyse les cas de non stationnarité à partir de deux types de processus :

- Les processus de type TS qui représentent la non stationnarité de type *déterministe*.
- Les processus de type DS qui représentent la non stationnarité de type *aléatoire (stochastique)*

1.1 Le processus de type TS (Trend Stationary) : [BOU 02]

Les processus TS sont des processus non stationnaires de type déterministe qui s'écrivent sous la forme : $X_t = f_t + \varepsilon_t$

Où :

- f_t : fonction polynomiale qui dépend du temps qui peut être linéaire ou non linéaire.
- ε_t : processus stationnaire.

Les étapes de la méthode utilisée pour le rendre stationnaire sont :

- Estimer les coefficients \hat{a}_0 et \hat{a}_1 par la méthode des moindres carrés ordinaires (MCO).
- Retrancher de X_t la valeur estimée $\hat{a}_0 + \hat{a}_1 t$

1.2 Le processus de type DS (Differency Stationary) : [BOU 02]

Les processus DS sont des processus que l'on peut rendre stationnaires en utilisant le filtre aux différences d'ordre d (d supérieur ou égal à 1) il s'écrit alors : $(1 - B)^d X_t = \beta + \varepsilon_t$

Où :

- B : opérateur décalage.
- d : l'ordre du filtre aux différences.
- ε_t : processus stationnaire.

Ces processus sont souvent représentés en utilisant le filtre aux différences premières ($d = 1$). Le processus est dit alors processus du 1^{er} ordre.

Il s'écrit: $(1 - B)X_t = \beta + \varepsilon_t \Leftrightarrow X_t = X_{t-1} + \beta + \varepsilon_t$

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

- $B = 0$:

Le processus DS est dit sans dérive, il s'écrit alors : $X_t = X_{t-1} + \varepsilon_t$

Ce processus porte le nom de modèle de marche au hasard ou de marche aléatoire.

- $B \neq 0$:

Le processus porte alors le nom de processus DS avec dérive, il s'écrit alors :

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

2. Test de racine unitaire : [BOU 02], [UPM 04]

Ces tests permettent de rechercher l'existence d'une racine unitaire dans les processus qui ont généré la série chronologique en permettant de déterminer le type de la non stationnarité qui peut être déterministe (de type TS) que l'on peut le rendre stationnaire par régression sur une tendance déterministe, ou stochastique (de type DS) que l'on peut les rendre stationnaires en utilisant un filtre au différence à l'ordre d

Il existe un grand nombre de tests de racine unitaire, les tests les plus utilisés sont ceux de Dicky-Fuller (1979) (cf annexe 4)

VI. Extension du modèle ARMA : (Le modèle ARIMA (p, d, q))

[CHA 03], [BOU 02]

Il convient de stationnariser une série par passage aux différences selon l'ordre d'intégration d (c'est-à-dire le nombre de fois qu'il faut différencier la série pour la rendre stationnaire).

Il est donc naturel de considérer la classe des processus dont la différence d'un certain ordre satisfait une représentation ARMA.

Si on note $\nabla^d X_t$ la différence d'ordre d de X_t c'est-à-dire : $\nabla^d X_t = (1 - B)^d X_t$

On va s'intéresser au processus X_t satisfaisant : $\Phi(B)\nabla^d X_t = \Theta(B)\varepsilon_t$

Un processus X_t est un modèle « autorégressif moyenne mobile intégré » d'ordre (p, d, q) ARIMA (p, d, q) s'il vérifie une équation de type : $\Phi(B)(1 - B)^d X_t = \Theta(B)\varepsilon_t$; pour tout $t \geq 0$

Avec :

- $\Phi(B)$: Polynôme autorégressif dont les racines sont de module supérieur à 1
 $\Phi(B) = 1 - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B$ où $\Phi_p \neq 0$
- $\Theta(B)$: Polynôme moyenne mobile dont les racines sont de module supérieurs à 1
 $\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B$ où $\theta_q \neq 0$
- Ces deux polynômes n'ont aucune racines communes.
- Les coefficients réels : $\Phi_i, i = 1, \dots, p$ et $\theta_j, j = 1, \dots, q$ sont fixés.
- ε_t : processus stationnaire

Remarque :

La famille ARIMA désigne parfois la classe de tous les modèles stationnaires et non stationnaires, en couvenant que les ARIMA (p, 0, q) sont des ARMA (p, q)

Après avoir défini les éléments de base de l'analyse des séries chronologiques, nous allons à présente donner un aperçu général des étapes de la méthodologie de Box and Jenkins.

Section II : La méthodologie de Box & Jenkins:

La méthodologie de Box and Jenkins (1976) est une étude systématique des séries chronologiques, à partir de leurs caractéristiques, afin de déterminer dans la famille des modèles *ARIMA*, le modèle le plus adapté à représenter le phénomène étudié.

Cette démarche comporte plusieurs étapes qui sont généralement répétées, autant de fois que nécessaire, ces étapes sont :

- La spécification (identification) du problème ;
- L'estimation des paramètres ;
- La validation du modèle ;
- Le choix du modèle ;
- Et la prévision.

I. L'identification du modèle : [BOU 02], [FIO 99]

Dans cette phase, il s'agit avant tout de vérifier la stationnarité de la série chronologique, ceci s'effectue essentiellement à partir de l'étude de sa représentation graphique et de l'étude des autocorrélations (les corrélogrammes).

Si l'étude des graphiques et du corrélogramme simple de la série présage que la série est affectée d'une tendance, il convient d'en étudier les caractéristiques selon le test de Dickey Fuller pour détecter de quel processus non stationnaire il s'agit puis de le rendre stationnaire par la suite. La méthode d'élimination de la tendance est fonction du processus DS ou TS sous jacent à la chronique étudiée.

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

C'est la phase la plus importante et la plus difficile, elle consiste essentiellement à déterminer le modèle le plus adéquat dans la famille des modèles *ARIMA*. Elle est fondée sur l'étude des corrélogrammes simple et partiel.

Nous pouvons essayer d'édicter quelques règles simples facilitant la recherche des paramètres p , d , q du modèle *ARIMA* :

- Si le corrélogramme simple n'a que ses q premiers termes significativement différents de zéro et que les termes du corrélogramme partiel diminuent lentement, nous pouvons pronostiquer un *MA* (q).
- Si le corrélogramme partiel n'a que ses p premiers termes significativement différents de zéro et que les termes du corrélogramme simple diminuent lentement, cela caractérise un *AR* (p).
- Si les fonctions d'autocorrélation simple et partielle n'apparaissent pas tronquées, alors il s'agit d'un processus *ARMA* (p , q) dont les paramètres dépendent de la forme particulière des corrélogrammes.

II. Estimation des paramètres: [BOU 02]

Après la phase d'identification, vient la phase de l'estimation des paramètres du modèle *ARIMA*. C'est la phase la plus « mécanique » de la démarche, où il s'agit d'estimer les paramètres du modèle sélectionné à l'étape de l'identification.

Dans le cas d'un modèle *AR* (p), nous pouvons appliquer la méthode des moindres carrés. Par contre l'estimation d'un modèle *MA* (q) s'avère plus complexe, Box & Jenkins suggèrent d'utiliser une procédure itérative de type balayage.

Nous pouvons aussi mentionner des méthodes d'estimation fondées sur une maximisation de fonction de vraisemblance.

De nos jours, la plupart des logiciels fournissent les valeurs estimées des paramètres du modèle.

III. Adéquation du modèle: [CHA 03], [BOU 02]

Cette partie consiste à examiner les résultats de l'estimation du modèle, et cela en faisant subir au modèle, des tests qui permettent la validation ou le rejet de ce dernier.

1. Test sur les paramètres du modèle :

Il s'agit de vérifier que :

- Tous les paramètres du modèle sont significativement différents de zéro c'est-à-dire :
 - La valeur absolue de la statistique de Student « *t- statistique* » doit être supérieure à 1.96 et 1.64 (valeurs critiques de loi normale centrée réduite) au seuil 5% et 10% respectivement
 - S'il y a des paramètres non significativement différents de zéro, il convient d'envisager une nouvelle spécification en éliminant l'ordre de *AR* et *MA* non valide, et ainsi de suite jusqu'à l'obtention d'un modèle valide.
- Les conditions de causalité et d'inversibilité (de stationnarité) sont vérifiées, en pratique on peut vérifier cela en s'assurant que les racines du modèle sont toutes en module supérieures à 1.

2. Test sur les résidus :

Il s'agit de vérifier la blancheur ainsi que la normalité des résidus par : les tests de Portemanteau, le test de Jarque & Bera, test de Skewness et de Kurtosis cités dans le chapitre II.

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

VI. Choix du modèle : [BOU 02], [FIO 99]

Dans la construction des séries temporelles, on estime souvent plusieurs spécifications. Il peut arriver que deux ou plusieurs modèles soient adéquats, dans ce cas, des tests supplémentaires devraient être utilisés pour déterminer la meilleure spécification, c'est ainsi que nous introduisons les critères de pouvoir prédictifs portant sur la qualité de l'information et du modèle.

1. Critère de pouvoir prédictif

Ces critères se basent sur l'erreur de prévision, et proposent des modèles qui conduisent à une erreur de prévision suffisamment petite.

Plusieurs critères sont proposés, à savoir :

1 - Le coefficient de détermination R^2 , correspond à une normalisation de la variance tel que :

$$R^2 = 1 - \frac{\hat{\sigma}^2}{\text{Var}(x)}$$

2- Le coefficient de détermination modifié \bar{R}^2 , tel que : $\bar{R}^2 = 1 - \frac{\hat{\sigma}^2/(n-p-q)}{\text{Var}(x)/(n-1)}$

Il s'agit de choisir le modèle fournissant les valeurs R^2 et \bar{R}^2 maximales

2. Critère d'information :

C'est une approche qui consiste à supposer que les modèles ARMA (p, q) fournissent des approximations de la réalité et que la vraie loi inconnue des observations ne satisfait pas forcément un tel modèle.

On peut fonder le choix du modèle sur une mesure de l'écart entre la vraie loi inconnue et le modèle proposé.

Plusieurs estimateurs de la quantité d'information ont été proposés :

- L'estimateur introduit par AKAIKE : $AIC(p, q) = \log \hat{\sigma}^2 + \frac{2(p+q)}{n}$
- L'estimateur introduit par Schwarz : $SC(p, q) = \log \hat{\sigma}^2 + (p+q) \frac{\log n}{n}$

Le modèle choisi est celui qui correspond à la valeur minimale de ces deux estimateurs.

Lorsque le modèle est validé, la prévision peut alors être calculée.

V. La prévision :

On entend par prévision une estimation de ce qui va arriver. Il existe trois types de prévisions selon l'intervalle de temps qui sépare le moment où est effectuée la prévision (l'origine de la prévision), et le moment pour lequel elle est souhaitée (l'horizon de la prévision) : prévision à court terme, moyen terme et long terme.

L'objectif de la méthode de Box & Jenkins est de réaliser des prévisions. Une fois le modèle ARIMA (p, d, q) choisi, estimé et validé pour les observations X_1, \dots, X_t , on calcule les prévisions.

Nous pouvons résumer les différentes étapes de la méthodologie de Box and Jenkins à travers le schéma suivant :

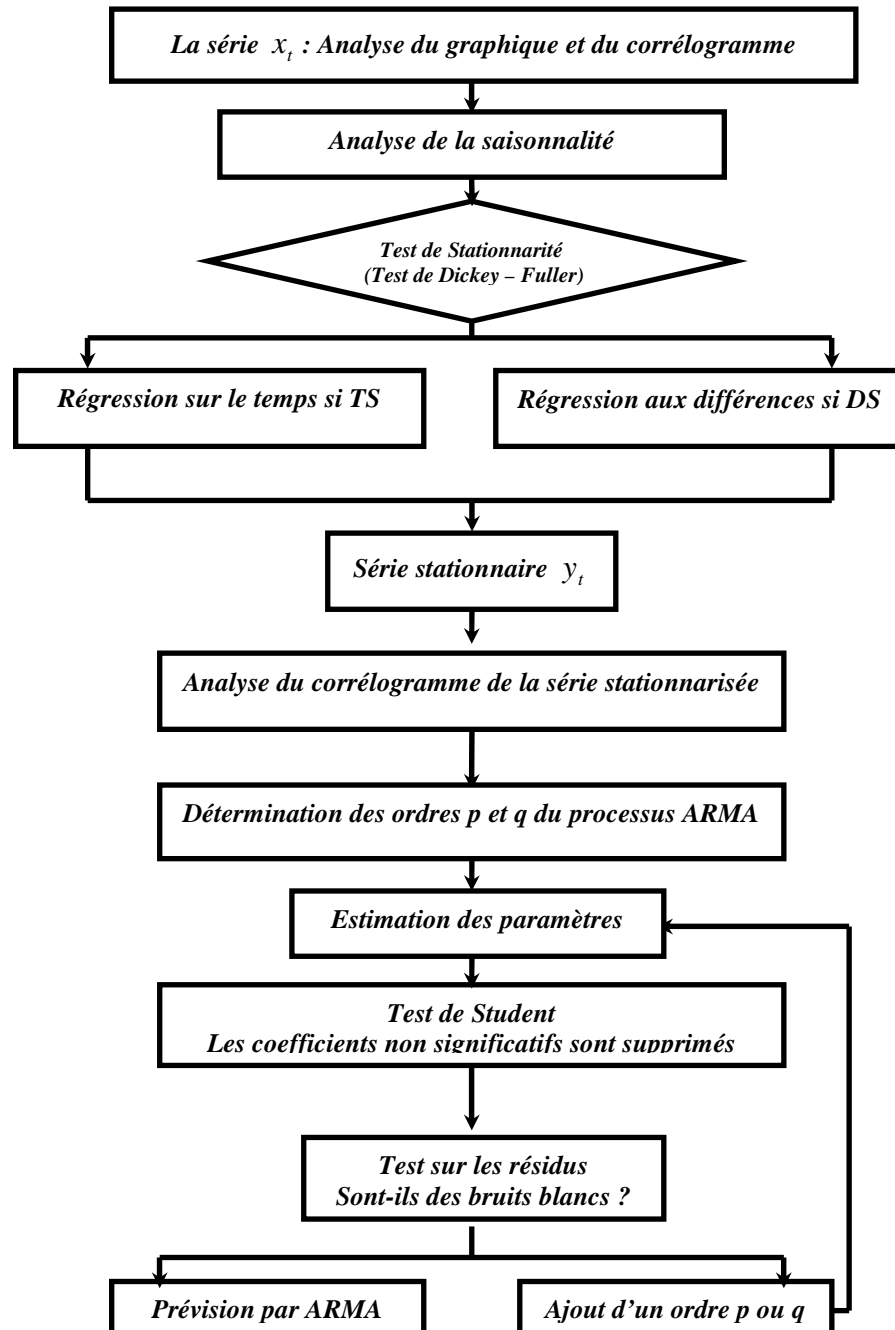


Figure III. 2 : Les étapes de la démarche prévisionnelle de Box & Jenkins

Section III: Mise en œuvre de la méthode de Box & Jenkins pour la prévision de la consommation du GPL c

Nous proposons dans ce qui suit, d'appliquer la méthodologie de Box & Jenkins pour chacune des variables explicatives retenue pour expliquer la consommation du GPLc dans le but d'aboutir à un modèle de prévision.

I. Mise en œuvre de la méthodologie Box & Jenkins sur les séries:

La méthodologie Box & Jenkins va être appliquée sur chacune des chroniques suivantes :

- La taille de la population utilisant le transport en commun POP_{tc} ;
- La taille de la population active POP_{act} ;
- Le PIB ;
- Le taux de croissance de la taille du parc automobile T_{cr} ;
- Le prix de l'essence P_{ess}

1. Etude de la série « la taille de la population utilisant le Transport en Commun » POP_{tc} :

1.1 Etude de la série brute POP_{tc} :

La série POP_{tc} représente l'évolution annuelle de la taille de la population qui utilise le transport en commun sur une période allant de 1980 à 2006.

a) Analyse préliminaire de la série POP_{tc} :

Pour avoir une idée globale mais non décisive sur la nature et les caractéristiques de la série (tendance, saisonnalité, etc...), il est nécessaire d'analyser le graphe représentant l'évolution de la série en question

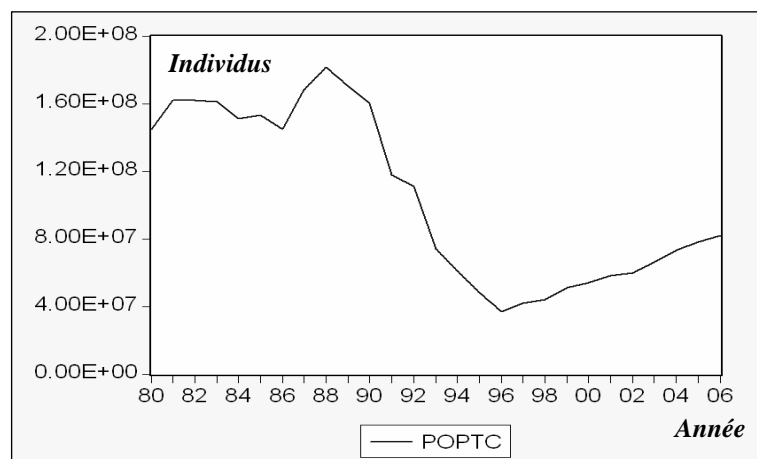


Figure III. 3 : Représentation graphique de la série POP_{tc}

La représentation graphique de la série POP_{tc} (figure III.3) nous montre que la série semble être affectée d'une tendance, il semble donc que la série POP_{tc} soit non stationnaire.

b) Examen du corrélogramme de la série POP_{tc}:

L'analyse du corrélogramme simple et partiel de la série brute (*figure III.4*) nous indique une non stationnarité de cette dernière.

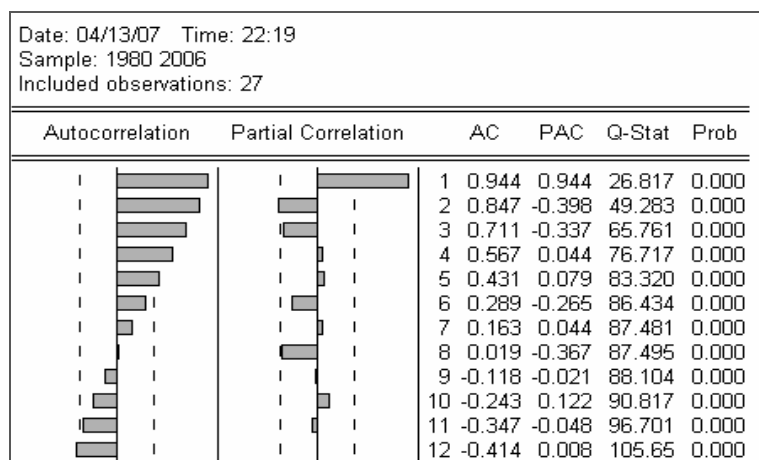


Figure III. 4 : Corrélogramme de la série la série POP_{tc}

En effet, les fonctions d'autocorrélation simple et partielle tendent lentement vers zéro, nous remarquons aussi de nombreux pics sortant de la l'intervalle de confiance, ceci est caractéristique d'une série non stationnaire.

Il devient donc nécessaire de vérifier la non stationnarité de la série POP_{tc} en appliquant le test de Dickey- Fuller et d'identifier la nature du processus non stationnaire puis de le rendre stationnaire.

C. Test de Dickey-Fuller sur la série POP_{tc}:

Le test de Dickey- Fuller permet de mettre en évidence le caractère stationnaire ou non d'une chronique par la détermination d'une tendance [BOU 02]

Nous allons appliquer le Test de Dickey – Fuller sur la série POP_{tc}.

Pour ce faire, nous utilisons le test programmé par le logiciel *EViews*. Le résultat est donné dans le *tableau III.1*

Tableau III. 1 : Test de Dickey-Fuller sur la POP_{tc}

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	1.38	2.79	$1.38 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	1.55	2.54	$1.55 \leq 2.54$	Constante non significativement différente de zéro
<i>Modèle 4</i>	-1.35	-1.95	$-1.35 \geq -1.95$	Série non stationnaire de type DS

Commentaires :

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que la statistique de Student (t- statistic = 1.38) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.

La tendance est donc non significativement différente de zéro, nous passons alors à l'estimation du modèle 5 pour tester la significativité de la constante.

- **Pour le Modèle 5 :** *modèle avec constante sans tendance déterministe*

Nous remarquons que la statistique de Student (t- statistic = 1.55) est inférieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.

La constante est donc non significativement différente de zéro, nous passons alors à l'estimation du modèle 4 pour vérifier l'existence ou non d'une racine unitaire.

- **Pour le Modèle 4 :** *modèle sans constante, ni tendance*

Nous testons l'hypothèse nulle H_0 : « Existence d'une racine unitaire » :

La statistique de Student (t- statistique = -1.35) est supérieure à la valeur critique lue dans la table de Dickey- Fuller (-1.95) au seuil 5%, on accepte alors l'hypothèse de non stationnarité H_0 . Il existe donc au moins une racine unitaire, la série POP_{tc} est donc non stationnaire de type : DS.

La meilleure façon de la rendre stationnaire est de la différencier une fois. La nouvelle série sera notée : $DPOP_{tc}$.

1.2 Etude de la série différenciée $DPOP_{tc}$:

a) Effet de la différenciation :

La représentation graphique de la série différenciée $DPOP_{tc}$ est donnée par la *figure III.5*

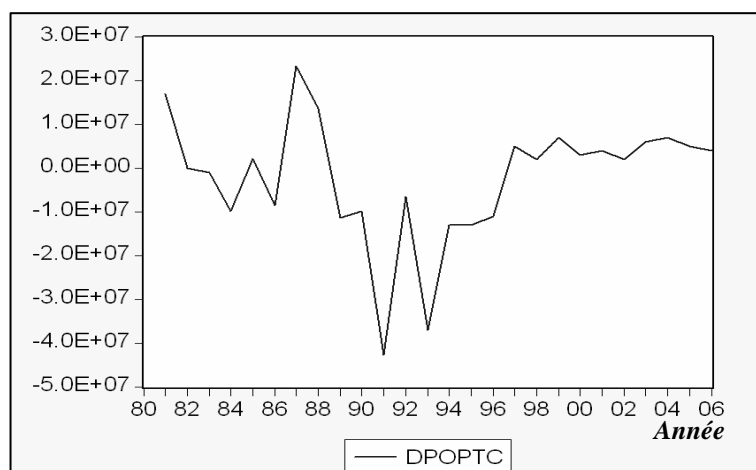


Figure III. 5 : Représentation graphique de $DPOP_{tc}$

Nous remarquons que l'effet de la tendance a été absorbé, la série semble être stationnaire, nous allons vérifier cela en appliquant le test de Dickey- Fuller de nouveau.

b) Test de Dickey- Fuller sur la série « DPOP_{tc} »:

Le test de Dickey Fuller appliqué à la série DPOP_{tc} a donné les résultats suivants (tableau III.2) :

Tableau III. 2 : Test de Dickey Fuller sur DPOP_{tc}

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	0.40	2.79	$0.40 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	0.52	2.54	$0.52 \leq 2.54$	Constante non significativement différente de zéro
<i>Modèle 4</i>	-1.73	-1.95	$-1.73 \geq -1.95$	Série non stationnaire de type DS

Commentaires :

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que la statistique de Student (t- statistic = 0.40) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.
Donc la tendance est non significativement différente de zéro

Nous passons alors à l'estimation du modèle 5 pour tester la significativité de la constante

- **Pour le Modèle 5 :** *modèle avec constante, sans tendance*

Nous remarquons que la statistique de Student (t- statistic = 0.52) est inférieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.
Donc la constante est non significativement différente de zéro.

Nous passons alors à l'estimation du modèle 4 pour tester l'hypothèse nulle H₀: « Existence d'une racine unitaire ».

- **Pour le Modèle 4 :** *modèle sans constante, ni tendance*

La statistique de Student (t- statistic = -1.73) est supérieure à la valeur critique (-1.95) lue sur la table de Dickey-Fuller au seuil 5%. Nous acceptons l'hypothèse de non stationnarité H₀. Il existe au moins une racine unitaire et donc la série DPOP_{tc} est non stationnaire de type DS.

La meilleure façon de la rendre stationnaire est de la différencier. Et la nouvelle série sera notée : DDPOP_{tc}.

1.3 Etude de la série brute différenciée deux fois DDPOP_{tc} :

a) Effet de la différenciation :

La représentation graphique de la série $DDPOP_{tc}$ est donnée par la *figure III.6*

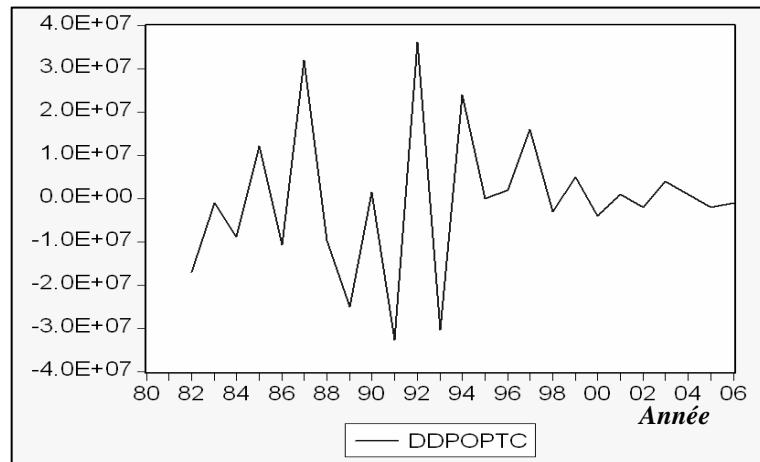


Figure III. 6 : Représentation graphique de la série $DDPOP_{tc}$

Nous remarquons bien que l'effet de la tendance a été absorbé, et que la série semble être stationnaire, nous allons appliquer le test de Dickey -Fuller pour vérifier cela.

b) Test de Dickey- Fuller sur la série $DDPOP_{tc}$:

Nous effectuons les mêmes étapes que celles de la série POP_{tc} et de la série $DPOP_{tc}$. Les résultats du test sont donnés par le *tableau III.3*

Tableau III. 3 : Test de Dickey Fuller sur $DDPOP_{tc}$

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	0.69	2.79	$0.69 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	0.042	2.54	$0.042 \leq 2.54$	Constante non significativement différente de zéro
<i>Modèle 4</i>	-9.52	-1.95	$-9.52 \leq -1.955681$	Série stationnaire

Commentaires :

- **Pour le Modèle 6 :** *modèle avec constante et tendance*

Nous remarquons que la statistique de Student (t- statistic = 0.69) est inférieure à la valeur critique (2.79) lue dans la tables de Dickey- Fuller au seuil 5%.
Donc la tendance est non significativement différente de zéro

Nous passons à l'estimation du modèle 5 pour tester la significativité de la constante

- **Pour le Modèle 5 :** *modèle avec constante sans tendance*

Nous remarquons que la statistique de Student (t- statistic = 0.042) est inférieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.

La constante est donc non significativement différente de zéro, nous passons alors à l'estimation du modèle 4 pour tester l'hypothèse H_0 : « existence d'une racine unitaire »

- **Pour le Modèle 4 : modèle sans constante, ni tendance**

La statistique de Student (t-statistique= -9.52) est inférieure à la valeur critique lue dans la table de Dickey-Fuller (-1.95) au seuil 5%,

D'après les résultats, nous rejetons l'hypothèse d'existence d'une racine unitaire H_0 , et nous acceptons l'hypothèse de stationnarité. La série $DDPOP_{tc}$ est donc stationnaire.

Après avoir stationnarisé la série POP_{tc} en la différenciant deux fois, nous pouvons passer alors à la modélisation du processus en utilisant la méthode Box & Jenkins.

1.4 Mise en œuvre de la méthode de Box and Jenkins sur $DDPOP_{tc}$:

a) Spécification du modèle :

Le corrélogramme de la série $DDPOP_{tc}$ fourni par *EViews* est donné par la *figure III.7*

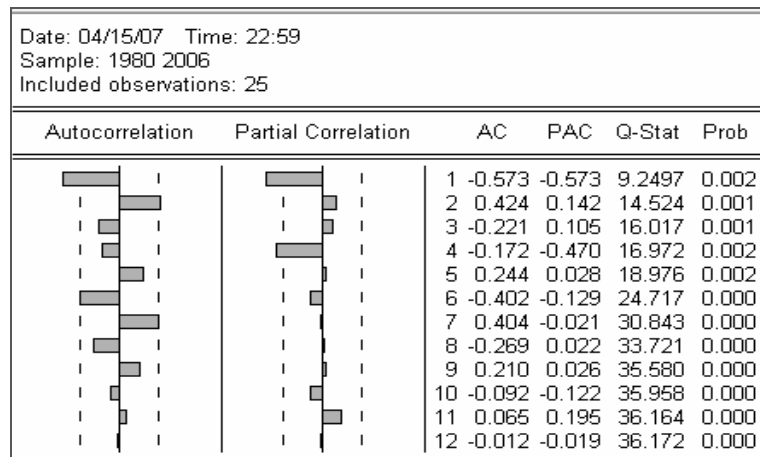


Figure III. 7: Corrélogramme de la série $DDPOP_{tc}$

Nous remarquons que :

- Le corrélogramme simple a quatre pics significatifs aux décalages : 1, 2, 6, et 7.
- Le corrélogramme partiel a deux pics significatifs aux décalages : 1 et 4.

Nous avons donc plusieurs modèles candidats pour représenter la série $DDPOP_{tc}$, parmi lesquels nous avons retenu les modèles suivants : *ARIMA* (1, 2, 2), *ARIMA* (1, 2, 7), *ARIMA* (1, 2, 6).

b) Estimation des paramètres des modèles :

La phase de l'estimation est la phase la plus « mécanique » de la démarche, où il s'agit d'estimer les paramètres des modèles sélectionnés à l'étape d'identification.

Nous utiliserons le logiciel *EViews* pour effectuer l'estimation des paramètres de chaque modèle.

Les résultats pour chaque modèle sont donnés dans les tableaux suivants:

- *Estimation des paramètres de ARIMA (1, 2, 2) :*

L'estimation du modèle *ARIMA (1, 2, 2)* est donnée par le *tableau III.4*

Tableau III. 4 : Estimation du modèle ARIMA (1, 2, 2) de la série DDPOP_{tc}

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.438411	0.193207	-2.269118	0.0334
MA(2)	0.406988	0.201828	2.016508	0.0561
R-squared	0.392135	Mean dependent var		166666.7
Adjusted R-squared	0.364505	S.D. dependent var		16541199
S.E. of regression	13186303	Akaike info criterion		35.70691
Sum squared resid	3.83E+15	Schwarz criterion		35.80508
Log likelihood	-426.4829	Durbin-Watson stat		2.029680
Inverted AR Roots		-.44		

Le coefficient de détermination $R^2 = 0.392135$

- *Estimation des paramètres de ARIMA (1, 2, 7) :*

L'estimation du modèle *ARIMA (1, 2, 7)* est donné par le *tableau (III.5)*

Tableau III. 5 : Estimation du modèle ARIMA (1, 2, 7) de la série DDPOP_{tc}

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.464534	0.174855	-2.656686	0.0144
MA(7)	0.913431	0.044233	20.65065	0.0000
R-squared	0.613230	Mean dependent var		166666.7
Adjusted R-squared	0.595650	S.D. dependent var		16541199
S.E. of regression	10518305	Akaike info criterion		35.25479
Sum squared resid	2.43E+15	Schwarz criterion		35.35296
Log likelihood	-421.0575	Durbin-Watson stat		1.820750
Inverted AR Roots		-.46		
Inverted MA Roots	.89+.43i	.89-.43i	.22-.96i	.22+.96i
	-.62-.77i	-.62+.77i		-.99

Le coefficient de détermination $R^2 = 0.613230$.

- **Estimation des paramètres de ARIMA (1, 2, 6) :**

L'estimation du modèle ARIMA (1, 2, 6) est donné par le *tableau (III.6)*

Tableau III. 6 : Estimation du modèle ARIMA (1, 2, 0) de la série DDPOP_{tc}

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.632255	0.160127	-3.948455	0.0007
MA(6)	-0.944430	0.038228	-24.70497	0.0000
R-squared	0.700674	Mean dependent var		166666.7
Adjusted R-squared	0.687069	S.D. dependent var		16541199
S.E. of regression	9253192.	Akaike info criterion		34.99849
Sum squared resid	1.88E+15	Schwarz criterion		35.09666
Log likelihood	-417.9819	Durbin-Watson stat		1.730616
Inverted AR Roots	-.63			
Inverted MA Roots	.99	.50-.86i	.50+.86i	-.50-.86i
		-.50+.86i	-.99	

Le coefficient de détermination $R^2 = 0.700674$

c) Adéquation des modèles:

Cette phase consiste à examiner les résultats de l'estimation des paramètres des modèles, et cela en faisant subir à chacun de ces modèles, des tests qui permettent la validation ou le rejet de ces derniers.

En effet, il s'agit de vérifier que :

- Les coefficients estimés dans chaque modèle sont significativement différents de zéro.
- Les conditions de causalité et d'inversibilité (de stationnarité) sont vérifiées pour chaque modèle.

- **Test concernant les paramètres du modèle ARIMA (1, 2, 2) :**

- Le rapport de Student du coefficient du modèle moyenne mobile est non significativement différent de zéro, car la probabilité de nullité de ce coefficient (0.0561) est supérieur à 0.05, on rejette donc ce modèle.

- **Test concernant les paramètres du modèle ARIMA (1, 2, 7) :**

- Les rapports de Student des coefficients du modèle (- 2.656686 et 20.65065) sont en valeur absolue supérieurs à 1.96, ils sont donc significativement différents de zéro, ceci est confirmé par les probabilités de nullité de ces coefficients (0.0144 et 0.0000) qui sont inférieures à 0.05.

- Les racines des polynômes autorégressif et moyenne mobile sont en module supérieures à 1 (leurs inverses fournis par *EViews* sont inférieurs à 1) donc les conditions de causalité et d'inversibilité sont vérifiées.

Le modèle *ARIMA* (1, 2, 7) est donc accepté.

- **Test concernant les paramètres du modèle *ARIMA* (1, 2, 6) :**

- Les rapports de Student des coefficients du modèle (-3.948455 et -24.70497) sont en valeur absolue supérieurs à 1.96, ils sont donc significativement différents de zéro. Ceci est confirmé par les probabilités de nullité de ces coefficients (0.0007 et 0.0000) qui sont inférieures à 0.05.
- Les racines du polynômes autorégressif et moyenne mobile sont en module supérieures à 1 (leurs inverses fournis par *EViews* sont inférieurs à 1) donc les conditions de causalité et d'inversibilité sont vérifiées.

Le modèle *ARIMA* (1, 2, 6) est donc accepté.

d) Choix du modèle :

Afin de déterminer la meilleure spécification, des tests supplémentaires devraient être utilisés pour. C'est ainsi que nous introduisons les critères du pouvoir prédictif portant sur la qualité de l'information et du modèle.

Ces tests sont résumés dans le *tableau III.7* :

Tableau III. 7 : Choix du meilleur modèle de la série DDPOPt

<i>Le modèle</i>	R^2	\bar{R}^2	<i>SC</i>	<i>AIC</i>
<i>ARIMA</i> (1, 2, 7)	0.613230	0.595650	35.35296	35.25479
<i>ARIMA</i> (1, 2, 6)	0.700674	0.687069	35.09666	34.99849

Suivant le critère du pouvoir prédictif et les critères d'informations mentionnés dans le *tableau III.7* le modèle *ARIMA* (1, 2, 6) sera choisi pour représenter le processus.

En effet :

- Les coefficients R^2 et \bar{R}^2 du modèle *ARIMA* (1, 2, 6) sont supérieurs à ceux du modèle *ARIMA* (1, 2, 7) ;
- Les critères d'information *AIC* et *SC* du modèle *ARIMA* (1, 2, 6) sont inférieurs à ceux de l'autre modèle.

Nous retiendrons donc le modèle : *ARIMA* (1, 2, 6)

e) Test sur les résidus :

Après avoir choisi le modèle le mieux adapté pour représenter la chronique, nous passons à l'analyse des résidus.

La représentation graphique de la série résiduelle, actuelle et estimée fournie directement par *EViews* est donnée par la *figure III.8*

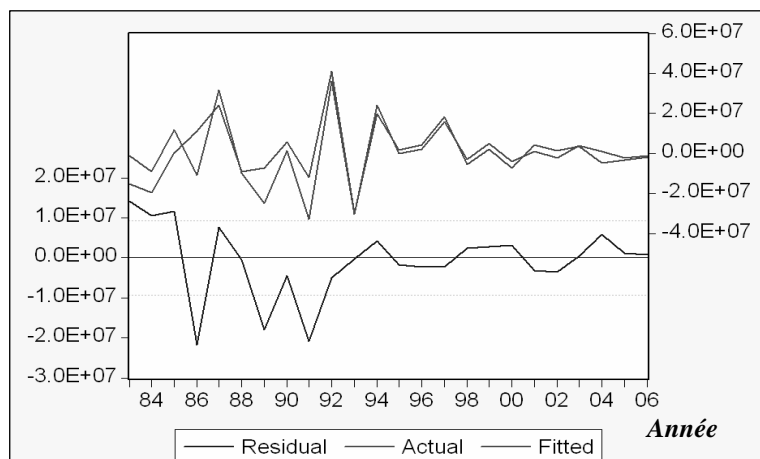


Figure III. 8 : Représentation graphique de la série résiduelle, actuelle et estimée de POPTc

A partir de la représentation graphique de la série résiduelle, actuelle et estimée, nous constatons que le graphe de la série estimée est semblable à celui de la série réelle à quelques pics près, donc le modèle explique à 70,08 % la variation total de la variable dépendante ($R^2 = 0.70$).

Les résidus calculés à partir du modèle estimé jouent un rôle important dans cette dernière étape. En effet il s'agira avant tout, de vérifier que les résidus forment bien un bruit blanc. Par la suite, il s'agira de vérifier que les résidus suivent une loi normale.

- **Corrélogramme des résidus du modèle ARIMA (1, 2, 6) :**

Le corrélogramme du modèle ARIMA (1, 2, 6) est représenté par la *figure III.9* :

Date: 04/24/07 Time: 21:32
 Sample: 1983 2006
 Included observations: 24
 Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.067	0.067	0.1211	
		2	0.100	0.096	0.4032	
		3	0.026	0.014	0.4238	0.515
		4	-0.095	-0.108	0.7029	0.704
		5	0.063	0.073	0.8339	0.841
		6	-0.209	-0.204	2.3461	0.672
		7	-0.186	-0.177	3.6162	0.606
		8	-0.234	-0.206	5.7579	0.451
		9	0.006	0.087	5.7597	0.568
		10	0.044	0.045	5.8442	0.665
		11	0.034	0.038	5.9008	0.750
		12	0.000	-0.072	5.9008	0.824

Figure III. 9 : Corrélogramme des résidus du modèle ARIMA (1, 2, 6)

L'analyse du corrélogramme des résidus, montre que tous les termes sont à l'intérieur de l'intervalle de confiance, les résidus forment donc un bruit blanc. Ceci est confirmé par la statistique de Ljung-Box (Q -stat) fournie directement par *EVIEWS* qui est inférieure à la valeur théorique de $\chi^2(h)$ quelque soit le retard h , et spécialement pour $h=12$. En effet, on a Q -stat(12) = 5.9008 inférieure à $\chi^2(12) = 21.026$ lue au seuil 0.05. Les résidus forment donc un bruit blanc.

• **Test de normalité sur les résidus de ARIMA (1, 2, 6) :**

L’histogramme de la distribution et les valeurs empiriques de Skewness et Kurtosis, de la statistique de Jarque- Bera sont donnés dans la *figure III.10* :

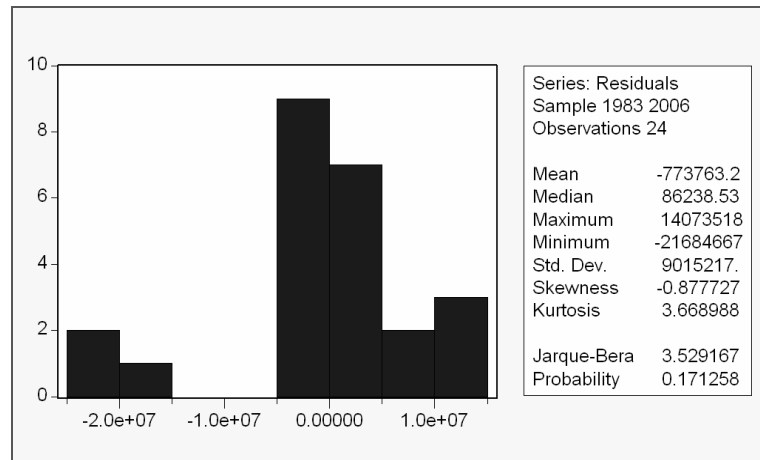


Figure III. 10 : Histogramme de la distribution des résidus du modèle ARIMA (1, 2, 6)

Le coefficient de Skewness (asymétrie) = - 0,877727

Le coefficient de Kurtosis (aplatissement) = 3,668988

On construit alors les statistiques centrées réduites correspondantes, que l’on compare à 1,96 qui représente la valeur de la loi normale au seuil 5%, on obtient :

$$v_1 = (\sqrt{0,877727} / (6 / 27)) = 0,6284 \text{ inférieure à } 1,96$$

$$v_2 = (3,668988 - 3) / \sqrt{(24 / 27)} = 0,7095 \text{ inférieure à } 1,96$$

L’hypothèse de normalité est donc acceptée, ceci est confirmé par la statistique de Jarque-Berra, en effet : $JB = 3.529167$ inférieure à $\chi^2(2) = 5.991$

Nous pouvons alors conclure que les résidus forment bien un bruit blanc gaussien.

Le modèle ARIMA (1, 2, 6) est donc validé.

Une fois le modèle validé, on aborde la dernière phase qui est la prévision.

1.5 La prévision de POP_{tc}:

Les valeurs prévues de la taille de la population qui utilise le Transport en Commun POP_{tc} fournies par *EVIIEWS* sur une période de 5 ans sont résumées dans le *tableau III.8*:

Tableau III. 8 : La prévision de la POP_{tc}

Année	Prévisions
2007	89611467
2008	98197336
2009	105668565
2010	108373675
2011	112915411

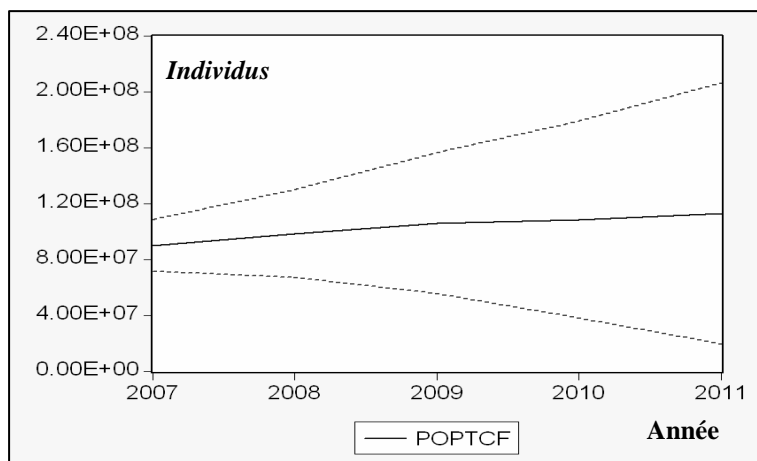


Figure III. 11 : La prévision de la POP_{tc}

La figure III.11 montre que pour les 5 prochaines années (2007, 2011) il est attendu que la taille de la population utilisant le transport en commun POP_{tc} progresse.

En effet, la population transportée par les transports en commun va enregistrer un taux de croissance positif pour atteindre le niveau prévisionnel 112 millions à l'horizon 2011.

La prévision de la taille de la population utilisant le transport en commun étant établie, nous abordons le même travail pour la prévision de la deuxième variable explicative : taille de la population active POP_{act}.

2. Etude de la série « la taille de la population active » POP_{act} :

2.1 Etude de la série brute POP_{act} :

La série POP_{act} représente l'évolution annuelle de la taille de la population active sur une période allant de 1980 à 2006.

a) Analyse préliminaire de la série POP_{act}:

La première étape consiste à l'analyse du graphe représentant l'évolution de la série. Ceci permettra d'avoir une idée sur la nature et les caractéristiques de cette dernière.

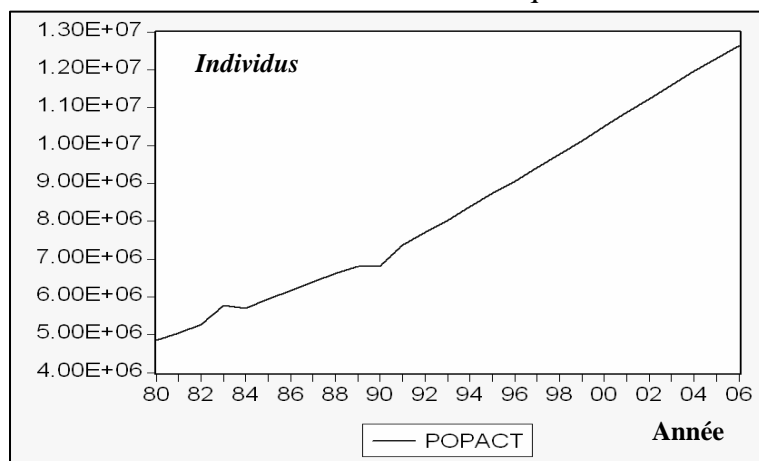


Figure III. 12 : Représentation graphique de la série POP_{act}

La représentation graphique de la série POP_{act} nous montre que la série est affectée d'une tendance, il semble donc que cette série soit non stationnaire.

b) Examen du corrélogramme de la série POP_{act} :

Le corrélogramme de la série POP_{act} est représenté par la *figure III.13*

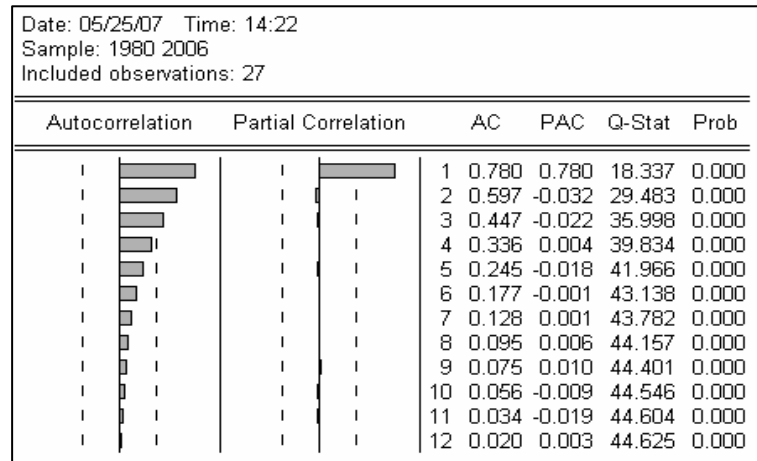


Figure III. 13 : Corrélogramme de la série POP_{act}

L'analyse des corrélogrammes simple et partiel de la série brute nous indique une non stationnarité de cette dernière. En effet, la fonction d'autocorrélation simple décroît lentement vers zéro, nous remarquons aussi de nombreux pics sortant de la bande de confiance, ceci est caractéristique d'une série non stationnaire.

Ceci sera vérifié en appliquant le Test de Dickey- Fuller afin de déterminer la nature du processus non stationnaire et par la suite de le rendre stationnaire.

c) Test de Dickey- Fuller sur la série POP_{act} :

Le Test de Dickey – Fuller a été appliqué par le biais du logiciel *EViews*. Les résultats sont regroupés dans le *tableau III.9*

Tableau III. 9 : Test de Dickey Fuller sur POP_{act}

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	1.70	2.79	$1.70 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	1.24	2.54	$1.24 \leq 2.54$	Constante non significativement différente de zéro
<i>Modèle 4</i>	12.39	-1.95	$12.39 \geq -1.955681$	Série non stationnaire de type DS

Commentaires :

- **Pour le Modèle 6 : modèle avec constante et tendance déterministe**

Nous remarquons que la statistique de Student (t- statistic = 1.70) est inférieure à la valeur critique (=2.79) lue dans la table de Dickey- Fuller au seuil 5%.

La tendance est donc non significativement différente de zéro. Nous passons alors à l'estimation du modèle 5 pour tester la significativité de la constante.

- **Pour le Modèle 5 : modèle avec constante, sans tendance déterministe**

Nous remarquons que :

La statistique de Student (t- statistic = 1.24) est inférieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.

La constante est donc non significativement différente de zéro, nous passons alors à l'estimation du modèle 4 pour vérifier l'existence ou non d'une racine unitaire.

- **Pour le Modèle 4 : modèle sans constante, ni tendance**

On teste l'hypothèse nulle H_0 « Existence d'une racine unitaire » :

La statistique de Student (t- statistic = 12.39) est supérieure à la valeur critique lue dans la table de Dickey–Fuller (-1.95) au seuil 5%, nous acceptons l'hypothèse de non stationnarité H_0 . Il existe donc au moins une racine unitaire. La série POP_{act} est donc non stationnaire de type : DS.

Pour la rendre stationnaire, il faut est la différencier une fois.

La nouvelle série sera notée : $DPOP_{act}$

2.2 Etude de la série différenciée $DPOP_{act}$:

a) Effet de la différenciation :

La représentation graphique de la série différenciée $DPOP_{act}$ est donnée par la figure III.14

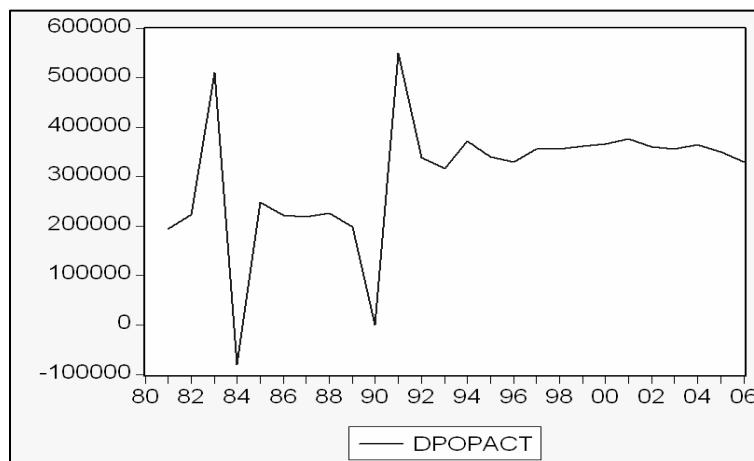


Figure III. 14: Représentation graphique de la série $DPOP_{act}$

Nous remarquons que l'effet de la tendance a été absorbé. Mais, il est nécessaire d'appliquer le Test de Dickey- Fuller pour vérifier la stationnarité de la nouvelle série $DPOP_{act}$.

b) Test de Dickey-Fuller sur la série $DPOP_{act}$:

L'application de Test de Dickey – Fuller sur la série $DPOP_{act}$ a donné les résultats représentés par le *tableau III.10* :

Tableau III. 10 : Test de Dickey Fuller sur $DPOP_{act}$

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	2.78	2.79	$2.78 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	4.94	2.54	$4.94 \geq 2.54$	Constante significativement différente de zéro
	-5.32	-2.98	$-5.32 \leq -2.98$	Série stationnaire

Commentaires :

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 2.78) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.

La tendance est donc non significativement différente de zéro.

- **Pour le Modèle 5 :** *modèle avec constante sans tendance*

Nous remarquons que :

La statistique de Student (t- statistic = 4.94) est supérieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la constante est significativement différente de zéro.

On passe alors directement au test de l'hypothèse H_0 : « existence d'une racine unitaire »

On remarque que la statistique de Student (t- statistique = -5.32) est inférieure à la valeur critique (- 2.98) lue dans la table de Dickey- Fuller au seuil 5%, on rejette alors l'hypothèse H_0 . La série $DPOP_{act}$ est donc stationnaire.

Une fois la série stationnarisée, nous pouvons passer à la modélisation du processus en utilisant la méthode de Box & Jenkins.

2.3 Mise en œuvre de la méthode de Box and Jenkins :

a) Spécification du modèle :

Nous allons examiner le corrélogramme de la série $DPOP_{act}$ représenté par la figure III.15:

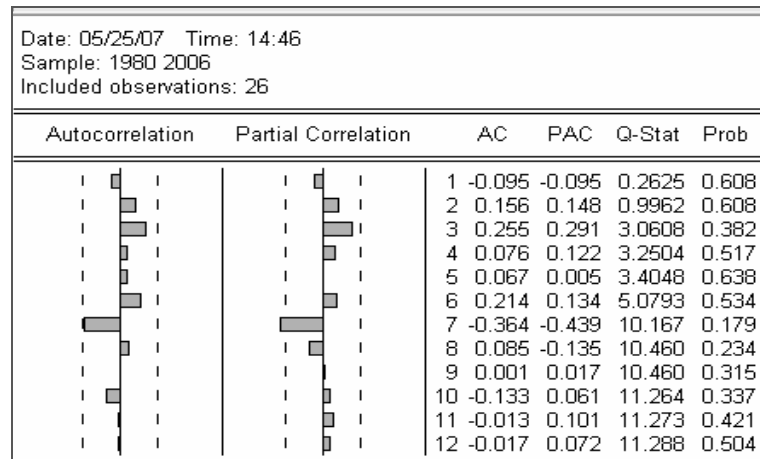


Figure III. 15 : Corrélogramme de la série $DPOP_{act}$

Nous remarquons que :

- Le corrélogramme simple a un pic significatif au décalage : 7.
- Le corrélogramme partiel a aussi un pic significatif au décalage : 7

Nous avons plusieurs modèles candidats pour représenter la série « $DPOP_{act}$ », parmi lesquels nous avons retenu les modèles suivants :

$ARIMA(7, 1, 7)$, $ARIMA(4, 1, 11)$, $ARIMA(3, 1, 7)$;

b) Estimation des paramètres des modèles :

Le logiciel *EViews* est utilisé pour effectuer l'estimation des paramètres de chaque modèle. Les résultats pour chaque modèle sont donnés dans les tableaux ci-dessous (III.11), (III.12), (III.13)

- Estimation des paramètres de $ARIMA(7, 1, 7)$:

Tableau III. 11: Estimation du modèle $ARIMA(7, 1, 7)$ de la variable $DPOP_{act}$

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR (7)	0.890988	0.096955	9.189687	0.0000
MA (7)	0.960330	0.021042	45.63911	0.0000
R-squared	0.339986	Mean dependent var		328581.5
Adjusted R-squared	0.301162	S.D. dependent var		104415.8
S.E. of regression	87287.97	Akaike info criterion		25.69111
Sum squared resid	1.30E+11	Schwarz criterion		25.79053
Log likelihood	-242.0656	Durbin-Watson stat		1.737713
Inverted AR Roots	.98	.61+.77i	.61-.77i	-.22+.96i
	-.22-.96i	-.89-.43i	-.89+.43i	
Inverted MA Roots	.90+.43i	.90-.43i	.22-.97i	.22+.97i
	-.62-.78i	-.62+.78i	-.99	

Le coefficient de détermination $R^2 = 0.339986$.

- **Estimation des paramètres de ARIMA (4, 1, 11) :**

Tableau III. 12: Estimation du modèle ARIMA(4, 1, 11) de la variable DPOP_{act}

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(4)	0.959915	0.021290	45.08679	0.0000
MA(11)	0.972388	0.012276	79.20778	0.0000
R-squared	0.949203	Mean dependent var		315029.5
Adjusted R-squared	0.946663	S.D. dependent var		102903.6
S.E. of regression	23765.42	Akaike info criterion		23.07636
Sum squared resid	1.13E+10	Schwarz criterion		23.17554
Log likelihood	-251.8400	Durbin-Watson stat		1.428264
Inverted AR Roots	.99	.00+.99i	-.00-.99i	-.99
Inverted MA Roots	.96-.28i	.96+.28i	.65+.75i	.65-.75i
	.14+.99i	.14-.99i	-.41+.91i	-.41-.91i
	-.84+.54i	-.84-.54i	-1.00	

Le coefficient de détermination $R^2 = 0.949203$

- **Estimation des paramètres de ARIMA (3, 1, 7) :**

Tableau III. 13 : Estimation du modèle ARIMA (3, 1, 7) de la variable DPOP_{act}

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(3)	0.940885	0.097111	9.688724	0.0000
MA(7)	0.946754	0.023344	40.55582	0.0000
R-squared	0.512847	Mean dependent var		297827.8
Adjusted R-squared	0.489649	S.D. dependent var		130051.8
S.E. of regression	92907.46	Akaike info criterion		25.79954
Sum squared resid	1.81E+11	Schwarz criterion		25.89828
Log likelihood	-294.6947	Durbin-Watson stat		1.511585
Inverted AR Roots	.98	-.49-.85i	-.49+.85i	
Inverted MA Roots	.89+.43i	.89-.43i	.22-.97i	.22+.97i
	-.62-.78i	-.62+.78i	-.99	

Le coefficient de détermination $R^2 = 0.512847$.

c) Adéquation des modèles :

Les paramètres de chaque modèle étant estimés, nous allons à présent examiner les résultats de l'estimation.

• **Test concernant les paramètres du modèle ARIMA (7, 1, 7) :**

- Les rapports de Student des coefficients du modèle (9.189687 et 45.63911) sont en valeur absolue supérieurs à 1.96, ils sont donc significativement différents de zéro. Ceci est confirmé par les probabilités de nullité de ces coefficients (0.0000 et 0.0000) qui sont inférieures à 0.05.
- Les racines des polynômes autorégressif et moyenne mobile sont en module supérieures à 1 (leurs inverses fournis par *EViews* sont inférieurs à 1) donc les conditions de causalité et d'inversibilité sont vérifiées.

Le modèle ARIMA (7, 1, 7) est donc accepté.

• **Test concernant les paramètres du modèle ARIMA (4, 1, 11) :**

- Les rapports de Student des coefficients du modèle (45.08679 et 79.20778) sont en valeur absolue supérieurs à 1.96, ils sont donc significativement différents de zéro, ceci est confirmé par les probabilités de nullité de ces coefficients (0.0000 et 0.0000) qui sont inférieures à 0.05.
- Les racines des polynômes autorégressif et moyenne mobile sont en module supérieures à 1 (leurs inverses fournis par *EViews* sont inférieurs à 1) donc les conditions de causalité et d'inversibilité sont vérifiées.

Le modèle ARIMA (4, 1, 11) est donc accepté.

• **Test concernant les paramètres du modèle ARIMA (3, 1, 7) :**

- Les rapports de Student des coefficients du modèle (9.688724 et 40.55582) sont en valeur absolue supérieurs à 1.96, ils sont donc significativement différents de zéro ceci est confirmé par les probabilités de nullité de ces coefficients (=0.0000 et 0.0000) qui sont inférieures à 0.05.
- Les racines du polynôme autorégressif et moyenne mobile sont en module supérieures à 1 (leurs inverses fournis par *EViews* sont inférieurs à 1) donc les conditions de causalité et d'inversibilité sont vérifiées).

Le modèle ARIMA (3, 1, 7) est donc accepté.

d) Choix du modèle :

Tableau III. 14 : Choix du meilleur modèle pour la variable $DPOP_{act}$

<i>Le modèle</i>	R^2	\bar{R}^2	<i>SC</i>	<i>AIC</i>
ARIMA (7, 1, 7)	0.339986	0.301162	25.79053	25.69111
ARIMA (4, 1, 11)	0.949203	0.946663	23.17554	23.07636
ARIMA (3, 1, 7)	0.512847	0.489649	25.89828	25.79954

Suivant le critère du pouvoir prédictif et les critères d'informations mentionnés dans le tableau ci-dessus, le modèle *ARIMA* (4, 1, 11) sera choisi pour représenter le processus.

En effet :

- les coefficients R^2 et \bar{R}^2 du modèle *ARIMA* (4, 1, 11) sont supérieurs à ceux des modèles *ARIMA* (7, 1, 7) et *ARIMA* (3, 1, 7)
- Les critères d'information AIC et SC du modèle *ARIMA* (4, 1, 11) sont inférieurs à ceux des deux autres modèles.

On retiendra donc le modèle : ***ARIMA* (4, 1, 11)** pour représenter la chronique.

e) Test sur les résidus :

La représentation graphique de la série résiduelle, actuelle et estimée est donnée par la *figure III.16*

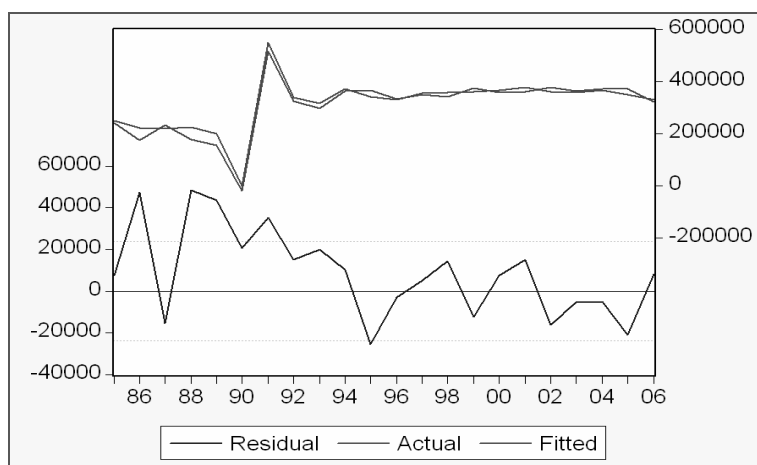


Figure III. 16 : Représentation graphique de la série résiduelle, actuelle et estimée de POP_{act}

Nous constatons que le graphe de la série estimée est semblable à celui de la série réelle à quelques pics près, donc le modèle explique bien à 94.92 % la variation totale de la variable dépendante ($R^2 = 0.949203$)

• **Corrélogramme des résidus du modèle *ARIMA* (4, 1, 11) :**

Le corrélogramme des résidus du modèle *ARIMA* (4, 1, 11) est donné par la *figure III.17*

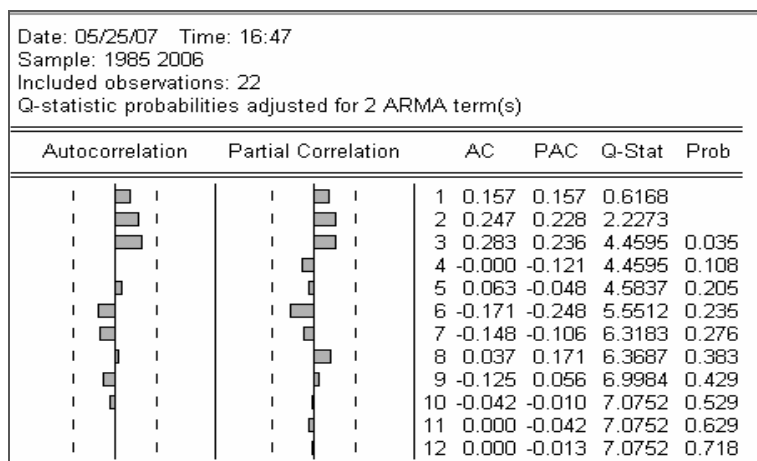


Figure III. 17 : Corrélogramme des résidus du modèle *ARIMA* (4, 1, 11)

L'analyse du corrélogramme des résidus montre que tous les termes sont à l'intérieur de l'intervalle de confiance, les résidus forment donc un bruit blanc. Ceci est confirmé par la statistique de Ljung-Box (Q -stat) qui est inférieure à la valeur théorique de $\chi^2(h)$ quelque soit le retard h , et spécialement pour $h=12$. En effet, nous avons :

Q -stat (12) = 7.0752 inférieure à $\chi^2(12) = 21.026$ lue au seuil 0.05, les résidus forment donc un bruit blanc.

- **Test de normalité sur les résidus de ARIMA (4, 1, 11) :**

L'histogramme de la distribution et les valeurs empiriques de Skewness, de Kurtosis et de la statistique de Jarque- Bera sont donnés par la *figure III.18*

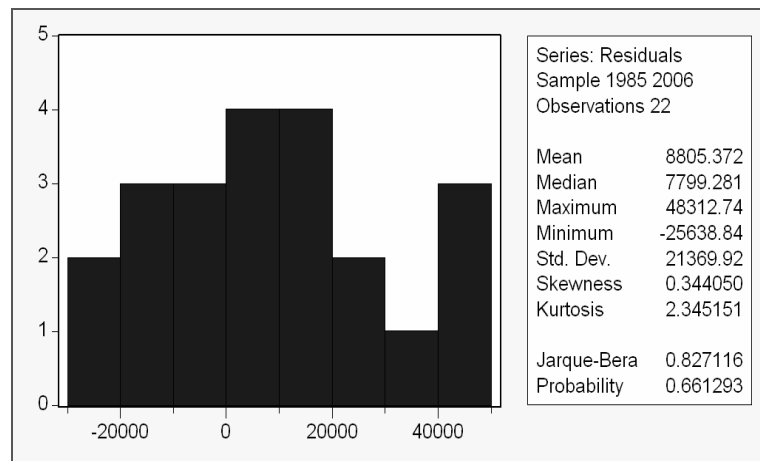


Figure III. 18 : Histogramme de la distribution des résidus de ARIMA (4, 1, 11)

Le coefficient de Skewness (asymétrie) = 0.344050

Le coefficient de Kurtosis (aplatissement) = 2.345151

On construit alors les statistiques centrées réduites correspondantes, que l'on compare à 1.96 qui représente la valeur de la loi normale au seuil 5%, nous avons :

$$v_1 = (\sqrt{0,344050}) / (6 / 27) = 1,2442 \text{ Inférieure à } 1,96$$

$$v_2 = (3 - 2,345151) / \sqrt{(24 / 27)} = 0,6548 \text{ Inférieure à } 1,96$$

L'hypothèse de normalité des résidus est donc acceptée, ceci est confirmé par la statistique de Jarque- Berra, en effet : $JB = 0.827116$, cette valeur est inférieure à $\chi^2(2) = 5.991$

Nous concluons donc que les résidus forment un bruit blanc gaussien.

Le modèle ARIMA (4, 1, 11) est donc validé.

Une fois le modèle validé, nous passons à la dernière phase qui est la prévision.

2.4 La prévision de la série POP_{act} :

Les valeurs prévues de la taille population active (POP_{act}), fournies par *EVIIEWS*, pour une période de 5 ans, sont réunies dans le *tableau III.15*

Année	Prévisions
2007	12968623
2008	13323814
2009	13673714
2010	13977277
2011	14310056

Tableau III. 15 : Prévision de POP_{act}

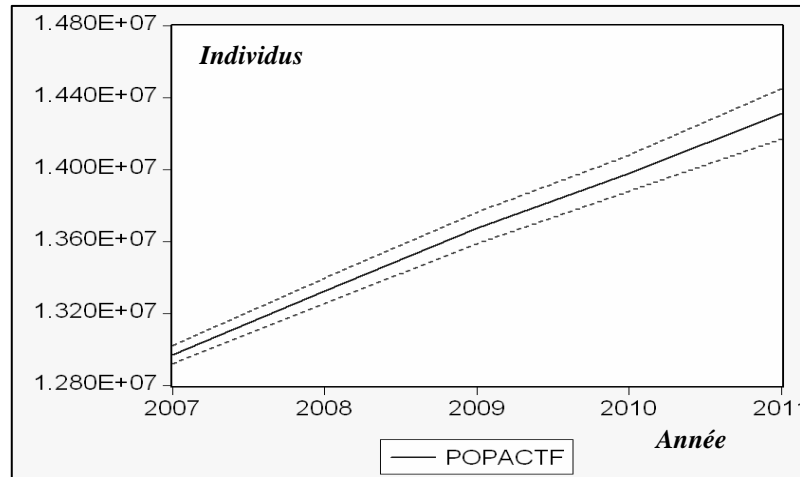


Figure III. 19 : La prévision de la Pop_{act}

Il ressort de la *figure III.19*, que la population active va enregistrer un taux de croissance positif. Elle sera en augmentation continue sur les 5 prochaines années pour atteindre le niveau prévisionnel de 14 millions, à l'horizon 2011.

La prévision de la taille de la population active étant établie, nous effectuerons le même travail pour la prévision des trois autres variables explicatives à savoir : le taux de croissance de la taille du parc auto, le PIB, et le prix de l'essence.

3. Etude de la série : « le taux de croissance du parc auto » T_{cr}:

L'étude de la série T_{cr} montre que cette dernière est stationnaire. Le modèle le plus adapté pour représenter la chronique est le modèle *ARIMA* (1, 0, 5), avec un $R^2 = 0.9459$.

En outre les résidus de ce dernier forment un bruit blanc gaussien, ce qui nous permet de valider le modèle. Les détails des résultats sont donnés en annexe 5.

Après avoir validé le modèle, passons à la prévision du taux de croissance de la taille du parc automobile. Les résultats de la prévision sont donnés par le *tableau III.16*.

Année	Prévisions
2007	0.031470
2008	0.035348
2009	0.031834
2010	0.034028
2011	0.040234

Tableau III. 16 : Prévision du T_{cr}

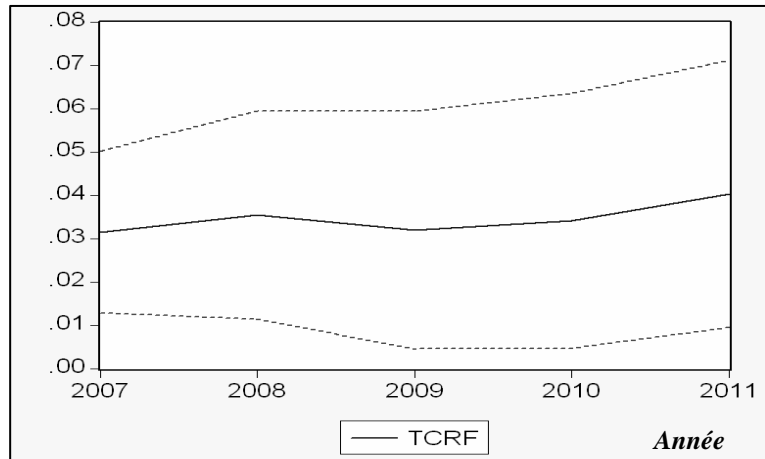


Figure III. 20 : La prévision de la T_{cr}

La figure III.20 montre que le taux de croissance du parc automobile devrait connaître une très légère progression pour les 5 prochaines années.

4. Etude de la série PIB:

La série PIB est non stationnaire. Après deux différenciations successives, la série devient stationnaire. Elle est notée : DDPIB. Le modèle le plus adapté pour représenter la chronique est le modèle *ARIMA* (2, 2, 4), avec un $R^2 = 0.4048$. Les résidus de ce dernier forment un bruit blanc gaussien, ce qui nous permet de valider le modèle. Les détails des résultats sont donnés en annexe 5.

Après validation du modèle, le calcul de la prévision du PIB conduit aux résultats donnés par le tableau III.17

Année	Prévisions
2007	2662608
2008	2731086
2009	2828509
2010	2958032
2011	3049949

Tableau III. 17 : Prévision du PIB

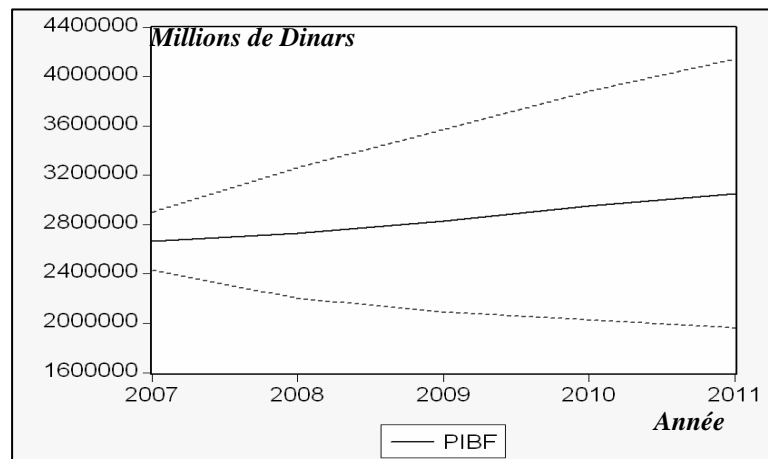


Figure III. 21: la prévision du PIB

La figure III. 21 montre que le PIB sera en légère progression durant les 5 prochaines années. Cet agrégat se rattachant étroitement à la notion de la valeur ajoutée, permet de conclure que la masse des biens et services produits, par les unités résidentes et mises à la disposition des utilisateurs finals, enregistrera un taux de croissance positif, pour atteindre le niveau prévisionnel de 3.10^6 Millions de dinars pour l'année 2011.

5. Etude de la série Prix essence P_{ess} :

La série P_{ess} est non stationnaire, différenciée deux fois, la série devient stationnaire. Elle est notée : DDP_{ess} . Le modèle le plus adapté pour représenter la chronique est le modèle $ARIMA(1, 2, 8)$, avec un $R^2 = 0.679015$, les résidus de ce dernier forment un bruit blanc gaussien, d'où la validation du modèle (cf. annexe 5).

Les résultats de la prévision sont donnés par le tableau III.18

Année	Prévisions
2007	22.17
2008	23.18
2009	24.42
2010	24.79
2011	24.70

Tableau III. 18: La prévision du P_{ess}

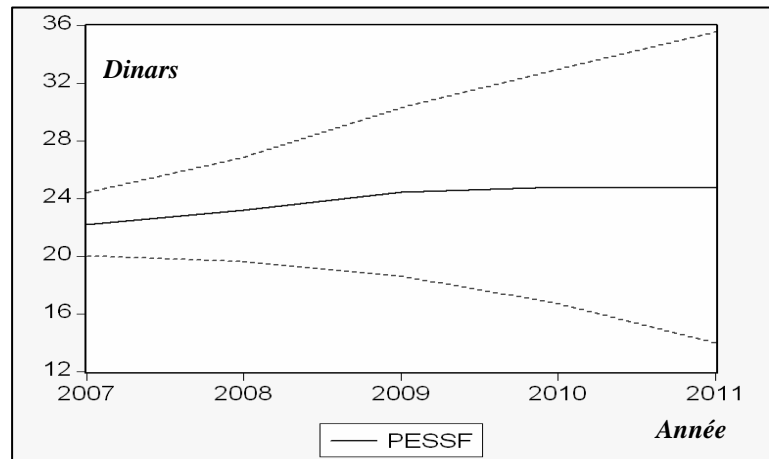


Figure III. 22: de La prévision du P_{ess}

La figure III.22 montre que le prix de l'essence augmenterait légèrement durant les 5 prochaines années. En effet, le prix administré du litre d'essence va enregistrer une très faible augmentation pour atteindre le niveau de 24-25 dinars, à partir de 2009.

Les prévisions des variables explicatives de la consommation du GPLc ayant été effectuées, nous pouvons passer à la prévision de cette dernière.

II. La prévision de la consommation du GPLc :

Les prévisions des variables explicatives sont reportées dans le tableau III. 19:

Tableau III. 19 : Tableau récapitulatif des Prévisions des variables explicatives

Année	POP_{tc}	POP_{act}	T_{cr}	PIB	P_{ess}
2007	89611467	12968623	0.031470	2662608	22.17
2008	98197336	13323814	0.035348	2731086	23.18
2009	105668565	13673714	0.031834	2828509	24.42
2010	108373675	13977277	0.034028	295803	24.79
2011	112915411	14310056	0.040234	3049949	24.70

Afin d'obtenir les prévisions concernant la consommation du GPLc, on injecte les prévisions obtenues pour chaque variable explicative dans le modèle, régi par ces dernières. Ce modèle est décrit comme suit :

$$C_{Gplc} = -342126.1286 + 0.06813749497*PIB + 0.0007059843269*POP_{tc} + 571136.6743*T_{cr} + 0.02742276076*POP_{act} + 3497.008951*P_{ess}$$

Les résultats de la prévision de la consommation du GPLc, fournis par *Eviews*, sont donnés dans le *tableau III.20* :

<i>Année</i>	<i>Prévisions</i>
2007	353699
2008	379921
2009	403778
2010	424898
2011	447221

Tableau III. 20: La prévision de la consommation du C_{GPLc}

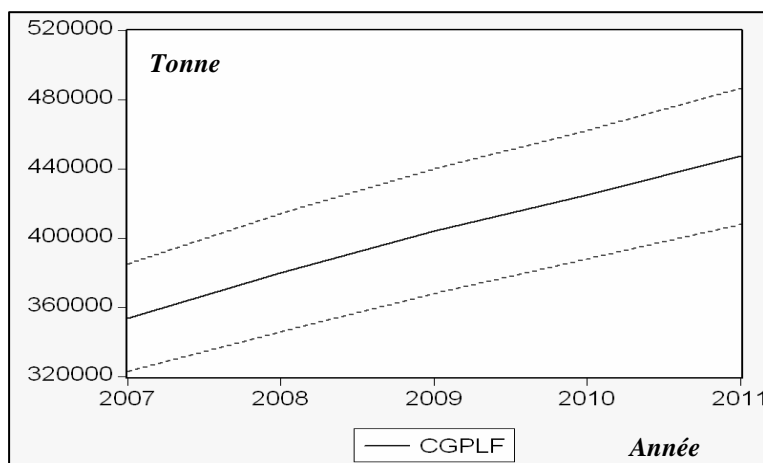


Figure III. 23: La prévision de la consommation du C_{GPLc}

La *figure III.23* montre que la consommation du GPL carburant va continuer sa progression pour atteindre plus de 4 Millions de tonnes à l'horizon 2011. Le taux annuel de cette progression, après avoir enregistré 7%, entre 2007 et 2008, tendra à se stabiliser autour de 5,2% par an. Cette stabilité pourrait être interprétée comme le reflet de la stabilité du comportement des usagers.

Afin d'appréhender l'attitude du consommateur algérien, vis-à-vis des carburants disponibles, nous avons procédé à un sondage sur un échantillon aléatoire de deux cents individus (cf. annexe 7), diversifiés en âge, sexe et situation sociale. Le sondage basé sur un questionnaire, effectué au niveau de diverses entités telles que le Ministère de l'Énergie et des Mines, Sonatrach, Naftal, l'hôpital Mustapha, les universités, les stations de service, les stations de taxi, les parkings etc.

Ainsi, une série de questions ont été posées aux citoyens, dont nous reproduisons, ci-après les plus significatives, accompagnées de leur réponse, dans le but de cerner l'attitude de ces derniers, à l'égard des carburants commercialisés :

- A la question n° 6 de savoir si les consommateurs sont satisfaits de leur carburant actuel:
- * 73 % des consommateurs de GPLc sont satisfaits, au plan économique, de leur carburant. Les 27 % restants ne le sont pas, en raison de problèmes de disponibilité, d'interdiction de garage en parking souterrain et du manque d'installateurs professionnels, agréés par les autorités compétentes.
- * 39 % des consommateurs essence sont satisfaits, principalement en raison de la disponibilité de ce carburant. Le reste, soit 61 % ne sont pas satisfaits, à cause du coût à la pompe et seulement de ce dernier.
- * 89 % des consommateurs de Gasoil sont plutôt satisfaits sur le plan économique et performance du diesel. 11 % ne le sont pas, évoquant le problème de pièces de rechange très chères.

- A la question n° 14 de savoir quel est le carburant le plus polluant, 51 % des personnes interrogées, considèrent que le GPLc fait partie des carburants les moins polluants, contre 49 % qui pensent le contraire.
- A la question n° 16 pourquoi le GPL c n'est pas très attractif:
 - 7 % ne savent pas ;
 - 18 % disent que c'est un problème de méconnaissance du produit ;
 - 41 % disent que c'est un problème de réduction du volume du coffre ;
 - 31 % disent que c'est un problème de disponibilité ;
 - 16 % disent que c'est un problème de conversion ;
 - 14 % disent que c'est un problème d'usure du moteur ;
 - 88 % disent que c'est un problème de sécurité (risque d'explosion) ;

Les résultats de ce sondage mettent en exergue le manque d'information des citoyens sur le GPLc. Certains le connaissent, mieux encore, quand ils ont entendu parler, c'est pour rappeler surtout ses inconvénients, souvent exagérés : risque d'explosion, volume du coffre du véhicule réduit à cause du réservoir supplémentaire pour le GPLc, restriction de garage dans les parkings souterrains etc... Les professionnels usagers de véhicules à GPLc, quant à eux, se plaignent du manque d'installateurs spécialisés agréés.

Aussi, le GPLc est complètement méconnu comme carburant moins polluant que l'essence et le diesel. Dans ce cadre, il faut bien souligner que l'échantillon est loin de paraître sensibilisé sur les questions de la pollution de l'environnement, ceci ne constitue pas une préoccupation majeure des consommateurs. Il apparaît à travers les réactions de l'échantillon que l'attitude des personnes interrogées est loin de pouvoir peser sur la stabilité de la consommation du GPLc.

Ainsi, et essentiellement sur la base des résultats de l'étude prévisionnelle menée et dans une moindre mesure de ceux du sondage effectué, pourrait-on présager qu'il n'y aura pas de bouleversement dans la hiérarchie de préférence des carburants à moyen terme, sauf si une politique volontariste des décideurs, en faveur du GPLc, était décidée ou un choc extérieur venait à perturber le présent équilibre.

C'est dans cette optique que, dans le chapitre qui suit, il sera considéré un scénario qui mettrait en œuvre un tel choc. Nous nous proposons d'illustrer son impact sur les liaisons dynamiques entre les valeurs explicatives, étudiées précédemment, ainsi que l'ampleur des changements structurels liés à ce choc, si changements il y a.

***Chapitre IV : La modélisation
Vectorielle AutoRégressive***

La méthodologie des Vecteurs Auto Régressifs « VAR » consiste à modéliser un ensemble de séries chronologiques à l'aide d'un processus AR (p). [BOU 02]

En pratique, on dispose souvent d'observations simultanées de plusieurs séries chronologiques qui nous permettent d'étudier les liaisons existant entre elles et de voir comment celles-ci évoluent dans le temps.

En effet, un modèle VAR se définit comme un outil économétrique particulièrement adapté pour rendre compte de l'ensemble des liaisons dynamiques à l'intérieur d'un groupe donné de variables. [EVR 01]

En règle générale, chacune des variables est expliquée non seulement par ses propres retards mais aussi par les valeurs présentes et passées des autres variables du système.

La méthodologie VAR a l'avantage de poser le moins de restrictions possibles sur les relations qui existent entre les variables du système et de laisser ainsi les données s'exprimer. De plus, les modèles VAR constituent un cadre d'analyse simple qui nous permet d'étudier, d'une manière systématique, les réponses dynamiques des variables considérées face à un certain nombre de chocs.

Nous nous sommes intéressées aux études traitant de ce type modélisation, dont le domaine d'application s'avère très vaste, couvrant des secteurs aussi variés que les marchés boursiers l'énergie, la météorologie et même l'hydrologie. A titre d'exemple, nous citerons le travail de KHELMI [KHE 05] qui a étudié la contribution des investissements en stocks à la volatilité du PIB réel au Canada, de même que celui de RAULT C. [RAU 02], dédié à la relation qui existe entre le groupe formé par les variables inflation, salaires et SMIC. Enfin, évoquons l'étude de AMBAPOUR S. [AMB 05], consacrée à une analyse, en terme de causalité, entre la croissance économique et la croissance de l'énergie au Congo.

Dans le cadre de l'analyse proposée, la modélisation VAR se fera entre les variables suivantes :

- Le prix du diesel ;
- La consommation du diesel ;
- La consommation de l'essence et
- La consommation du GPLc.

Notre analyse est suscitée par la forte pression observée ces dernières années sur la consommation des carburants, particulièrement celle du gasoil dont la demande affiche des taux de croissance très élevés, dus au fait de son efficacité et surtout **de son faible coût**. Ce dernier s'est maintenu stable durant ces dix dernières années, ce qui a eu pour conséquence la multiplication de la consommation du gasoil par deux.

Il paraît donc intéressant d'étudier les liaisons dynamiques à l'intérieur du groupe formé par les variables énumérées ci-dessus, de chercher à identifier et à comprendre les sources des fluctuations de ces dernières en utilisant l'analyse VAR, ainsi que de mesurer l'importance relative des chocs dans l'explication des fluctuations.

A cet effet, notre travail sera scindé en deux parties: la première sera consacrée à la présentation d'une manière succincte des principes de base de la modélisation vectorielle autorégressive, la seconde portera sur l'analyse VAR entre les variables sus-citées à savoir : le prix du gasoil, la consommation du GPLc, la consommation de l'essence, et la consommation du gasoil.

Rappelons que dans notre étude, nous nous sommes intéressées qu'aux véhicules particuliers, mais suite à l'indisponibilité des données concernant la consommation des

particuliers du **gasoil**, nous étions donc contraints de considérer la consommation totale qui inclut le transport de marchandises et de voyageurs

Section 1: Principes de base de la modélisation vectorielle Autorégressive : [BOU 02]

I. Représentation d'un modèle VAR : [BOU 02]

Soit une représentation VAR à k variables $y_{1t}, y_{2t}, \dots, y_{kt}$ (chacune de ces variables est fonction de ses propres valeurs passées et de celles des autres variables) et p décalages [notée VAR(p)] notée sous la forme matricielle :

$$Y_t = A_0 + A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + E_t$$

Avec :

$$Y_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{k,t} \end{bmatrix}; A_p = \begin{bmatrix} a_{1p}^1 & a_{1p}^2 & \dots & a_{1p}^k \\ a_{2p}^1 & a_{2p}^2 & & a_{2p}^k \\ \vdots & & & \vdots \\ a_{kp}^1 & a_{kp}^2 & \dots & a_{kp}^k \end{bmatrix}; A_0 = \begin{bmatrix} a_1^0 \\ a_2^0 \\ \vdots \\ a_k^0 \end{bmatrix}; E_t = \begin{bmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{kt} \end{bmatrix}$$

Où

Y_t : est un vecteur colonne d'ordre k ;

A_k : est une matrice carrée d'ordre k telle que $k=1, \dots, p$;

A_0 : est un vecteur colonne de dimension k ;

E_t : est un vecteur colonne de dimension k, représentant le bruit blanc ;

$y_{1t}, y_{2t}, \dots, y_{kt}$ sont considérées comme étant stationnaires, les perturbations $e_{1t}, e_{2t}, \dots, e_{kt}$ sont des bruits blancs de variances constantes et non autocorrélées.

On note $\Sigma_e = E(e_t, e_t')$, la matrice de dimension (k, k) des variances covariances des erreurs. Cette matrice est bien sûr inconnue.

II. Condition de stationnarité du modèle : [BOU 02]

Un modèle VAR est stationnaire, s'il satisfait aux trois conditions classiques :

- $E(Y) = \mu, \forall t$;
- $Var(Y_t) < \infty$;
- $Cov(Y_t, Y_{t+k}) = E(Y_t - \mu)(Y_{t+k} - \mu) = \Gamma_k, \forall t$

En pratique on démontre qu'un processus VAR(p) est stationnaire si le polynôme défini à partir du déterminant : $\det(I - A_1 z - A_2 z^2 - \dots - A_p z^p) = 0$ a ses racines à l'intérieur du cercle unité du plan complexe.

III. Estimation des paramètres : [BOU 02], [KHE 05]

Dans le cas d'un processus VAR, chacune des équations peut être estimée par les MCO, indépendamment des autres (ou par la méthode de maximum de vraisemblance)
Soit le modèle VAR estimé :

$$Y_t = \hat{A}_0 + \hat{A}_1 Y_{t-1} + \hat{A}_2 Y_{t-2} + \dots + \hat{A}_p Y_{t-p} + e_t$$

e_t : étant le vecteur de dimension $(k, 1)$ des résidus d'estimation $e_{1t}, e_{2t}, \dots, e_{kt}$

Les paramètres du processus VAR ne peuvent être estimés que sur des séries chronologiques stationnaires. Ainsi, après étude des caractéristiques des séries, les séries sont soit stationnarisées par différence dans le cas d'un processus DS, soit il est possible d'ajouter une composante tendance à la spécification VAR dans le cas d'un processus TS.

IV. Détermination du nombre de retard p : [BOU 02]

Pour déterminer le nombre de retards, nous allons présenter les critères de AKAIKE (AIC) et de SCHWARTZ (SC). Dans le cas de la représentation VAR, ces critères peuvent être utilisés pour déterminer l'ordre p du modèle. La procédure de sélection de l'ordre de la représentation consiste à estimer tous les modèles VAR pour un ordre allant de 1 à h (h étant le retard maximum admissible par la théorie économique ou par les données disponibles)

Les fonctions AIC (p) et SC (p) sont calculées de la manière suivante :

$$AIC(p) = Ln[\det[\Sigma_e]] + \left(\frac{2k^2 p}{n} \right)$$

$$SC(p) = Ln[\det[\Sigma_e]] + \left(\frac{k^2 p \ln(n)}{n} \right)$$

Avec :

k : le nombre de variables du système.

n = nombre d'observations.

p = le nombre de retards.

Σ_e = matrice des variance covariance des résidus du modèle.

Le retard p retenu est le retard qui minimise les critères AIC ou SC.

V. Dynamique d'un modèle VAR:

Les modèles VAR permettent d'analyser les effets de la politique économique, cela au travers de simulation de chocs aléatoires (d'innovations) et de la décomposition de la variance de l'erreur. Cependant, cette analyse s'effectue en postulant la constance de l'environnement économique « toutes choses étant égales par ailleurs ».

V.1 Représentation VMA d'un processus VAR : [RAU 02]

Comme un modèle AR (1) à une représentation MA (∞), par analogie un modèle VAR(1) à une représentation VMA (∞). Le modèle sous cette forme va permettre de mesurer l'impact sur les valeurs présentes d'une variation des innovations (ou des chocs) de ε_{1t} et de ε_{2t} .

Soit la représentation VAR(p) stationnaire :

$$Y_t = A_0 + A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + E_t$$

Sa représentation VMA (∞) est donnée par :

$$Y_t = \mu + \varepsilon_t + M_1 \varepsilon_{t-1} + M_2 \varepsilon_{t-2} + \dots = \mu + \sum_{i=1}^{\infty} M_i \varepsilon_{t-i}$$

$$\text{Avec } \mu = (1 - A_1 - A_2 - \dots - A_p) = \sum_{j=1}^{\min(p,i)} \hat{A}_j M_{i-j}, i = 1, 2, \dots \text{ et } M_0 = 1$$

Sous cette forme, la matrice M apparaît comme un « multiplicateur d'impact », c'est-à-dire que c'est au travers de cette matrice qu'un choc se répercute tout au long du processus.

Une variation à un instant donné t de ε_t affecte toutes les valeurs suivantes de Y_t , l'effet d'un choc (ou d'une innovation) est donc permanent et va en s'amortissant.

V.2 Analyse des chocs : [BOU 02] [AMB 05]

L'analyse d'un choc consiste à mesurer l'impact de la variation d'une innovation sur les variables.

Soit : VAR(1)

$$y_{1t} = a_1^0 + a_{11}^1 y_{1t-1} + a_{11}^2 y_{2t-1} + \varepsilon_{1t}$$

$$y_{2t} = a_2^0 + a_{21}^1 y_{1t-1} + a_{21}^2 y_{2t-1} + \varepsilon_{2t}$$

Une variation de ε_{1t} à un instant donné a une conséquence immédiate sur y_{1t} puis sur y_{1t+1} et y_{2t+1} . Par exemple s'il se produit en t un choc sur ε_{1t} égal à 1 (unitaire) nous avons l'impact suivant:

$$\text{En } t : \begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{En } t + 1 : \begin{bmatrix} \Delta y_{1t+1} \\ \Delta y_{2t+1} \end{bmatrix} = \begin{bmatrix} a_{11}^1 & a_{11}^2 \\ a_{21}^1 & a_{21}^2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{En } t + 2 : \begin{bmatrix} \Delta y_{1t+2} \\ \Delta y_{2t+2} \end{bmatrix} = \begin{bmatrix} a_{11}^1 & a_{11}^2 \\ a_{21}^1 & a_{21}^2 \end{bmatrix} \begin{bmatrix} \Delta y_{1t+1} \\ \Delta y_{2t+1} \end{bmatrix}$$

Etc ...

Les différentes valeurs ainsi calculées constituent « la fonction de réponse impulsionnelle »

Le choix du sens de l'impact est important et conditionne les valeurs obtenues, nous pouvons observer que l'effet d'une innovation s'estompe au cours du temps.

La généralisation à un modèle VAR à k variables nécessite le recours à des procédures d'orthogonalisation de matrice et s'avère donc complexe. Les logiciels d'économétrie tel que *EViews* offrent la possibilité de choisir l'ordre des variables et permettent ainsi de simuler les différents cas de figure.

V.3 La décomposition de la variance : [BOU 02] [KHE 05]

La décomposition de la variance de l'erreur de prévision a pour objectif de calculer pour chacune des innovations sa contribution à la variance de l'erreur. Par une technique mathématique, on peut écrire la variance de l'erreur de prévision à un horizon h en fonction de la variance de l'erreur attribuée à chacune des variables ; il suffit ensuite de rapporter chacune de ces variances à la variance totale pour obtenir son poids relatif en pourcentage.

Soit Var (1) à deux variables y_{1t} et y_{2t} , la variance de l'erreur de prévision pour y_{1t+h} s'écrit :

$$\sigma_{y_1}^2(h) = \sigma_{y_1}^2 [m_{11}^2(0) + m_{11}^2(1) + \dots + m_{11}^2(h-1)] + \sigma_{y_2}^2 [m_{22}^2(0) + m_{22}^2(1) + \dots + m_{22}^2(h-1)]$$

Où les m_{ii} sont les termes de la matrice M de la section V.1.

A l'horizon h, la décomposition de la variance, en pourcentage, des propres innovations de y_{1t} sur y_{1t} , est donnée par :

$$\frac{\sigma_{y_1}^2 [m_{11}^2(0) + m_{11}^2(1) + \dots + m_{11}^2(h-1)]}{\sigma_{y_1}^2(h)}$$

Et la décomposition de la variance, en pourcentage, des innovations de y_{1t} sur y_{2t} , est donnée par :

$$\frac{\sigma_{y_2}^2 [m_{22}^2(0) + m_{22}^2(1) + \dots + m_{22}^2(h-1)]}{\sigma_{y_1}^2(h)}$$

- Si un choc sur ε_{1t} n'affecte pas la variance de l'erreur de y_{2t} quel que soit l'horizon de prévision, alors y_{2t} peut être considéré comme exogène car y_{2t} évolue indépendamment de ε_{1t} .
- A contrario, si un choc sur ε_{1t} affecte fortement la variance de l'erreur de y_{2t} , alors y_{2t} est considéré comme endogène.

V.4 La causalité : [BOU 02]

Au niveau théorique, la mise en évidence de relation causale entre les variables économétriques fournit des éléments de réflexions propices à une meilleure compréhension des phénomènes économétriques. De manière pratique, « the causal knowledge » est nécessaire à une formulation correcte de la politique économétrique.

Nous nous limiterons à l'exposé de la causalité au sens de Granger qui est la plus fréquemment utilisée en économétrie.

- **Causalité au sens de Granger**

Granger (1969) a proposé les concepts de causalités et d'exogénéité. On dit que la variable x cause au sens de Granger la variable y si et seulement si la connaissance du passé de x améliore la prévision de y à tout horizon.

Soit le modèle VAR(p) pour lequel les variables y_{1t} et y_{2t} sont stationnaires :

$$y_{1t} = a_1^0 + \sum_{i=1}^P a_{1i}^1 y_{1t-i} + \sum_{i=1}^P a_{1i}^2 y_{2t-i} + \varepsilon_{1t}$$

$$y_{2t} = a_2^0 + \sum_{i=1}^P a_{2i}^1 y_{1t-i} + \sum_{i=1}^P a_{2i}^2 y_{2t-i} + \varepsilon_{2t}$$

Un test d'hypothèse permet de conclure sur le sens de la causalité, ainsi y_{2t} cause y_{1t} au sens de Granger si l'hypothèse nulle définie ci-dessous est rejetée au profit de l'hypothèse alternative :

$$H_0 : y_{2t} \text{ ne cause pas } y_{1t} \text{ si l' hypothèse : } a_{1i}^2 = 0 \text{ avec } i = \overline{1, p}$$

$$H_1 = \text{non } H_0 : \exists a_{1i}^2 \neq 0$$

De façon analogue, y_{1t} cause au sens de Granger y_{2t} si l'hypothèse nulle définie ci-dessous est rejetée au profit de l'hypothèse alternative :

$$H_0 : y_{1t} \text{ ne cause pas } y_{2t} \text{ si l' hypothèse : } a_{2i}^1 = 0 \text{ avec } i = \overline{1, p}$$

$$H_1 = \text{non } H_0 : \exists a_{2i}^1 \neq 0$$

Ces tests peuvent être conduits à l'aide d'un test de Fisher classique. Par ailleurs, si on est mené à rejeter les deux hypothèses nulles, on a une causalité bidirectionnelle, on parle de boucle rétroactive (feedback effect)

Section 2 : Analyse VAR entre les variables (P_{gas} , C_{GPLC} , C_{ess} , C_{gas})

Le but de cette section est d'étudier les liaisons dynamiques existant à l'intérieur du groupe formé par les variables : prix du gasoil, consommation du GPLC, consommation de l'essence et consommation du gasoil (P_{gas} , C_{GPLC} , C_{ess} , C_{gas} respectivement). En effet, nous allons tenter d'identifier et de comprendre les sources des fluctuations de ces dernières en utilisant l'analyse VAR.

Pour ce faire, nous allons analyser la série chronologique multi variée Y_t qui couvre la période allant de 1980 à 2006. Elle est composée des 4 séries suivantes :

P_{gas} : Le prix du gasoil ;

C_{GPLC} : La consommation du GPLC;

C_{ess} : La consommation de l'essence ;

C_{gas} : La consommation du gasoil.

Avec : $k = 4$ (nombre de variables), $N = 27$ (nombre d'observations)

La première partie de notre analyse visera la spécification de l'ordre du modèle VAR, la seconde traitera des relations dynamiques existantes entre les variables qui le composent à travers les trois instruments d'analyse que sont le test de causalité de Granger, la décomposition de la variance de l'erreur de prévision et les fonctions aux chocs.

I. Spécification du modèle VAR :

La spécification de l'ordre du modèle VAR, repose sur les étapes suivantes :

- Vérification de la stationnarité des séries composant le modèle VAR à savoir : P_{gas} , C_{GPLC} , C_{ess} et C_{gas} .
- Détermination de l'ordre du modèle VAR optimal.

1. Stationnarisation des séries :

Il s'agit de vérifier la stationnarité des séries composants le modèle VAR, à savoir :

P_{gas} , C_{gas} , C_{ess} et C_{GPLC} .

Après application du test de Dickey Fuller sur les séries (cf. annexe 6), nous constatons que :

- La série C_{GPLC} est non stationnaire, différenciée deux fois, la série devient stationnaire, elle est notée : DDC_{GPLC} ;
- La série C_{ess} est non stationnaire, différenciée une fois, la série devient stationnaire, elle est notée : DC_{ess} ;
- La série C_{gas} est non stationnaire, différenciée deux fois, la série devient stationnaire elle est notée : DDC_{gas} ;
- La série P_{gas} est non stationnaire, différenciée deux fois, la série devient stationnaire elle est notée : DDP_{gas} .

Après avoir stationnarisé ces séries, nous abordons la deuxième étape de la spécification qui est la détermination de l'ordre du modèle VAR optimal.

2. Détermination de l'ordre du modèle VAR optimal :

L'objectif de cette partie est la construction d'un modèle VAR à 4 variables et p décalages noté : $X_t = (DDP_{gas}, DDC_{gas}, DC_{ess}, DDC_{GPLC})$ avec X_t un vecteur 4×1

Détermination de p :

La détermination du retard « p » se fait en sélectionnant le retard p qui minimise les critères d'Akaike (AIC) et de Schwartz (SC)

Tableau IV. 1: Choix du retard p

p	AIC	SC
1	76.09	77.07
2	76.81	78.59
3	75.43	78.01
4	71.14	74.52

La comparaison des modèles suivant les deux critères d'Akaike et de Schwartz, montre qu'au décalage $p=4$ ces deux critères sont minimaux, donc le modèle VAR (4) pourrait convenir ($p=4$).

2.1 Validation du modèle VAR(4) :

Les racines des polynômes autorégressifs du modèle VAR(4) ne sont pas toutes supérieures à 1 en module, car leurs inverses calculés par *EViews* ne sont pas tous inférieurs à 1.

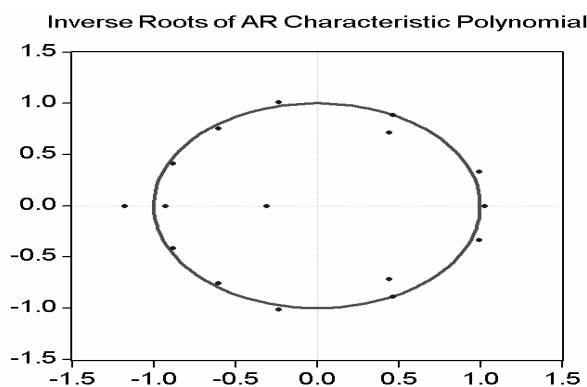


Figure IV. 1: Racines inverses des polynômes autorégressifs du modèle VAR(4)

Les conditions de stationnarité et d'inversibilité ne sont pas vérifiées. Le modèle est non stationnaire, il n'est donc pas accepté.

On revient à la comparaison des modèles restants en prenant le modèle qui minimise les deux critères (AIC et SC), soit le modèle VAR (3).

2.2 Validation du modèle VAR(3) :

Après avoir testé les racines du modèle Var (3), nous avons remarqué que les racines des polynômes autorégressifs des séries ne sont pas toutes supérieures à 1 en module, car leurs inverses calculés par EVIEWS ne sont pas tous inférieurs à 1.

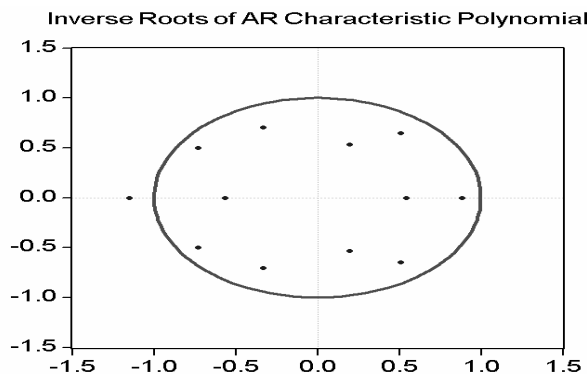


Figure IV. 2 : Racines inverses des polynômes autorégressifs du modèle VAR(3)

Les conditions de stationnarité et d'inversibilité ne sont pas vérifiées. Le modèle VAR est non stationnaire, il est donc rejeté.

On revient encore une fois à la comparaison des modèles restants en prenant le modèle qui minimise les deux critères (AIC et SC), soit le modèle VAR (1)

2.3 Validation du modèle VAR(1) :

- **Test sur les racines :**

Les racines du modèle VAR(1) sont tous supérieures à 1 en module, car leurs inverses calculés par Eviews sont tous inférieurs à 1.

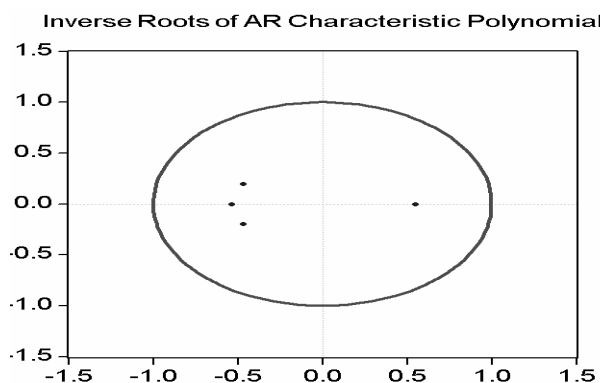


Figure IV. 3 : Racines inverses des polynômes autorégressifs du modèle VAR(1)

Les conditions de stationnarité et d'inversibilité sont vérifiées. Le modèle VAR(1) est stationnaire, il est donc retenu.

Après avoir déterminé le modèle VAR optimal, il s'agira de vérifier la blancheur de ses résidus.

• **Test sur les résidus :**

Pour vérifier la blancheur des résidus nous allons analyser les corrélogrammes des résidus de chaque série composant le VAR.

Cogramme des résidus de la série DDC_{GPLc} :

Le corrélogramme des résidus de la série DDC_{GPLc} est donné par la *figure IV.4*:

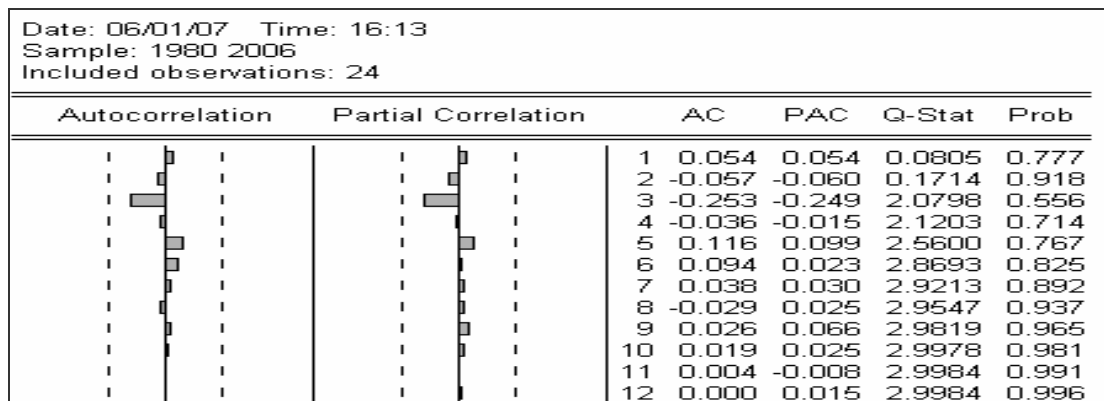


Figure IV. 4 : Corrélogramme des résidus de la série DDC_{GPLc}

L'analyse du corrélogramme des résidus, montre que tous les termes sont à l'intérieur de l'intervalle de confiance, donc les résidus forment un bruit blanc.

Ceci est confirmé par la statistique de Ljung- Box (Q - stat) qui est inférieure à la valeur théorique de $\chi^2(h)$ quelque soit le retard h , et spécialement pour $h=12$ on a :

Q - stat (12) = 2.9984 inférieure à $\chi^2(12) = 21.026$ lue au seuil 0.05, donc les résidus de la série DDC_{GPLc} forment un bruit blanc.

Corrélogramme des résidus de la série DC_{ess} :

Le corrélogramme des résidus de la série DC_{ess} est donné par la *figure IV.5*:

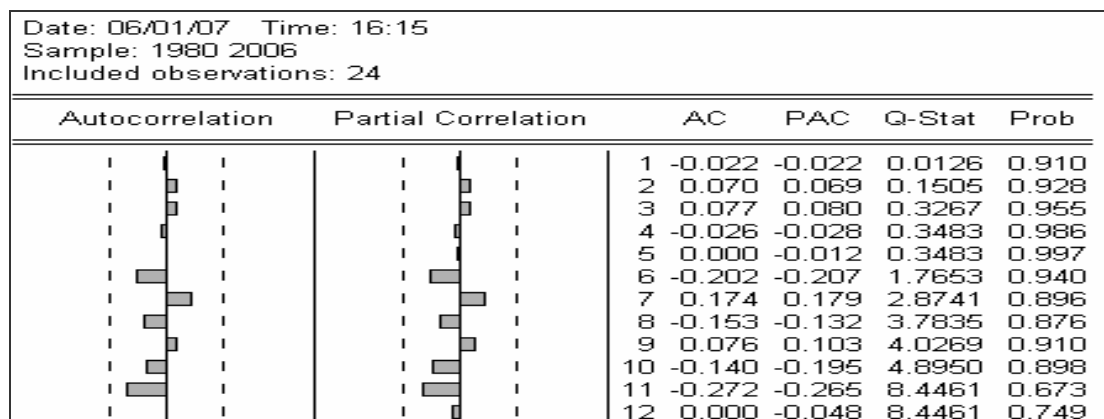


Figure IV. 5 : Corrélogramme des résidus de la série DC_{ess}

L'analyse du corrélogramme des résidus, montre que tous les termes sont à l'intérieur de l'intervalle de confiance, donc les résidus forment un bruit blanc

Ceci est confirmé par la statistique de Ljung- Box (Q - stat) qui est inférieure à la valeur théorique de $\chi^2(h)$ quelque soit le retard h , et spécialement pour $h=12$ on a :

Q - stat (12) = 8.4461 inférieure à χ^2 (12) = 21.026 lue au seuil 0.05, donc les résidus de la série DC_{ess} forment un bruit blanc.

Corrélogramme des résidus de la série DDC_{gas} :

Le corrélogramme des résidus de la série DDC_{gas} est donné par la figure *figure IV.6*:

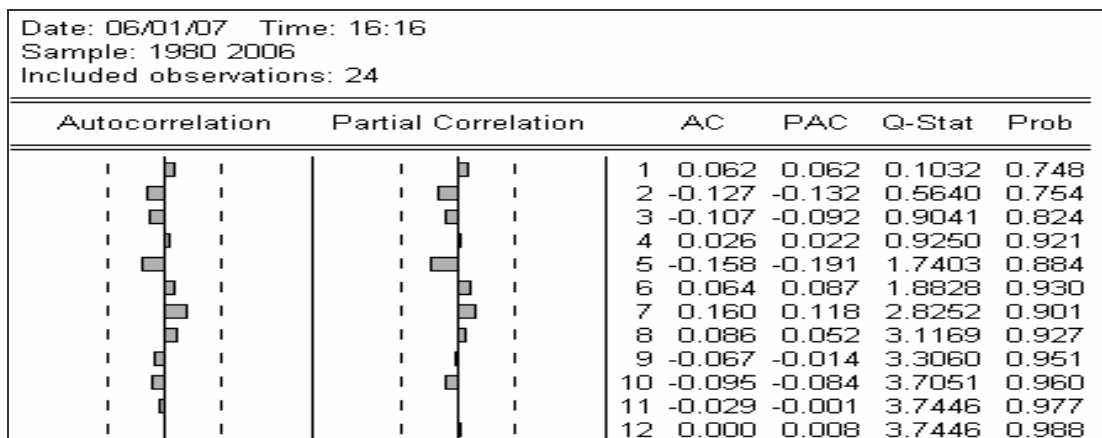


Figure IV. 6 : Corrélogramme des résidus de la série DDC_{gas}

L'analyse du corrélogramme des résidus montre que tous les termes sont à l'intérieur de l'intervalle de confiance, donc les résidus forment un bruit blanc

Ceci est confirmé par la statistique de Ljung- Box (Q - stat) qui est inférieure à la valeur théorique de χ^2 (h) quelque soit le retard h, et spécialement pour h= 12 on a :

Q - stat (12) = 3.7446 inférieure à χ^2 (12) = 21.026 lue au seuil 0.05, donc les résidus de la série DDC_{gas} forment un bruit blanc.

Corrélogramme des résidus de la série DDP_{gas} :

Le corrélogramme des résidus de la série DDP_{gas} est donné par la *figure IV.7* :

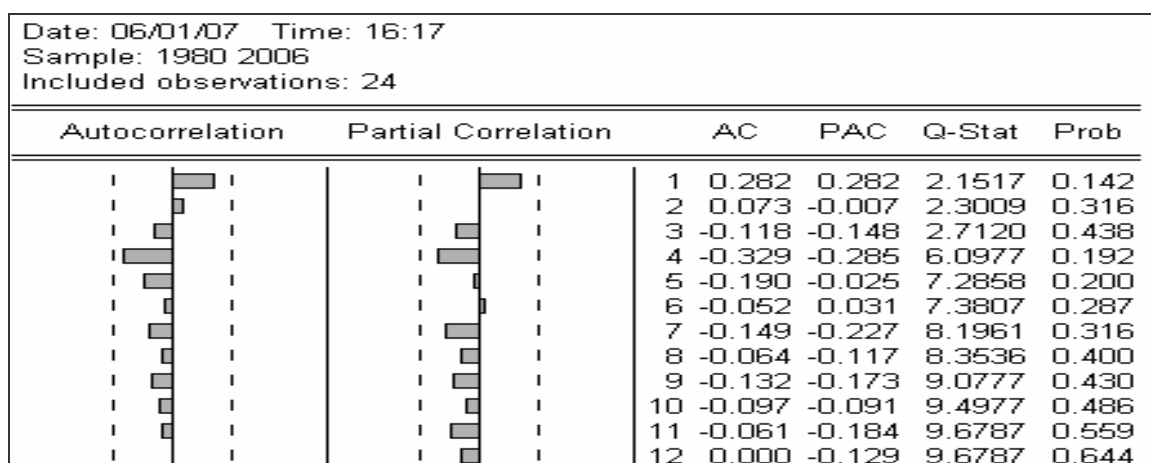


Figure IV. 7 : Corrélogramme des résidus de la série DDP_{ess}

L'analyse du corrélogramme des résidus, montre que tous les termes sont à l'intérieur de l'intervalle de confiance, donc les résidus forment un bruit blanc

Ceci est confirmé par la statistique de Ljung- Box qui est inférieure à la valeur théorique de $\chi^2(h)$ quelque soit le retard h , et spécialement pour $h=12$ on a :

Q -stat (12) = 9.6787 inférieure à $\chi^2(12) = 21.026$ lue au seuil 0.05, donc les résidus de la série DDP_{gas} forment un bruit blanc.

Sur la base de l'étude des corrélogrammes simples des résidus, nous pouvons conclure que les résidus associés au modèle $VAR(I)$ retenu forment un processus bruit blanc. Donc le modèle VAR(1) est validé.

Il s'écrit sous sa forme standard:

$$DDP_{gas}(t) = -0,79 * DDP_{gas}(t-1) - 3,65 E^{-5} * DDC_{GPLC}(t-1) + 3,26 E^{-6} * DC_{ess}(t-1) - 1,17 E^{-7} * DDC_{gas}(t-1) + 8 E^{-3} \dots\dots\dots(1)$$

$$DDC_{GPLC} = 5646,46 * DDP_{gas}(t-1) - 0,04 * DDC_{GPLC}(t-1) - 0,01 * DC_{ess}(t-1) + 7 E^{-3} * DDC_{gas}(t-1) + 670,80 \dots\dots\dots(2)$$

$$DC_{ess} = -2715,17 * DDP_{gas}(t-1) + 1,74 * DDC_{GPLC}(t-1) + 0,51 * DC_{ess}(t-1) - 0,14 * DDC_{gas}(t-1) + 7806,9 \dots\dots\dots(3)$$

$$DDC_{gas} = 1715,22 * DDP_{gas}(t-1) + 4,12 * DDC_{GPLC}(t-1) - 0,43 * DC_{ess}(t-1) - 0,61 * DDC_{gas}(t-1) + 23693,72 \dots\dots\dots(4)$$

VAR(1) s'écrit aussi sous sa forme matricielle :

$$\begin{pmatrix} DDP_{gas,t} \\ DDC_{GPLC,t} \\ DC_{ess,t} \\ DDC_{gas,t} \end{pmatrix} = \begin{pmatrix} -0,79 & -3,65 E^{-5} & 3,26 E^{-6} & -1,17 E^{-7} \\ 5646,46 & -0,04 & -0,01 & 7,00 E^{-3} \\ -2715,17 & 1,74 & 0,51 & -0,14 \\ 1715,22 & 4,12 & -0,43 & 0,61 \end{pmatrix} \begin{pmatrix} DDP_{gas,t} \\ DDC_{GPLC,t} \\ DC_{ess,t} \\ DDC_{gas,t} \end{pmatrix} + \begin{pmatrix} 8,00 E^{-3} \\ 670,80 \\ 7806,9 \\ 23693,72 \end{pmatrix}$$

Notre principal objectif est d'analyser l'effet du prix de gasoil sur les consommations des différents carburants. Cela est possible à travers les équations (2), (3) et (4) où la variable DDP_{gas} est exogène. Il est envisageable par ailleurs de prendre en compte les possibilités de feedback sur le prix de gasoil à travers l'influence des différentes consommations dans la première équation où le prix est une variable endogène.

Le modèle VAR(1) étant correctement spécifié, nous pouvons l'utiliser pour appréhender les relations dynamiques existant entre les variables qui le composent, à savoir: DDC_{GPLC} , DC_{ess} , DDC_{gas} , DDP_{gas} à l'aide des trois instruments d'analyse à savoir :

- Le test de causalité de Granger ;
- La décomposition de la variance de l'erreur de prévision ;
- Les fonctions de réponses aux chocs ;

II. Analyse des relations dynamiques entre les variables :

1. Test de causalité de Granger :

On entend par causalité, une variable qui a un impact différé sur une autre variable. Le test de causalité de Granger permet de distinguer les sens de causalité pertinents dans les équations du VAR et d'affiner cette analyse.

Les résultats du test fournis directement par *EViews* sont donnés par le tableau IV.2 :

Tableau IV. 2 : Test de Granger

Pairwise Granger Causality Tests
Date: 06/01/07 Time: 16:27
Sample: 1980 2006

Null Hypothesis:	Obs	F-Statistic	Probability
H_0^1 : DDPGAS does not Granger Cause DDCGPL	24	11.2721	0.06298
H_0^2 : DDCGPL does not Granger Cause DDPGAS		6.27509	0.02055
H_0^1 : DDPGAS does not Granger Cause DCESS	24	0.00739	0.93229
H_0^2 : DCESS does not Granger Cause DDPGAS		4.74575	0.04091
H_0^1 : DDCGAS does not Granger Cause DCESS	24	2.63309	0.11958
H_0^2 : DCESS does not Granger Cause DDCGAS		1.85473	0.18767
H_0^1 : DCESS does not Granger Cause DDCGPL	24	1.22580	0.28075
H_0^2 : DDCGPL does not Granger Cause DCESS		0.72696	0.40349
H_0^1 : DDCGAS does not Granger Cause DDCGPL	24	0.20137	0.65821
H_0^2 : DDCGPL does not Granger Cause DDCGAS		2.03826	0.16809
H_0^1 : DDPGAS does not Granger Cause DDCGAS	24	0.05893	0.81055
H_0^2 : DDCGAS does not Granger Cause DDPGAS		0.00898	0.92541

Interprétation :

D'après le tableau IV.2:

On remarque l'existence d'une causalité au sens de Granger entre le prix du gasoil et la consommation du GPLc, le prix du gasoil et la consommation de l'essence.

Il n'en est pas de même pour la consommation du gasoil et le prix de ce dernier pour lesquels on remarque une absence de causalité au sens de Granger. En effet :

- Dans la 1^{ème} partie du tableau, nous rejetons l'hypothèse H_0^2 : et donc *le prix du diesel cause au sens de Granger la consommation du GPLc*, la probabilité associée est de 0.02055 et elle est inférieure au seuil statistique usuel de 5%, il n'y a donc causalité au sens de Granger de DDP_{gas} vers DDC_{GPLc} , on conclut donc que DDP_{gas} explique significativement la variable DDC_{GPLc} .
- Dans la 2^{ème} partie du tableau nous rejetons l'hypothèse H_0^2 : et donc : *le prix du diesel cause au sens de Granger la consommation de l'essence* car la probabilité associée est de 0.04091, elle est inférieure au seuil statistique usuel de 5%, il y a donc causalité au sens de Granger de DDP_{gas} vers DDC_{ess} . On conclut donc que DDP_{gas} explique significativement la variable DDC_{ess} .
- Cependant, dans la 3^{ème} partie du tableau, nous acceptons l'hypothèse H_0^2 : c'est à dire que le prix du diesel ne cause pas au sens de Granger la consommation de ce dernier, en effet la probabilité associée est de 0.92541, cette dernière est supérieure au seuil statistique usuel de 5%, il n'y a donc pas de causalité au sens de Granger entre le prix du diesel et la consommation de ce dernier.

De la même manière, on remarque une absence de causalité au sens de Granger entre les trois consommations à savoir : la Consommation du GPLc, la consommation de l'essence et la consommation du gasoil. Aucune consommation n'explique significativement l'autre, car les probabilités associées sont toutes supérieures au seuil usuel 5%.

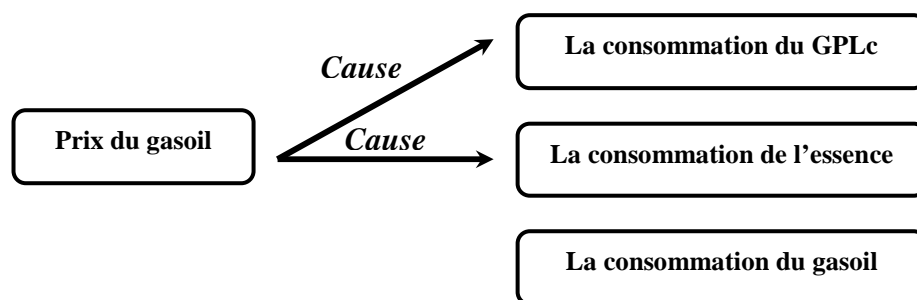


Figure IV. 8 : La causalité entre les variables selon Granger

En conclusion (figure IV.8), il ressort de ce test:

- Une absence de causalité entre les trois consommations, aucune consommation n'explique significativement l'autre.
- Présence d'une causalité entre le prix du gasoil et la consommation du GPLc, le prix du gasoil et la consommation de l'essence.
- Absence de causalité entre le prix du gasoil et la consommation de ce dernier.

Il apparaît aussi dans le test de causalité de Granger que les deux équations précédentes (2) et (3) du modèle VAR(1) sont les plus pertinentes à interpréter :

* Equation (1) :

$$DDP_{\text{gas}}(t) = -0,79 * DDP_{\text{gas}}(t-1) - 3,65 E^{-5} * DDC_{\text{GPLc}}(t-1) + 3,26 E^{-6} * DC_{\text{ess}}(t-1) - 1,17 E^{-7} * DDC_{\text{gas}}(t-1) + 8 E^{-3} \dots\dots\dots(1)$$

Absence d'effet de feedback sur le prix de gasoil (les consommations ne causent pas le prix du gasoil).

* Equation (2) :

$$DDC_{\text{GPLc}} = 5646,46 * DDP_{\text{gas}}(t-1) - 0,04 * DDC_{\text{GPLc}}(t-1) - 0,01 * DC_{\text{ess}}(t-1) + 7 E^{-3} * DDC_{\text{gas}}(t-1) + 670,80 \dots\dots\dots(2)$$

Causalité affirmée par le test de Granger, l'effet du prix de gasoil sur la consommation du GPLc est positif.

* Equation (3) :

$$DC_{\text{ess}} = -2715,17 * DDP_{\text{gas}}(t-1) + 1,74 * DDC_{\text{GPLc}}(t-1) + 0,51 * DC_{\text{ess}}(t-1) - 0,14 * DDC_{\text{gas}}(t-1) + 7806,9 \dots\dots\dots(3)$$

Causalité affirmée par le test Granger, l'effet du prix du gasoil sur la consommation de l'essence est négatif, ce qui n'est pas un signe attendu théoriquement, mais qui correspond bien à ce qui a été observé dans le passé. (Figure IV.9)

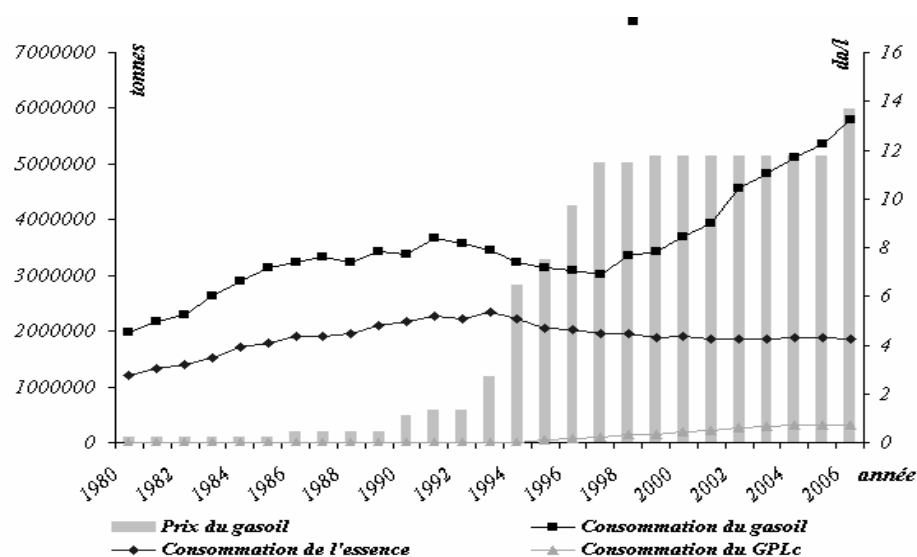


Figure IV. 9 : Evolution historique des consommations de carburants avec le prix du gasoil [MEM]

* Equation (4) :

$$DDC_{\text{gas}} = 1715,22 * DDP_{\text{gas}}(t-1) + 4,12 * DDC_{\text{GPLc}}(t-1) - 0,43 * DC_{\text{ess}}(t-1) - 0,61 * DDC_{\text{gas}}(t-1) + 23693,72 \dots\dots\dots(4)$$

La causalité n'a pas été affirmée par le test de Granger. On peut remarquer quand même que le signe du prix de gasoil est positif, ce qui contredit la théorie économique (le prix devrait

avoir un impact négatif sur la consommation). Cependant ce signe correspond bien à ce qui a été observé par le passé

2 L'analyse impulsionnelle :

Le paragraphe suivant se propose d'affiner les premiers résultats fournis par le test de Granger par une analyse impulsionnelle, en effet :

Le test de causalité au sens de Granger effectué précédemment repose principalement sur une notion d'antériorité entre les variables et peut s'interpréter comme un test de causalité intra- échantillon. Ce test ne peut révéler que la causalité au sein de l'échantillon, mais il ne fournit aucune information sur les propriétés dynamiques du système de variables considéré.

En revanche, **l'exercice de la décomposition de la variance de l'erreur de prévision** ainsi que **l'exercice d'analyse des fonctions de réponses impulsionnelles** peuvent quant à eux être considérés comme des tests de causalité hors échantillon, étant donné qu'ils ont pour objectif :

- de décomposer la variance de l'erreur de prévision d'une variable donnée pour différents horizons de prévisions, en parts attribuables aux innovations (ou chocs) dues à chaque variable du système, y compris elle-même.
- d'identifier et de comprendre les sources des fluctuations des différentes variables en illustrant l'impact de la variation d'une innovation sur ces dernières.

Nous nous proposons dans ce qui suit d'utiliser ces deux instruments afin d'appréhender les relations dynamiques existant entre les variables qui composent le modèle VAR(1).

2.1 La décomposition de la variance de l'erreur de prévision:

La décomposition de la variance permet d'évaluer pour plusieurs horizons temporels, l'importance relative des différents chocs sur les fluctuations des variables du modèle.

En effet, la décomposition de la variance de l'erreur de prévision a pour objectif de calculer pour chacune des innovations (chocs), sa contribution à la variance de l'erreur.

La décomposition de la variance pour chaque variable se fera sur une période de 10 ans, les résultats sont donnés par les tableaux ci-dessous :

- **Décomposition de la variance de l'erreur de prévision de DDC_{GPLc} :**

La décomposition de la variance de l'erreur de prévision de la variable DDC_{GPLc} est donnée par le tableau IV.3:

Tableau IV. 3 : Décomposition de la variance de la variable DDC_{GPLc}

Period	S.E.	DDCGPL	DCESS	DDCGAS	DDPGAS
1	8537.383	87.28843	0.000000	0.000000	12.71157
2	10012.43	63.46499	1.033176	1.502983	33.99886
3	11227.72	52.92443	1.065826	1.439968	44.56978
4	11686.45	50.47821	1.128427	1.329205	47.06416
5	11787.16	50.15068	1.125881	1.384028	47.33941
6	11803.52	50.10976	1.127847	1.476546	47.28585
7	11808.24	50.07674	1.126951	1.544175	47.25214
8	11811.54	50.04900	1.126324	1.575813	47.24886
9	11813.61	50.03318	1.126060	1.588171	47.25258
10	11814.53	50.02689	1.125920	1.592016	47.25517

La décomposition de la variance indique que la variance de l'erreur de prévision de **la consommation du GPLc** est due à : **50%** à ses propres innovations, **47.25 %** à celles du prix du gasoil, **1.12 %** à celles de la consommation du gasoil et **1.59 %** à celles de la consommation de l'essence.

Ainsi l'effet d'un choc au niveau de la consommation du gasoil et au niveau de la consommation de l'essence n'a qu'une influence très peu significative sur la consommation du GPLc.

Ce qui est remarquable, c'est plutôt l'effet d'un choc au niveau du prix du diesel sur la consommation du GPLc. Cet effet prend de l'importance à mesure que l'on s'éloigne de l'instant de sa survenance, il a une influence de 12% environ sur un horizon de 1 an, qui passe à 44 % environ sur un horizon de 2- 3 ans. Ceci nous amène à penser que l'impact des chocs au niveau du prix du gasoil sur la consommation du GPLc est retardé et qu'il faut attendre quelques années pour voir son effet agir.

- **Décomposition de la variance de l'erreur de prévision de DC_{ess} :**

La décomposition de la variance de l'erreur de prévision de DC_{ess} est donnée par le tableau IV.4

Tableau IV. 4 : Décomposition de la variance de la variable DC_{ess}

Period	S.E.	DDCGPL	DCESS	DDCGAS	DDPGAS
1	78258.55	1.650439	77.17859	0.000000	21.17097
2	90043.16	1.690213	73.25469	6.982098	18.07300
3	92317.47	1.611039	74.29894	6.882247	17.20778
4	93996.28	1.565988	73.06904	7.630486	17.73449
5	94254.46	1.597597	73.03148	7.643762	17.72716
6	94478.38	1.606033	72.82301	7.711519	17.85944
7	94509.05	1.615024	72.80680	7.711197	17.86698
8	94530.32	1.616898	72.78677	7.715634	17.88070
9	94532.85	1.617824	72.78576	7.715394	17.88102
10	94534.33	1.617955	72.78460	7.715632	17.88182

La décomposition de la variance indique que la variance de l'erreur de prévision de **la consommation de l'essence** est due à : **72.78%** à ses propres innovations, **17.88 %** à celles du prix du gasoil, **1.61 %** à celles de la consommation du GPLc et **7.71 %** à celles de la consommation du gasoil.

Ainsi l'effet d'un choc au niveau de la consommation du diesel et au niveau de la consommation du GPLc a très peu d'influence sur la consommation de l'essence, contrairement à un choc au niveau du prix du gasoil, dont l'effet sur la consommation de l'essence est significatif et remarquable de par son caractère permanent et constant (17%) à mesure que l'on s'éloigne de l'instant de sa provenance, restant presque du même ordre.

Ceci nous amène à penser que l'impact des chocs au niveau du prix du gasoil a son effet sur la consommation de l'essence au moment de sa provenance, son influence reste presque du même ordre au cours du temps.

- **Décomposition de la variance de l'erreur de prévision de DDC_{gas}:**

Tableau IV. 5 : Décomposition de la variance de la variable DC_{gas}

Period	S.E.	DDCGPL	DCESS	DDCGAS	DDPGAS
1	176747.9	1.303074	0.033186	97.64766	1.016085
2	214229.2	2.240746	2.235776	91.25592	4.267554
3	229605.7	2.601274	1.946831	91.73678	3.715111
4	236598.4	2.482963	1.913191	91.52292	4.080923
5	240994.4	2.416817	1.854654	90.56199	5.166539
6	243164.9	2.445803	1.821861	89.79615	5.936185
7	244223.1	2.485181	1.810728	89.32879	6.375304
8	244614.3	2.511154	1.805105	89.13440	6.549340
9	244752.7	2.522457	1.803764	89.06091	6.612866
10	244791.7	2.526608	1.803215	89.03921	6.630967

La décomposition de la variance indique que la variance de l'erreur de prévision de **la consommation du Diesel** est due à : **89%** à ses propres innovations, **2.52 %** à celles de la consommation du GPLc, **1.80 %** à celle de la consommation de l'essence et **6.63 %** à celle de la consommation du gasoil.

Ainsi, l'effet d'un choc, au niveau de la consommation du GPLc, de l'essence ainsi qu'au niveau du prix du gasoil, a une influence peu significative sur la consommation de ce dernier.

- **Décomposition de la variance de l'erreur de DDP_{gas}:**

La décomposition de la variance de DDP_{gas} est donnée par le tableau ci-dessous :

Tableau IV. 6 : Décomposition de la variance de la variable DDP_{gas}

Period	S.E.	DDCGPL	DCESS	DDCGAS	DDPGAS
1	0.829955	0.000000	0.000000	0.000000	100.0000
2	1.274278	6.574335	3.087322	0.025907	90.31244
3	1.388211	9.484380	2.626970	0.478075	87.41057
4	1.415410	10.23930	2.717951	0.955311	86.08744
5	1.420579	10.34145	2.699206	1.469520	85.48982
6	1.422825	10.31375	2.695334	1.719866	85.27105
7	1.424613	10.29165	2.690903	1.841908	85.17554
8	1.425483	10.28732	2.687622	1.881539	85.14352
9	1.425882	10.28709	2.686655	1.894902	85.13135
10	1.426009	10.28782	2.686183	1.898151	85.12785

La décomposition de la variance indique que la variance de l'erreur de prévision de **la consommation du Diesel** est due à : **85.12%** à ses propres innovations, **10.28 %** à celles de la consommation du GPLc, **2.68 %** à celle de la consommation de l'essence et **1.89 %** à celle de la consommation du gasoil.

Ainsi l'effet des chocs au niveau de la consommation du GPLc, au niveau de la consommation de l'essence, et au niveau de la consommation du Diesel, à une influence peu significative sur le prix de ce dernier.

Nous remarquons que les résultats de l'exercice de l'analyse de la décomposition de la variance de l'erreur de prévision viennent confirmer les résultats déjà obtenus dans l'exercice précédent à savoir : l'analyse de causalité de Granger.

Si la décomposition de la variance donne déjà des indications sur les sources des fluctuations des variables, l'analyse des fonctions de réponses impulsionnelles qui va suivre permettra une meilleure interprétation des résultats.

2.2 Les fonctions de réponses impulsionnelles:

L'analyse des chocs cherche à prévoir le comportement des différentes variables considérées sur un horizon donné, en se basant sur les comportements observés par le passé.

Cette analyse se base sur l'hypothèse de système économique stable, c'est-à-dire pas de changement structurel important pouvant influencer les relations et les causalités établies dans le passé.

L'objectif de cette partie est donc d'obtenir des indications sur les sources des fluctuations des différentes variables. Les graphiques ci-dessous, retracent les fonctions impulsionnelles, nous nous intéressons aux effets de chocs sur 10 périodes.

Ces graphiques révèlent que le modèle estimé VAR(1) possède une capacité rapide de résorption des chocs : ceux-ci sont en effet totalement absorbés au bout de 10 périodes après avoir vu leur impact décroître de manière quasi linéaire. Cela confirme la stabilité du modèle.

- **L'effet d'un choc au niveau de la consommation de l'essence :**

L'effet d'un choc au niveau de la consommation de l'essence sur les autres variables est représenté par la figure IV.10 :

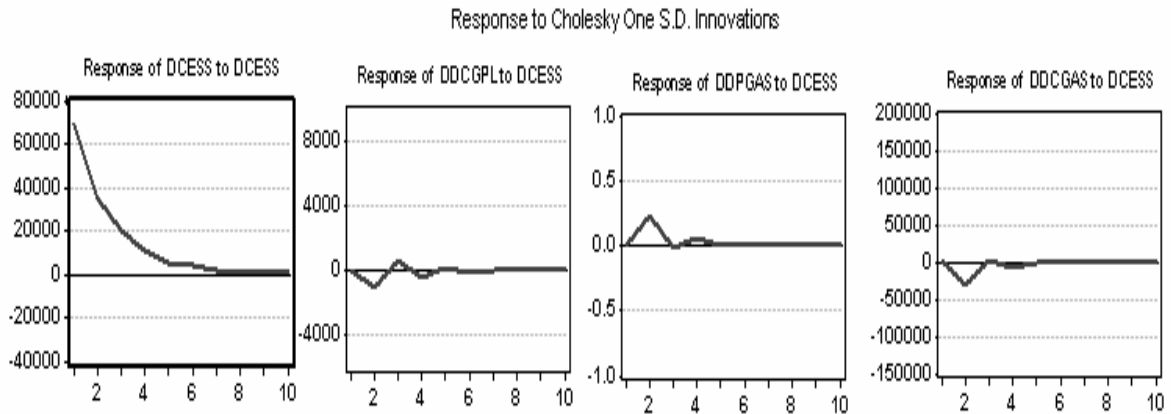


Figure IV. 10 : Effet d'un choc au niveau de la consommation de l'essence

Il ressort de cette figure, qu'un choc au niveau de la consommation de l'essence semble n'avoir qu'un impact infime et d'ailleurs peu significatif sur la consommation du GPLc, celle du gasoil ainsi que sur le prix de ce dernier.

- **L'effet d'un choc au niveau de la consommation du GPLc:**

L'effet d'un choc au niveau de la consommation du GPLc sur les autres variables est représenté par la figure IV.11

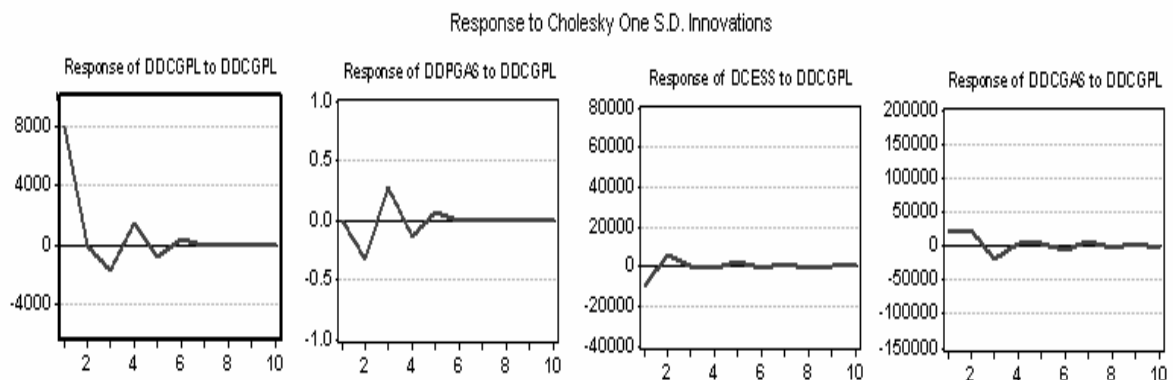


Figure IV. 11: Effets d'un choc au niveau de la consommation du GPLc

On remarque, là aussi, qu'un choc sur la consommation du GPLc semble avoir un impact qu'on pourrait qualifier d'infime et non significatif sur la consommation de l'essence, celle du gasoil ainsi que sur le prix de ce dernier.

- **L'effet d'un choc au niveau de la consommation du gasoil:**

L'effet d'un choc au niveau de la consommation du gasoil sur les autres variables est représenté par les graphiques suivants :

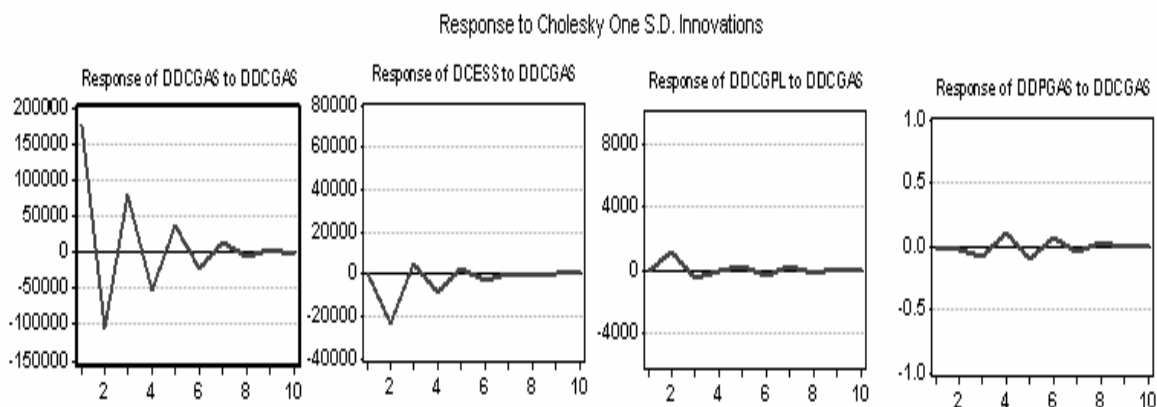


Figure IV. 12 : Effets d'un choc au niveau de la consommation du gasoil

De même, on remarque qu'un choc au niveau de la consommation du gasoil semble n'avoir qu'un impact faible sur la consommation de l'essence, sur la consommation du GPLc, ainsi que sur le prix du gasoil.

Nous nous intéressons en particulier à l'effet d'un choc aux niveau du prix du gasoil sur les autres variables pour rester en cohérence avec les résultats obtenus avec le test de causalité.

- **L'effet d'un choc au niveau du prix du gasoil :**

L'effet d'un choc au niveau du prix du gasoil sur les autres variables est représenté par le graphique ci-dessous :

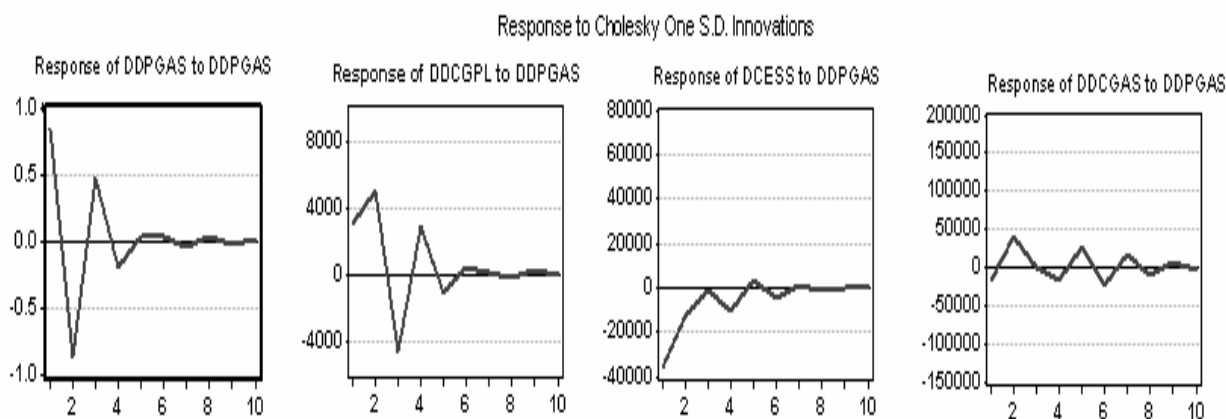


Figure IV. 13 : Effets d'un choc au niveau de prix du gasoil

Interprétations :

On peut remarquer qu'un choc sur la variable 'prix de gasoil' différencié 2 fois, traduisant **l'accélération de la croissance des prix**, engendre un fort ajustement dès l'année suivante, avec une accélération négative (décélération) du prix.

Cela apparaît clairement dans le graphique ci-dessous, où une accélération est corrigée rapidement par une décélération (figure IV.14)

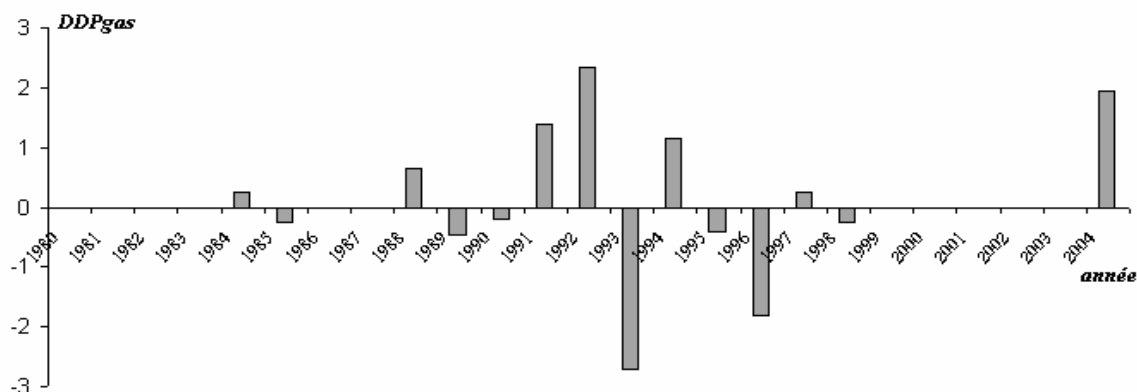


Figure IV. 14 : L'accélération de la croissance du prix du gasoil

La décomposition de la variance de l'erreur de DDP_{gas} montre effectivement que l'erreur de prévision est en grande partie due aux innovations du prix du gasoil.

L'absorption du choc sur le prix du gasoil se traduit par le retour vers l'équilibre représenté par une accélération nulle des prix, au bout de 5 périodes.

- **L'impact sur la consommation du GPLc :**

La consommation du GPLc répond instantanément au prix de gasoil. En effet, un choc sur la variable 'prix de gasoil' se traduit par à une accélération de la consommation de GPLc. Cet impact significatif a été confirmé par les tests de causalité et la décomposition de la variance.

Cette réponse de la consommation du GPLc a été effectivement observée historiquement avec la forte pénétration du GPLc, concordant avec la forte augmentation des prix pendant le plan d'ajustement structurel, imposée à l'Algérie à partir de 1994. (figure IV.15).

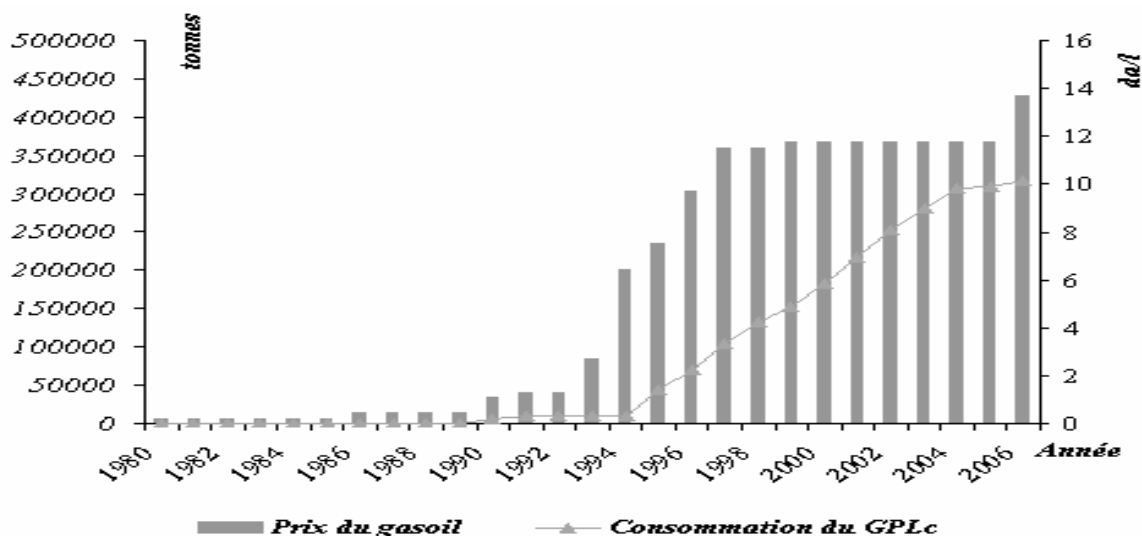


Figure IV. 15 : Evolution historique de la consommation du GPLc avec le prix du gasoil

- **L'impact sur la consommation de l'essence :**

La consommation de l'essence répond instantanément au prix de gasoil, puisque un choc sur la variable du prix du gasoil se traduit par une baisse de la consommation d'essence. Cet impact significatif a été confirmé par les tests de causalité et la décomposition de la variance.

Cette variation négative de la consommation d'essence en réponse à un choc de prix de gasoil n'était pas attendue, mais elle se justifie dans les séries passées observées, illustrées par le graphique (figure IV.16). En effet, cette réponse de la consommation de l'essence a été effectivement observée historiquement.

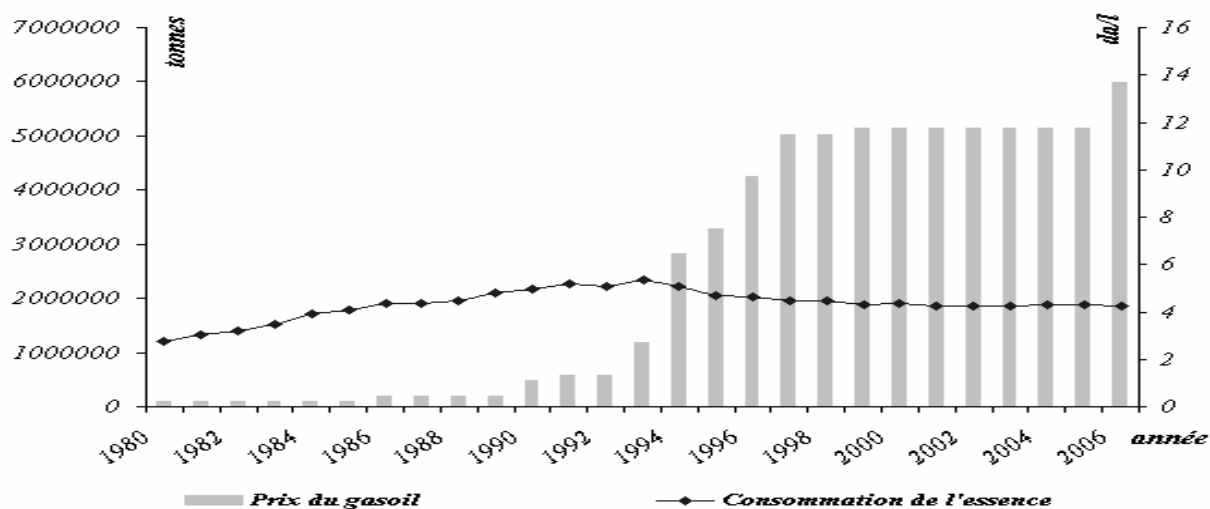


Figure IV. 16 : Evolution historique de la consommation de l'essence avec le prix du gasoil

Bien que la causalité soit validée statistiquement, nous pensons que la baisse de la consommation d'essence en réponse à un choc de prix de gasoil n'est pas liée directement à la hausse des prix, mais elle reproduit plutôt dans le futur les répercussions de la crise économique pendant les années 90 et à la forte diésélisation du parc automobile observée dans le passé (l'analyse des chocs exploite les informations du passé pour les reproduire dans le futur).

- **L'impact sur la consommation du gasoil :**

D'après le graphique gasoil et conformément au test de causalité, l'impact d'un choc de prix de gasoil sur la consommation de gasoil n'est pas très significatif relativement aux autres variables. C'est un résultat qui contredit la théorie économique, qui stipule une réaction plutôt forte et négative de la consommation au prix.

On peut lier le résultat de non causalité des prix de gasoil et du faible impact d'un choc sur le prix de gasoil aux raisons suivantes :

- La série gasoil inclut le transport de marchandises et de voyageurs qui sont captifs du carburant gasoil et qui ont été historiquement plus dépendants de variables de croissance économique et de déplacement des voyageurs que de prix.

- La forte diésélisation du parc automobile ces dernières années qui fait que la consommation de gasoil est plutôt influencée par la taille du parc automobile que par le prix.

Conclusion générale

Le GPLc revêt une importance considérable dans la stratégie d'approvisionnement en énergie du secteur des transports en Algérie. Cette importance est liée d'une part aux avantages économiques, écologiques du GPLc ainsi aux ressources disponibles et, d'autre part, à son potentiel de substitution aux carburants classiques notamment le gasoil. En effet, ce dernier fait l'objet d'une tension élevée, qui risque de peser lourdement sur les capacités de sa production dans les années à venir. Si la tendance actuelle de la consommation se maintient, il faut s'attendre d'ici les prochaines années à l'importer.

A travers cette étude, nous avons tenté de comprendre l'évolution de la consommation du carburant GPLc sur le marché national, dans un contexte caractérisé par une augmentation significative de la consommation de tous les carburants, principalement celle du gasoil.

La première étape de notre étude consistait à identifier les principales variables déterminantes, dont le pouvoir explicatif sur la consommation de GPLc a été testé à travers l'utilisation d'un modèle de régression linéaire multiple. Les variables explicatives de la consommation de GPLc obtenues dans notre modèle sont :

- Le PIB : cette variable exprime la richesse économique du pays,
- La population active : c'est une variable qui prend en compte les personnes actives qui se déplacent et qui ont des revenus pour le faire.
- Le nombre de voyageurs utilisant le transport en commun : cette variable prend en compte les déplacements par les moyens de transport collectifs, elle s'avère une variable pertinente dans notre modèle compte tenu de la part des taxis dans le transport collectifs et la part importante du GPLc comme carburant pour ces moyens de transport
- Taux de croissance du parc automobile : cette variable traduit le fait que les véhicules convertis au GPLc sont plutôt des véhicules neufs ;
- Le prix de l'essence : cette variable s'avère pertinente puisqu'elle encourage l'équipement en kit GPLc et donc la consommation de ce dernier.

Le prix de gasoil et le prix du GPLc n'étaient pas explicatifs dans notre modèle. En effet, le fait que ces prix soient administrés (constants sur une longue période) génère des problèmes statistiques d'estimation (multi colinéarité).

La deuxième étape a constitué à effectuer les prévisions des différentes variables explicatives. Ces prévisions ont été menées en utilisant la méthode de Box & Jenkins à travers des modèles de type ARMA (p, q) pour chacune des variables. Les résultats obtenus ont permis de prévoir la consommation du GPLc sur un horizon de 5 ans. Celle-ci devrait atteindre 4 M tonnes environ en 2011, marquant un taux de croissance annuel moyen de 5%. Nous noterons que cette prévision prend en compte les facteurs de demande, et ne prend pas en compte les facteurs agissant sur l'offre tel que l'effet d'une politique de promotion du GPLc.

Le modèle de régression que nous avons spécifié ne prend pas en compte un effet dynamique engendré par l'intégration de variables explicatives retardées. Cela pourrait exprimer une rigidité de la réponse de la consommation de GPLc aux variables explicatives choisies.

Pour prendre en compte l'effet de prix de gasoil, qui constitue l'élément principal de notre problématique, nous avons introduit dans la troisième partie de ce travail la modélisation multi variée, qui permet de tester les relations de causalité entre variables et de modéliser l'impact d'un choc sur le prix de gasoil. Cette modélisation permet entre autres de prendre en compte l'effet retardé de la variable prix de gasoil. Les tests de causalité de Granger et la décomposition de la variance dans le modèle VAR ont permis d'approuver l'influence significative du prix de gasoil sur la consommation du GPLc, l'analyse impulsionnelle confirme cette influence et nous renseigne sur la réponse de ces consommations à un éventuel choc sur le prix de gasoil.

Nous tenons à préciser que les prévisions et analyses ont été conduites sur la base de l'hypothèse de non changement structurel important, faisant que les modèles spécifiés sont supposés stables sur l'horizon de prévision considéré. Cette hypothèse permet de reproduire dans le futur les phénomènes observés dans le passé.

Ce que nous retenons de cette étude est que le potentiel de développement du GPLc est non négligeable, il dépend à la fois de facteurs économiques, que nous avons pu tester dans les modèles élaborés, mais aussi de facteurs comportementaux (réticences des consommateurs) que nous avons identifiés à travers le questionnaire. Concernant l'influence des prix, il ressort aussi de notre analyse que les consommateurs du GPLc arbitrent en sa faveur en fonction du prix de l'essence et du prix de gasoil. Ce dernier, joue un rôle important que nous avons pu expliciter à travers l'analyse des chocs.

Ce travail fournit donc une base de réflexion sur les déterminants de la consommation du GPLc et le potentiel de son évolution sur la base des informations antérieures. Les relations modélisées pourraient être sensiblement modifiées en cas de changement structurel.

Ce dernier serait par exemple le résultat d'une politique de libéralisation des prix ou de taxation des gasoils (l'alignement des prix de gasoil sur les autres carburants par des taxes : taxation des émissions de CO₂, promotion de carburants alternatifs au détriment des produits pétroliers, etc...), ou d'un changement de comportement des consommateurs. Celui ci se traduirait par une plus grande sensibilité au prix des carburants, une moindre préférence pour les motorisations diesel qui conduirait à une consommation plus équilibrée entre les différents carburants. Ceci pourrait constituer un levier important à la croissance de la consommation du GPL carburant.

LES ANNEXES

ANNEXE 1 : **Questions / Réponses sur le GPL c**

Nous avons recueilli les principales questions que chacun de nous peut se poser sur le GPLc et sur le fonctionnement des véhicules à bicarburation.

Les réponses sont données ci dessous:[CFBP]

Q1 : La consommation du GPL c est-elle plus élevée, par rapport aux autres carburants?

Le pouvoir calorifique du GPLc étant moins élevé au litre que celui d'un carburant conventionnel, cela entraîne une surconsommation de l'ordre de **20 %** en moyenne à performances égales, mais celle-ci est très largement compensée par un prix au litre le plus bas du marché.

Q2 : Le GPL c a-t-il une incidence sur la longévité du moteur?

L'huile moteur conserve ses propriétés plus longtemps grâce à la combustion propre du GPLc. L'indice d'octane supérieur du GPL permet un fonctionnement plus doux évitant les vibrations. Ces deux facteurs permettent d'accroître sensiblement la vie du moteur fonctionnant au GPLc, en obtenant une longévité équivalente aux moteurs diesels.

Q3 : Les performances sont-elles modifiées, lors de la conversion?

Une très légère baisse de puissance (de l'ordre de 2 à 5 %) rendue insensible lors de la conduite par une souplesse d'utilisation (couple moteur disponible à plus bas régime) et un silence exemplaire caractérisant le fonctionnement des véhicules GPL.

Q4 : Les voitures GPL c sont-elles soumises aux restrictions de circulation en cas de pollution ?

Pas en Algérie, mais en Europe Les automobiles GPLc bénéficient automatiquement de la " pastille verte " qui autorise de rouler les jours de restriction de circulation à cause du taux de pollution.

Q5 : L'accès des parkings souterrains est-il interdit ?

L'accès des parkings souterrains ouverts au public est normalement autorisé aux véhicules GPLc, à l'exception de ceux qui ne sont pas équipés de soupapes de sécurité. Malheureusement, par méconnaissance, les propriétaires des parkings souterrains n'y autorisent pas l'accès.

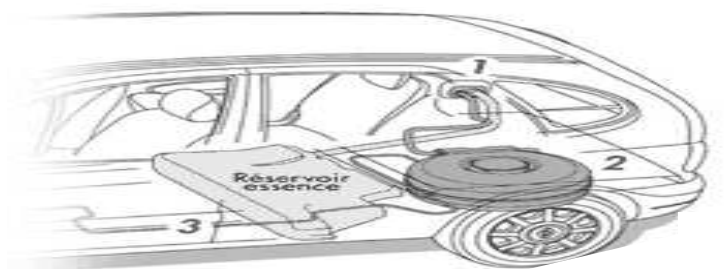
Q6 : L'installation GPL c diminue-t-elle le volume du coffre ?

De nombreux types de réservoirs GPLc sont disponibles, permettant une intégration optimum selon la morphologie du véhicule :

- **Les réservoirs toriques**, (logé en lieu et place de la roue de secours)
- **Les réservoirs cylindriques** (généralement adaptés au fond de coffre, derrière la banquette)

- **Les réservoirs « twin »**, composés de deux réservoirs cylindriques réunis (cette configuration est adaptée pour les logements rectangulaires des roues de secours ou sous châssis).
- **Les réservoirs polymorphes installés sous châssis** (optimisation du volume de GPL et de l'autonomie).

L'installation, désormais courante, de réservoirs **toriques** à la place de la roue de secours préserve entièrement le volume du coffre du véhicule GPLc.



1. Remplissage
2. Réservoir GPLc : en forme torique.
3. Ligne de gaz



Figure 1. 1 : La partie arrière d'un véhicule de motorisation GPL carburant

Q7 : Toutes les voitures peuvent-elles être transformées ?

Les automobiles fonctionnant à l'essence catalysée peuvent être transformées au GPLc. Il est également recommandé de ne procéder à l'installation d'un kit GPL que sur les voitures de moins de 60 000 km afin d'utiliser des bases mécaniques fiables.

Q8 : Une voiture GPLc a-t-elle moins d'autonomie ?

La capacité du réservoir GPLc, en règle générale, est légèrement inférieure à celle du réservoir essence mais totalement utilisable grâce au principe de la bicarburant (le réservoir d'origine étant utilisé comme réserve).

A noter : la contenance moyenne d'un réservoir GPLc permet de parcourir 400 à 600 kms.

Q9 : Une voiture GPLc est-elle rapidement amortie ?

Le GPLc est le carburant le moins cher actuellement disponible.

Par ailleurs, l'utilisation du GPLc est favorisée par un ensemble de dispositions fiscales avantageuses, tant pour les particuliers que pour les professionnels. En dépit d'une surconsommation de l'ordre de 20% dû à sa différence de pouvoir calorifique au litre par

rapport aux carburants traditionnels, l'acquisition d'un véhicule GPLc ou la transformation d'un véhicule essence peut être ainsi amortie dès les premiers kilomètres.

Q10 : Le GPLc représente-t-il un risque ?

Un réservoir GPL, réalisé en matériau très résistant (acier), est équipé de 5 accessoires contribuant à la sécurité : un clapet anti-retour, un limiteur de remplissage à 80% de la capacité totale du réservoir, un limiteur de débit, une électrovanne et une soupape de sécurité "gros débit", désormais obligatoire sur toutes les nouvelles installations GPL depuis le 1er janvier 2000. L'ensemble est construit en fonction de normes très strictes.

Q11 : A quoi sert la soupape de sécurité ?

En cas d'incendie, la chaleur intense due au feu du véhicule fait monter la pression à l'intérieur du réservoir GPLc. Comme pour une « cocotte-minute », la soupape permet de lâcher la pression vers l'extérieur et empêche le risque d'explosion.

Q12 : Est-ce que le GPLc est plus économique que le Diesel ?

Les différentes incitations et mesures fiscales mises en place pour le GPLc, tant pour les particuliers que pour les professionnels, et le faible coût d'entretien des véhicules GPL comparé aux versions diesel équivalentes, lui confère quasiment toujours un net avantage économique. Sans parler écologie, silence et propreté pour faire le plein.

Q13 : Quelles différences y a-t-il entre le GPLc et le GNV (Gaz Naturel Véhicule) ?

Le GPLc et le GNV sont 2 gaz issus des champs de gaz et de pétrole, mais dont les caractéristiques sont très différentes. Le GPLc peut être stocké facilement à l'état liquide à environ 5 bar pour une température extérieure de 15°C, alors que le GNV doit être comprimé à au moins 200 bar pour être contenu dans un réservoir automobile.

Pour un volume extérieur identique un réservoir GPLc assure une autonomie environ 1,5 fois supérieure à celle procurée par un réservoir GNV et le temps de remplissage en station-service est 2 à 3 fois moins long pour le GPL c.

Q14 : Tout garagiste peut-il installer un kit GPL ?

NON, l'installation doit impérativement être effectuée par un professionnel agréé par l'état.

L'équipement du véhicule GPLc se fait par l'installation d'un réservoir spécifique et d'un kit de transformation GPL qui doivent répondre à des réglementations très strictes.

[Décret n°83 - 496 du 13 Août 1983 relatif aux conditions d'utilisation et de distribution du GPL comme carburant sur les véhicules automobiles]

Q15 : Existe-t-il des voitures GPLc de série ?

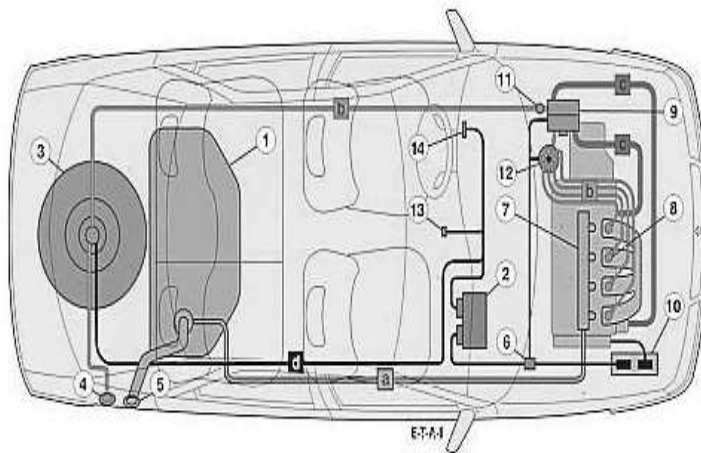
La plupart des constructeurs automobiles proposent dans leurs catalogues, des voitures GPLc. Les modèles commercialisés bénéficient tous d'équipements étudiés et adaptés à chaque véhicule, et de la garantie du constructeur.

Q16 : Le GPLc a-t-il tendance à détériorer prématurément les soupapes du moteur ?

La différence de température de combustion du GPL dans un moteur essence étant plus élevée, les soupapes de certains moteurs fragiles peuvent effectivement devoir être réglées plus souvent qu'en fonctionnement avec de l'essence. C'est pourquoi, pour ces moteurs « fragiles », les renforcements de sièges de soupapes sont proposés. Mais la généralisation des moteurs 16 soupapes et le renforcement de ces soupapes sur la plupart des moteurs actuels ne posent plus de problèmes.

Q17 : Comment est faite une installation GPL sur un véhicule ?

Exemple d'une bicarburation (GPLc + essence) sur une CITROËN XANTIA break : [OUTF]



Bleu:

Circuit essence

Vert:

Circuit GPL

Rouge:

Circuit réchauffage du vapo-détendeur

Noir:

Circuit électrique GPL

1/ Réservoir d'essence: 65 litres

2/ Calculateur GPL

3/ Réservoir GPL: 57 litres

4/ Remplissage GPL

5/ Remplissage essence

6/ Boîtier de relais de commande GPL

7/ Rampe d'injecteurs d'essence

8/ Diffuseurs GPL

9/ Vapo-détendeur (0,96 bar en sortie)

10/ Intercommunication calculateurs GPL/essence

11/ Electrovanne d'entrée

12/ Distributeur GPL (boîtier + vanne + moteur pas à pas)

13/ Commutateur GPL/essence

14/ Jauge de carburant GPL

Figure 1. 2: Installation GPL c sur une Citroën Xantia break

ANNEXE 2 : Les données

<i>Années</i>	<i>Consommation essence (tonnes)</i>	<i>Consommation gasoil (tonnes)</i>	<i>Consommation GPLc (tonnes)</i>	<i>Masse salariale (Milliers de da)</i>	<i>Parc automobile total (véhicules)</i>	<i>Prix de l'essence* à la pompe (da/l)</i>	<i>Prix du gasoil à la pompe (da/l)</i>
1980	1212970.	1974127.	0.000000	57.10000	1196381.	0.100000	0.200000
1981	1326221.	2161911.	0.000000	65.80000	1366056.	0.500000	0.200000
1982	1393504.	2284736.	0.000000	76.90000	1511506.	1.390000	0.200000
1983	1528789.	2635585.	0.000000	88.60000	1667853.	1.390000	0.200000
1984	1715085.	2906280.	0.000000	94.20000	1863257.	1.390000	0.200000
1985	1784843.	3127540.	0.000000	103.00000	1999576.	2.300000	0.200000
1986	1897157.	3236904.	0.000000	120.10000	2110601.	2.300000	0.450000
1987	1907341.	3337156.	0.000000	125.70000	2228544.	2.300000	0.450000
1988	1962230.	3246494.	0.000000	137.50000	2311667.	2.300000	0.450000
1989	2093969.	3435207.	0.000000	152.20000	2404333.	2.300000	0.450000
1990	2169015.	3390503.	5021.000	180.00000	2491771.	3.000000	1.100000
1991	2267384.	3673728.	9487.000	255.50000	2554662.	5.000000	1.300000
1992	2217840.	3577702.	10000.00	341.30000	2601439.	5.000000	1.300000
1993	2333354.	3446211.	10000.00	412.50000	2650662.	6.000000	2.700000
1994	2209085.	3225275.	10000.00	469.90000	2692427.	7.000000	6.450000
1995	2049359.	3131433.	43567.00	568.80000	2724160.	12.000000	7.500000
1996	2022983.	3097795.	70333.00	667.20000	2742306.	13.000000	9.700000
1997	1958481.	3023384.	103927.0	772.10000	2786257.	18.000000	11.500000
1998	1944286.	3351496.	133049.0	794.60000	2818681.	18.400000	11.500000
1999	1870744.	3434132.	153579.0	847.60000	2886064.	19.150000	11.750000
2000	1903945.	3697428.	182194.0	884.60000	2914272.	19.150000	11.750000
2001	1864496.	3929919.	217267.0	972.40000	2938252.	19.150000	11.750000
2002	1847649.	4560486.	253770.0	1047.30000	2977857.	20.150000	11.750000
2003	1850061.	4833707.	281989.0	1116.50000	3027445.	20.150000	11.750000
2004	1872343.	5124643.	307636.0	1244.60000	3107526.	21.200000	11.750000
2005	1881500.	5370000.	308900.0	1349.60000	3211052.	21.200000	11.750000
2006	1860020.	5790000.	316950.0	1429.60000	3252003.	21.200000	13.700000

Annexe 2 : Les données des séries

Années	Prix du GPLc à la pompe (da/l)	PIB en millions de da	Population (individus)	Population active (individus)	Population transportée par le transport en commun (individus)	Taux de croissance annuel du parc automobile
1980	0.00	8821.570	18666000	4853160.	1.45E+08	0.196381
1981	0.00	100943.4	19262000	5046644.	1.62E+08	0.141823
1982	0.00	112778.7	19883000	5268995.	1.62E+08	0.106474
1983	0.00	126837.6	20522000	5779374.	1.61E+08	0.103437
1984	0.00	142001.2	21185000	5698765.	1.51E+08	0.117159
1985	0.00	154407.8	21863000	5946736.	1.53E+08	0.073161
1986	0.00	159495.0	22512000	6168288.	1.45E+08	0.055524
1987	0.00	167138.6	23139000	6386364.	1.68E+08	0.055881
1988	0.00	186309.8	23783000	6611674.	1.82E+08	0.037299
1989	0.00	219380.6	24409000	6810111.	1.70E+08	0.040086
1990	1.10	292670.0	25022000	6810111.	1.60E+08	0.036366
1991	1.30	451362.1	25643000	7359541.	1.18E+08	0.025223
1992	1.30	551735.7	26271000	7697403.	1.11E+08	0.018310
1993	2.70	613315.5	26894000	8014412.	74000000	0.018992
1994	4.00	757453.9	27496000	8386280.	61000000	0.015756
1995	4.80	1010529.	28060000	8726660.	48000000	0.011786
1996	5.00	1276568.	28566000	9055422.	37000000	0.006661
1997	7.00	1416138.	29045000	9410580.	42000000	0.016027
1998	7.20	1422509.	29507000	9766817.	44000000	0.011636
1999	7.20	1645600.	29965000	10128170	51000000	0.023905
2000	7.20	2128601.	30416000	10493520	54000000	0.009773
2001	7.20	2187406.	30879000	10869408	58000000	0.008228
2002	7.20	2248986.	31357000	11229408	60000000	0.013479
2003	7.20	2318986.	31848000	11585645	66000000	0.016652
2004	7.20	2377791.	32364000	11950645	73000000	0.026451
2005	7.20	2643830.	32600000	12300651	78000000	0.033314
2006	9.00	2676901.	33200000	12629413	82000000	0.012753

Tableau 2. 1 : Données historiques des données 1980-2006

Annexe 3: Guide Eviews 5

[EVI][LAD 06]



EViews est un logiciel d'économétrie complet, moderne et convivial offrant toutes les fonctions nécessaires en analyse financière, prévision macro-économique et simulation.

Ce logiciel présente plusieurs avantages à savoir :

- Une interface utilisateur souple et rapide.
 - Des outils d'économétrie.
 - Techniques d'estimation d'équations pour séries chronologiques et tableaux croisés.
 - Evaluation de modèles : tests d'hypothèses, modèle ARCH.
 - Prévision et simulation : simulation stochastique, méthode de Monte-Carlo, prévisions statistiques et dynamiques.
 - Gestion des données : base de données intégrée, conversion automatique des fréquences, import de fichiers ASCII et Excel, 4 millions d'observations par série.
 - Etc...
- **Environnement Général de la Workfile d'Eviews :**

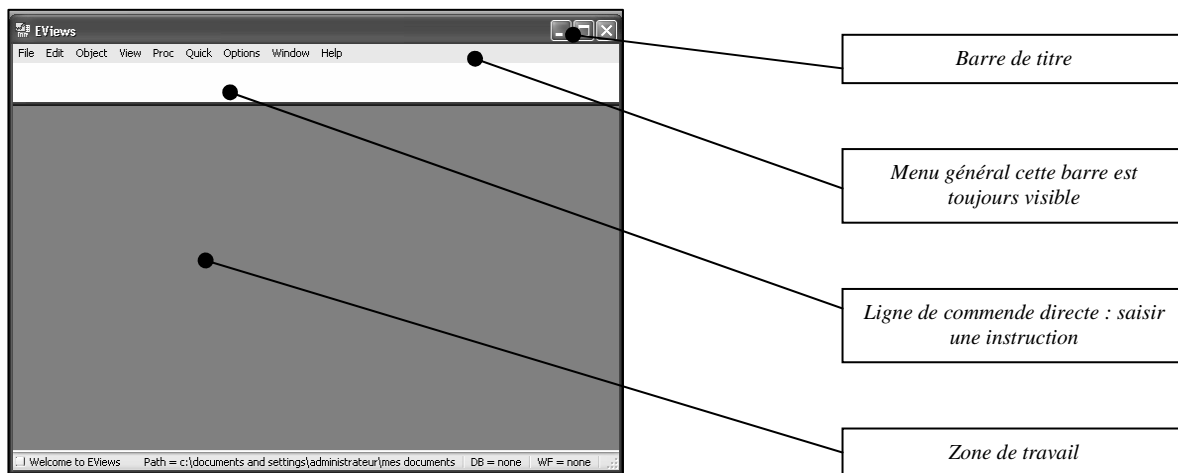


Figure 3. 1: L'environnement de l'espace de travail (Workfile) de Eviews 5

- **Barre de titre** : Elle est en haut de la fenêtre principale, avec le label *Eviews*, quand cette barre est mise en bleu foncé, *Eviews* est le programme en cours d'exécution.
- **Menu général** : Il est situé en dessous de la barre de titre. Il y a les menus descendants pour trouver les fonctions disponibles d'*Eviews* pour le workfile courant.
- **Ligne de commande** : Elle est en dessous du menu principal, elle est affichée comme panneau blanc, ou vous pouvez écrire les commandes d'*Eviews*.
- **Zone de travail** : Elle est en bas de la fenêtre où *Eviews* montrera les diverses fenêtres d'objet qu'il crée.

○ **Création d'un espace de travail :**

La première étape consiste à créer un espace de travail « *WORKFILE* » par : <File> <New> <Workfile... > (figure 3.2), c'est à dire créer un espace d'accueil pour les séries statistiques et les résultats de calcul.



Figure 3. 2: Création d'un workfile

Dans la fenêtre proposée (figure3.2) : il convient d'indiquer le format des séries qui seront utilisées (mensuelles, semestrielles, annuelles, sans dates ...), ainsi le type de structure des séries [dans notre cas : « dated- regular frequency » : séries datées (1980 – 2006) avec une fréquence régulière] et nommer l'espace de travail [Régression].

Par défaut, deux séries seront alors créées, la série c qui correspond par défaut a la constante, et la série *resid* qui correspondra lors de la régression à la série des résidus (figure3.3).

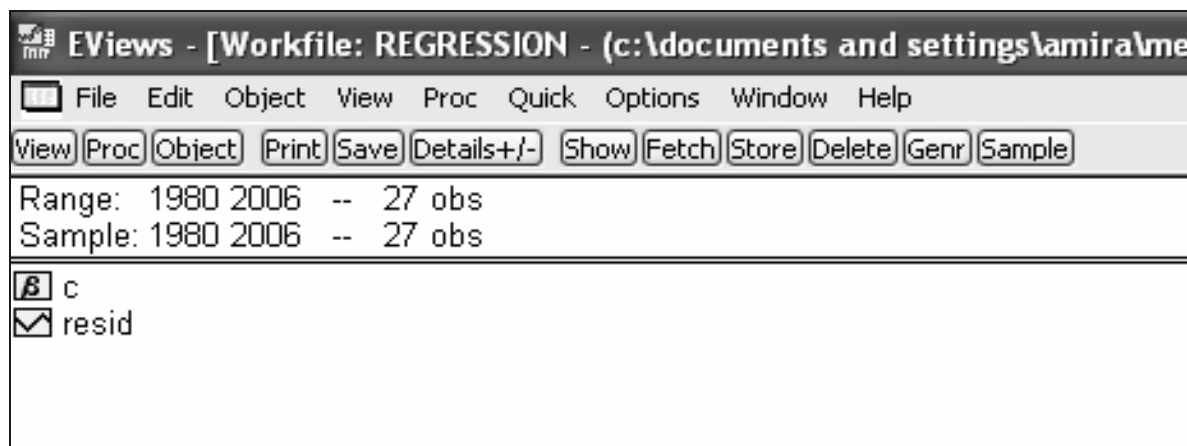


Figure 3. 3: création des séries constante et résiduelle par défaut

- **Création d'une série :**

Pour créer une série, faire : <Object> <New object> : après l'apparition de la fenêtre, un clique sur <series>, ensuite il s'agira de la nommer et de cliquer sur ok (Figure 3.4).

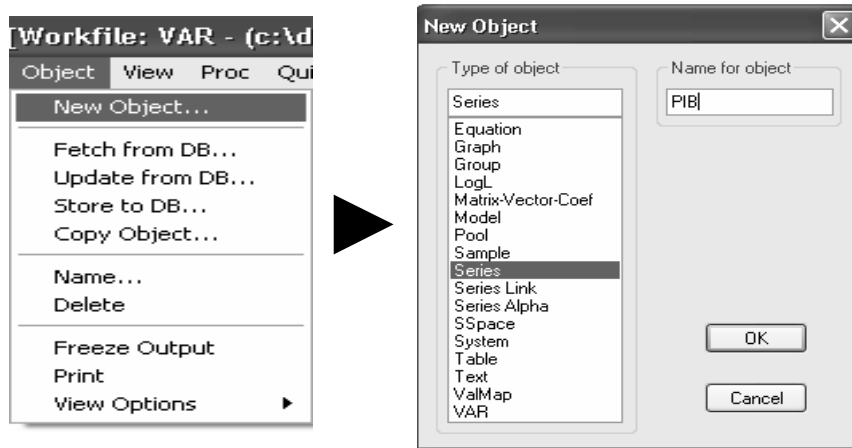


Figure 3. 4: Création d'une série

- **La saisie des données d'une série :**

Pour faire la saisie des données d'une série, il suffit de cliquer sur <Edit +/-> et faire la saisie (figure 3.5).

On peut même importer des données d'EXCEL en procédant par : <File> <Import>

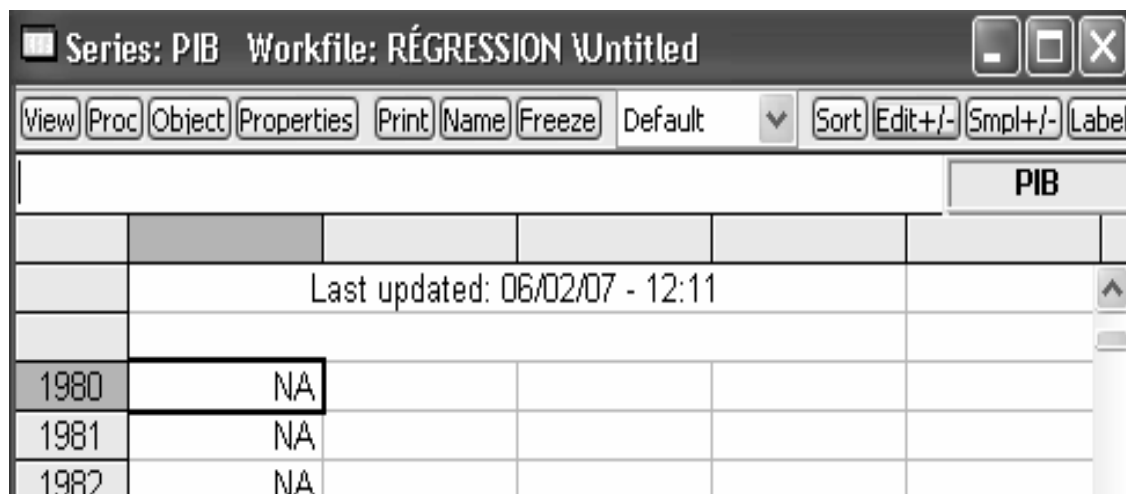


Figure 3. 5: La saisie des données d'une série

- **Faire un graphique de la série :**

Dans la barre d'outil de la base de données cliquer sur <View> puis <Graph>. (Figure 3.6).

Le graphique apparaît dans une fenêtre réduite. Pour sauvegarder le graphique, il suffit de cliquer sur *Name* dans la boîte de dialogue qui apparaît. Le graphique figurera alors dans votre *workfile* comme un objet graphique aux côtés des objets séries des données déjà présents.



Figure 3. 6: Graphe d'une série

- **Faire le corrélogramme de la série :**

Après avoir sélectionné la série en cliquant dessus, sur la barre d'outil associé faire : <View> <Correlogramm> (figure3.7). Une fenêtre s'ouvre pour préciser quel type de corrélogramme voulu : *level* : série bute, *1st difference* : série différenciée une fois, *2nd difference* : série différenciée deux fois.

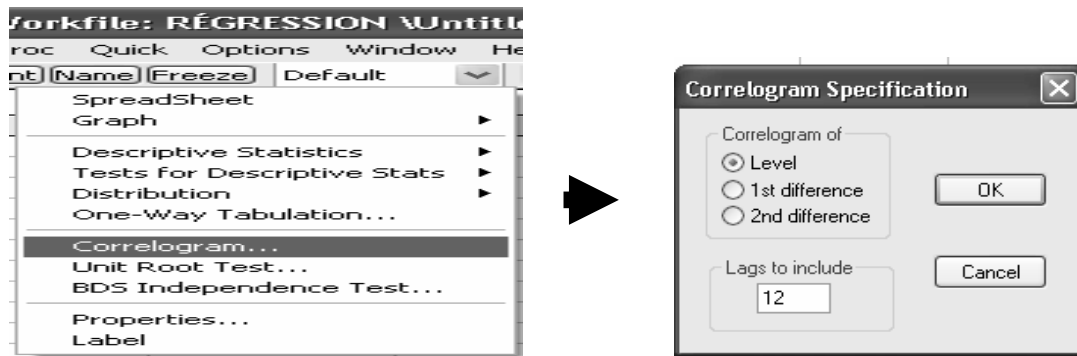


Figure 3. 7: Corrélogramme d'une série

- **Création d'une nouvelle série à partir d'une série existante :**

Pour générer une série à partir d'une autre: dans le menu général, faire : <Quick> <Generate series> (figure 3.8).

Par exemple pour créer la série DDBIP (la différenciée deux fois de la série existante PIB), il faut saisir la formule suivante :

$$DDBIP = d(PIB, 2) \quad \text{ou bien} \quad DDBIB = DPIB - DPIB(-1)$$

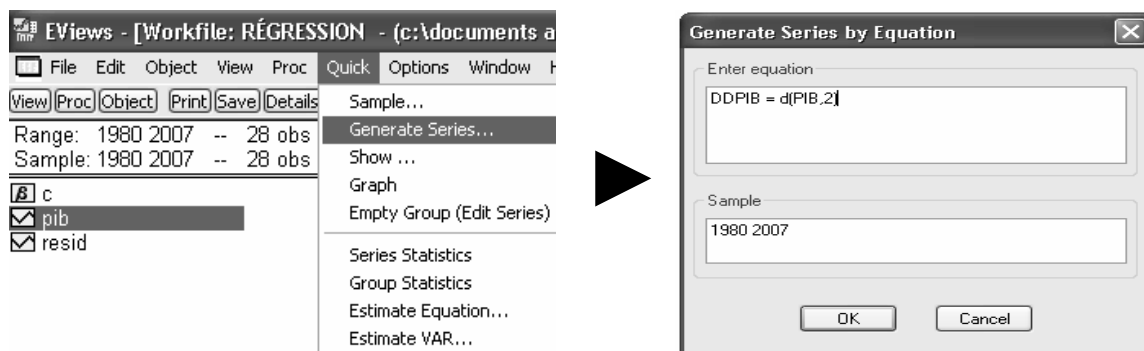


Figure 3. 8 : Génération d'une nouvelle série

- **Estimation d'un modèle :**

Pour estimer un modèle faire : <Quick> <Estimate equation>.

Par exemple, pour un modèle ARIMA faire entrer dans cette ordre : le nom de la série à modéliser, et les paramètres du modèle en les séparant par un espace. (Figure 3.9).

Sinon pour un modèle de régression multiple faire entrer : la variable endogène, la constante c, suivie des variables exogènes du modèle en les séparant par un espace.

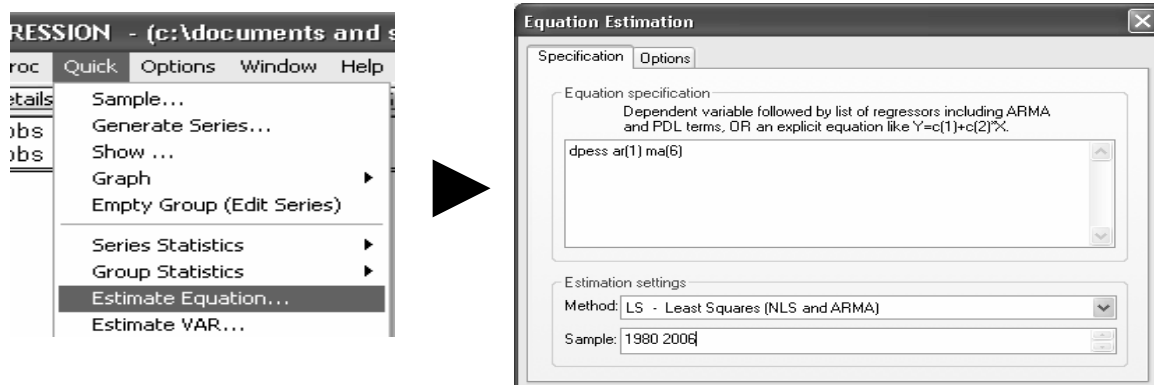


Figure 3. 9: Estimation d'un modèle

- **Les Tests statistiques :**

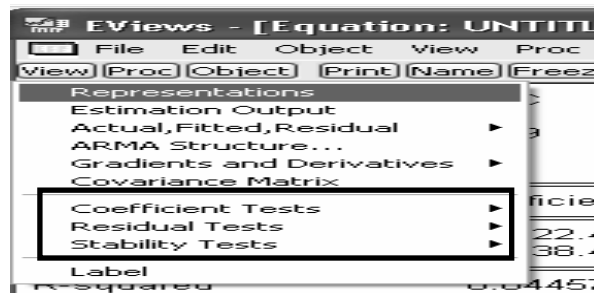


Figure 3. 10 : Tests statistiques

Le T- test, F- test et le test de Durbin-Watson figurent directement dans la fenêtre *equation*. Pour calculer d'autres tests statistiques, cliquez sur le bouton *view* de la fenêtre *equation*, parmi la liste proposée à chaque fois, sélectionnez soit les 'Coefficient Tests', soit les 'Residuals Tests', soit les 'Stability Tests' (Figure 3.10)

Pour le test de Dickey Fuller d'une série : de la barre d'outil de la série faire : <view> <Unit Root Test>.

- **La modélisation VAR :**

- * Estimation du VAR et choix du nombre de retards : on utilise les critères d'information AIC et de Schwarz SC donnés dans l'estimation du VAR : <Quick><Estimate VAR>.(figure 3.11)

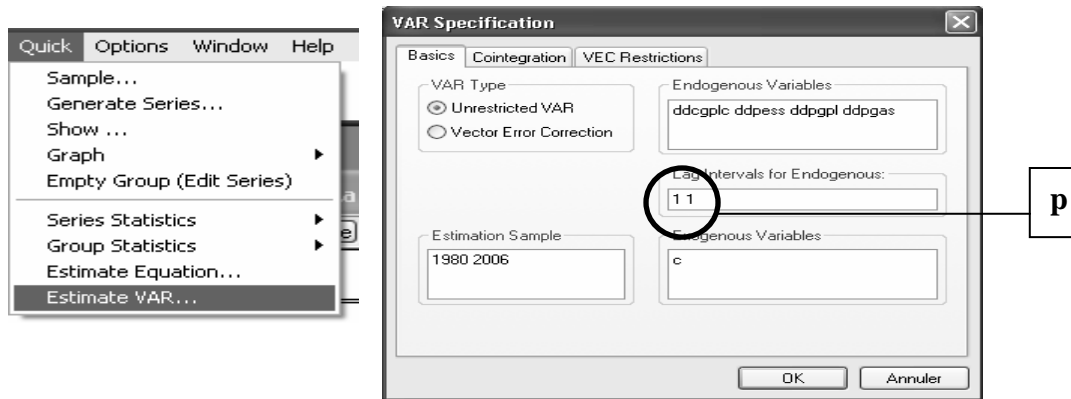


Figure 3. 11: Estimation du modèle VAR

- * Pour le test de causalité de Granger faire : <Quick><Group statistics> <Granger Causality Test> (figure 3.12)

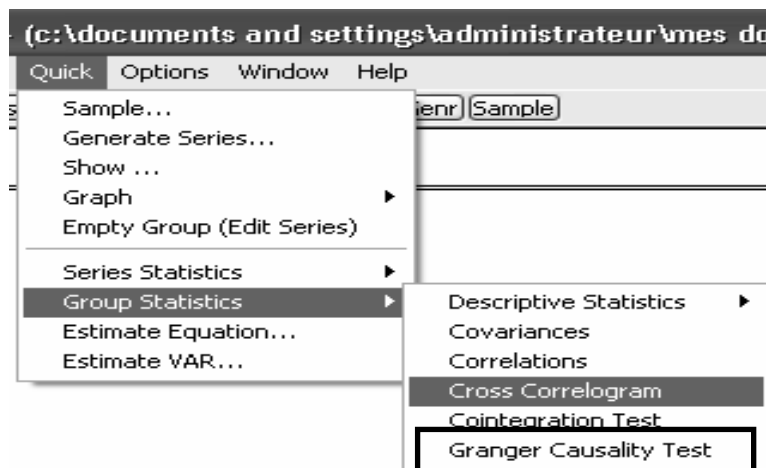


Figure 3. 12: Test de Granger

- * Pour la décomposition de la variance et l'analyse impulsionnelle faire : <View> <Variance Decomposition> <Impulse response> (respectivement)

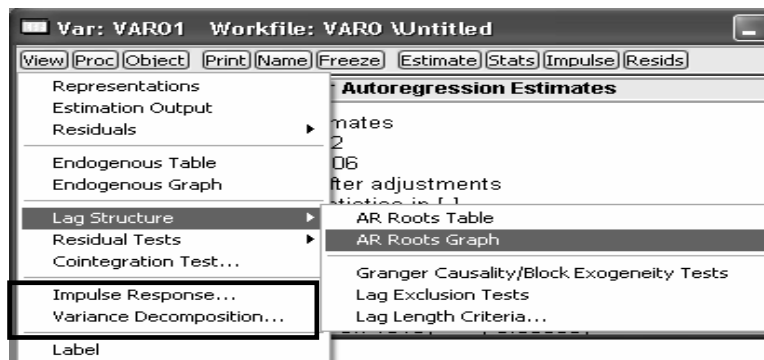


Figure 3. 13: Validation des paramètres et l'analyse dynamique d'un VAR

Annexe 4 : Test de Dickey- Fuller

I. Le test de Dickey- Fuller simple : [BOU 02]

Les modèles servant de base à la construction de ce test sont au nombre de trois.

[1] : Modèle sans constante ni tendance déterministe : $X_t = \rho X_{t-1} + \varepsilon_t$

[2] : Modèle avec constante et sans tendance déterministe : $X_t = c + \rho X_{t-1} + \varepsilon_t$

[3] : Modèle avec constante et avec tendance déterministe : $X_t = c + bt + \rho X_{t-1} + \varepsilon_t$

Où ε_t est supposé être un bruit blanc.

L'hypothèse du test est formulée comme suit :

$H_0 : \rho = 1$

$H_1 : |\rho| < 1$

En s'inspirant du modèle 1 : $X_t = \rho X_{t-1} + \varepsilon_t \dots \dots \dots (1)$

Et en retranchant X_{t-1} à chaque membre de l'équation (1), on obtient :

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

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

En pratique, en supposant : $\Phi = (\rho - 1)$ on estime les modèles suivants :

Modèle 4 : $\Delta X_t = \Phi X_{t-1} + \sum_{j=1}^p \Phi_j \Delta X_{t-j} + \varepsilon_t$

Modèle 5 : $\Delta X_t = \Phi X_{t-1} + \sum_{j=1}^p \Phi_j \Delta X_{t-j} + \varepsilon_t + c$

Modèle 6 : $\Delta X_t = \Phi X_{t-1} + \sum_{j=1}^p \Phi_j \Delta X_{t-j} + \varepsilon_t + c + bt$

Ce qui revient à dire que le test de racine unitaire repose sur le test de l'hypothèse nulle : $\Phi = 0$ (non stationnaire) contre l'hypothèse alternative : $|\Phi| \neq 0$ (stationnaire)

Et donc le système d'hypothèse devient :

$H_0 : \Phi = 0$

$H_0 : |\Phi| \neq 0$

▪ **Principes du test de Dickey-Fuller :**

Sous l'hypothèse H_0 , le processus X_t n'est pas stationnaire quelque soit le modèle retenu. Dickey et Fuller ont étudié la distribution asymptotique de l'estimateur du paramètre ρ sous l'hypothèse H_0 à l'aide des simulations de Monte-Carlo, ils ont tabulé les valeurs critiques pour des échantillons de tailles différentes.

Soit la t -statistique notée $(t_{\hat{\phi}})$ tel que : $t_{\hat{\phi}} = \frac{\hat{\Phi} - 1}{\sigma_{\hat{\phi}}^2}$

On la compare alors avec la valeur critique t tabulée :

si $t_{\hat{\phi}} \geq t_{\text{tabulé}}$ on accepte l'hypothèse H_0 , il existe une racine unitaire, d'où le processus est non stationnaire.

Sinon on rejette l'hypothèse.

Remarque :

- Les principaux logiciels d'analyse de séries temporelles calculent automatiquement les valeurs critiques ($t_{\hat{\phi}}$)
- Ces tests révèlent l'existence d'une racine unitaire (une non stationnarité) mais restent insuffisants pour faire la différence entre les processus TS et DS, c'est ainsi qu'on adopte un algorithme en trois étapes.

▪ **Enoncé de l'algorithme du test de Dickey Fuller:**

Avant d'énoncer l'algorithme, il est important de préciser que :

- On dit que la tendance est significativement différente de 0, si et seulement si **et seulement si t- statique > valeur critique** lue dans la table de Dickey Fuller
Alors la tendance existe, sinon elle est dite non significativement différente de zéro
- On dit que la constante est non significativement différente de 0, **si et seulement si t- statique < valeur critique** lue dans la table de Dickey- Fuller sinon elle est dite significativement différente de zéro

Etape 1 :

Dans cette étape on estime le modèle [3] et on teste la signification de la tendance déterministe.

- Si la tendance n'est pas significativement différente de zéro (t- statistique de la tendance est inférieure aux valeurs critiques de la tendance tabulées par Dickey- Fuller), on passe à l'étape 2.

- Si la tendance est significativement différente de zéro, on teste l'hypothèse nulle de racine unitaire en comparant : ($t_{\hat{\phi}}$) aux valeurs tabulées par Dickey fuller.

Si on accepte H_0 alors X_t est non stationnaire de type DS.

Si on rejette H_0 alors X_t est non stationnaire de type TS.

Etape 2 :

On estime le modèle [2] et on teste la signification de la constante.

- Si la constante n'est pas significative on passe à l'étape 3.

- Si la constante est significative, on teste l'hypothèse de racine unitaire.

Si H_0 est acceptée, donc X_t est non stationnaire de type DS.

Si H_0 n'est pas acceptée alors X_t est stationnaire.

Etape 3 :

On estime dans ce cas le modèle [1] et on teste l'hypothèse nulle de racine unitaire.

Si H_0 est acceptée alors le processus X_t est non stationnaire de type DS

Si H_0 est rejetée alors le processus X_t est stationnaire.

II. Test de Dickey-Fuller Augmenté (ADF):

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

Le test découle de manière similaire aux tests DF simples, seules les tables statistiques diffèrent.

Remarque

Les principaux logiciels d'analyse de séries temporelles calculent automatiquement les valeurs critiques et la valeur de p.

Annexe 5 : Application de la méthode de Box & Jenkins sur les séries T_{cr} , PIB et P_{ess}

I. Etude de la série : « le taux de croissance du parc auto » T_{cr} :

1. Etude de la série brute :

La série T_{cr} représente l'évolution annuelle du taux de croissance du parc auto sur une période allant de 1980 à 2006.

1.1 Analyse préliminaire de la série :

La première étape consiste à l'analyse du graphe de la série, ceci permettra d'avoir une idée sur la nature et les caractéristiques de la série (tendance, saisonnalité.....etc.).

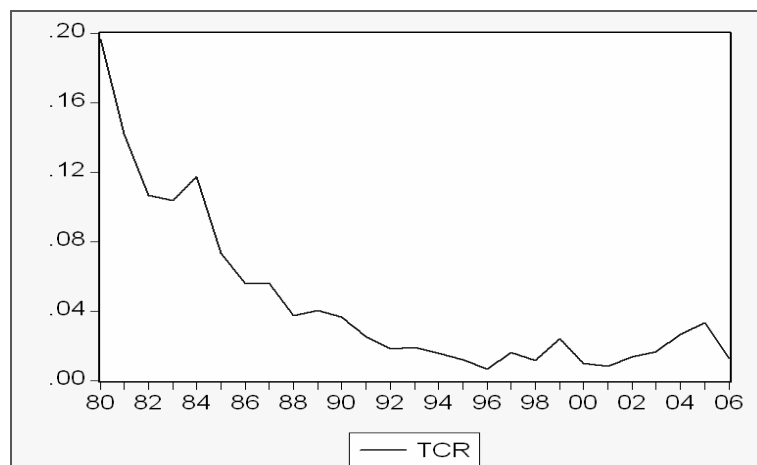


Figure 5. 1 : Représentation graphique de la série T_{cr}

La représentation graphique de la série T_{cr} (figure 5. 1) nous montre que la série semble être affectée d'une tendance, il semble donc que la série T_{cr} soit non stationnaire.

1.2 Examen du corrélogramme de la série 'Tcr' :

Le corrélogramme de la série T_{cr} est représenté par la figure 5.2

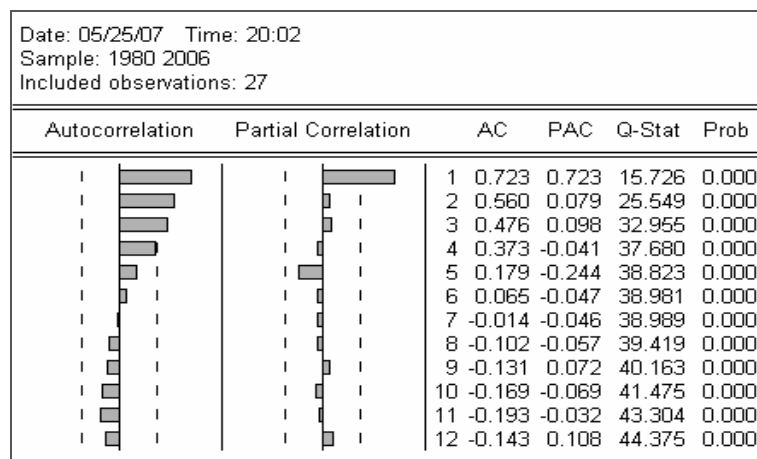


Figure 5. 2 : Corrélogramme de la série T_{cr}

L'analyse du corrélogramme simple et partiel de la série brute nous indique une non stationnarité de la série. En effet, la fonction d'autocorrélation simple et partielle diminue lentement vers zéro, nous remarquons aussi l'existence de quelques pics sortants de la bande de confiance, ceci est indicatif d'une série non stationnaire.

Il devient donc nécessaire de vérifier ceci en appliquant le Test de Dickey- Fuller pour déterminer si le processus est stationnaire ou pas, de quel processus non stationnaire il s'agit et par la suite de le rendre stationnaire.

1.3 Test de Dickey- Fuller sur la série T_{cr} :

L'application du Test de Dickey-Fuller sur la série T_{cr} a donné les résultats suivants *tableau 5.1*

Tableau 5. 1 : Test de Dickey- Fuller sur T_{cr}

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	0.88	2.79	$0.88 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	2.00	2.54	$2.00 \leq 2.54$	Constante non significativement différente de zéro
<i>Modèle 4</i>	-5.40	-1.95	$-5.40 \leq -1.95$	Série stationnaire

Commentaires :

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 0.88) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%, donc la tendance est non significativement différente de zéro.

- **Pour le Modèle 5 :** *modèle avec constante sans tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 2.00) est inférieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%. La constante est donc non significativement différente de zéro.

- **Pour le Modèle 4 :** *modèle sans constante, ni tendance*

On teste l'hypothèse nulle H_0 : « Existence d'une racine unitaire »:

La statistique de Student (t- statistique $t = -5.40$) est inférieure à la valeur critique lue dans la table de Dickey-Fuller (-1.95) au seuil 5%, on rejette alors l'hypothèse de non stationnarité H_0 , la série T_{cr} est donc stationnaire.

La série étudiée étant stationnaire, nous pouvons aborder la phase de modélisation par la méthode de Box & Jenkins.

2. Mise en œuvre de la méthode de Box and Jenkins :

2.1 Spécification du modèle :

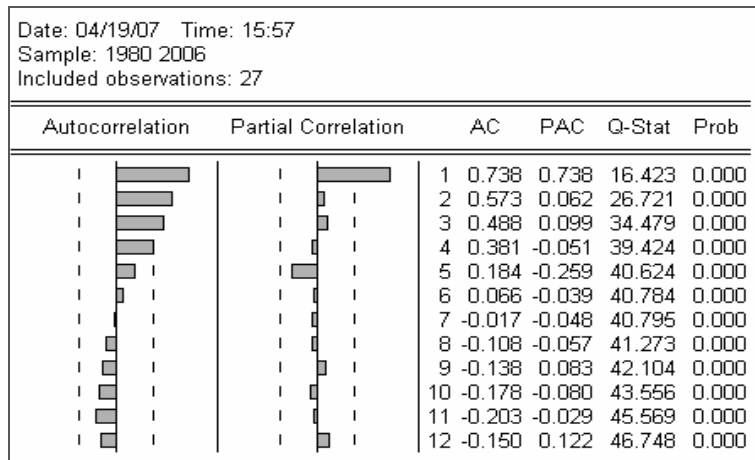


Figure 5.3 : Corrélogramme de la série T_{cr}

Le corrélogramme de la série T_{cr} « le taux de croissance du parc automobile » est donné par la figure 5.3

- Le corrélogramme simple a des pics significatifs aux 4 premiers décalages.
- Le corrélogramme partiel a un seul pic significatif au décalage 1

Nous avons plusieurs modèles candidats qui peuvent représenter la série « T_{cr} », parmi lesquels nous avons retenu les modèles suivants : $ARIMA(1, 0, 4)$, $ARIMA(1, 0, 5)$.

2.2 Estimation des paramètres :

Le logiciel *Eviews* a permis d'effectuer l'estimation des paramètres de chaque modèle. Les résultats pour chaque modèle sont donnés dans les tableaux (5.2) et (5.3)

- **Estimation des paramètres de $ARIMA(1, 0, 4)$:**

Tableau 5.2 : Estimation du modèle $ARIMA(1,0,4)$ de T_{cr}

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.748791	0.075514	9.915973	0.0000
MA(4)	0.859474	0.048230	17.82015	0.0000
R-squared	0.913872	Mean dependent var		0.041050
Adjusted R-squared	0.910283	S.D. dependent var		0.037432
S.E. of regression	0.011212	Akaike info criterion		-6.069897
Sum squared resid	0.003017	Schwarz criterion		-5.973120
Log likelihood	80.90866	Durbin-Watson stat		1.639781
Inverted AR Roots	.75			
Inverted MA Roots	.68-.68i	.68-.68i	-.68+.68i	-.68+.68i

$R^2 = 0.910283$

• **Estimation des paramètres de ARIMA (1, 0, 5) :**

Tableau 5. 3 : Estimation du modèle ARIMA (1,0,5) de T_{cr}

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.840694	0.065820	12.77263	0.0000
MA(5)	0.901273	0.030874	29.19210	0.0000
R-squared	0.945925	Mean dependent var		0.041050
Adjusted R-squared	0.943672	S.D. dependent var		0.037432
S.E. of regression	0.008884	Akaike info criterion		-6.535369
Sum squared resid	0.001894	Schwarz criterion		-6.438592
Log likelihood	86.95980	Durbin-Watson stat		2.403299
Inverted AR Roots	.84			
Inverted MA Roots	.79-.58i -.98	.79+.58i	-.30+.93i	-.30-.93i

$R^2 = 0.943672$

2.3 Adéquation des modèles :

• **Test concernant les paramètres du modèle ARIMA (1, 0, 4) :**

- Les rapports de Student des coefficients du modèle sont en valeur absolue (9.915973 et 17.82015) supérieurs à 1.96, ils sont donc significativement différents de zéro. Ceci est confirmé par les probabilités de nullité de ces coefficients (0.0000 et 0.0000) qui sont inférieures à 0.05.
- Les racines du polynômes autorégressif et moyenne mobile sont en module supérieures à 1 (leurs inverses fournis par *Eviews* sont inférieurs à 1) donc les conditions de causalité et d'inversibilité sont vérifiées.

Le modèle ARIMA (1, 0, 4) est donc accepté.

• **Test concernant les paramètres du modèle ARIMA (1, 0, 5) :**

- Les rapports de Student des coefficients du modèle sont en valeur absolue (12.77263 et 29.19210) supérieurs à 1.96, ils sont donc significativement différents de zéro, chose confirmée par les probabilités de nullité de ces coefficients (0.0000 et 0.0000) qui sont inférieures à 0.05.
- Les racines du polynômes autorégressif et moyenne mobile sont en module supérieures à 1 (leurs inverses fournis par *Eviews* sont inférieurs à 1) donc les conditions de causalité et d'inversibilité sont vérifiées.

Le modèle ARIMA (1, 0, 5) est donc accepté.

2.4 Choix du modèle :

Tableau 5. 4 : Choix du meilleur modèle pour T_{cr}

Le modèle	R^2	\bar{R}^2	BIC	AIC
ARIMA (1, 0, 4)	0.913872	0.910283	-5.973120	-6.069897
ARIMA (1, 0, 5)	0.945925	0.943672	-6.438572	-6.535369

Suivant le critère du pouvoir prédictif et les critères d'information mentionnés dans le *tableau V.21*, le modèle **ARMA (1, 0, 5)** sera choisi pour représenter le processus.

En effet :

- Les coefficients R^2 et \bar{R}^2 du modèle *ARIMA (1, 0, 5)* sont supérieurs à ceux de *ARIMA (1, 0, 4)*.
- Les critères AIC et BIC du modèle *ARIMA (1, 0, 5)* leur sont inférieurs.

2.5 Test sur les résidus :



Figure 5. 4 : Représentation graphique de la série résiduelle, actuelle et estimée de T_{cr}

Nous constatons d'après la représentation graphique de la série résiduelle, actuelle et estimée *figure 5.4* que le graphe de la série estimée est semblable à celui de la série réelle à quelques pics près, donc le modèle explique à 94,59 % la variation totale de la variable dépendante ($R^2 = 0.945903$).

- **Corrélogramme des résidus du modèle *ARIMA (1, 0, 5)* :**

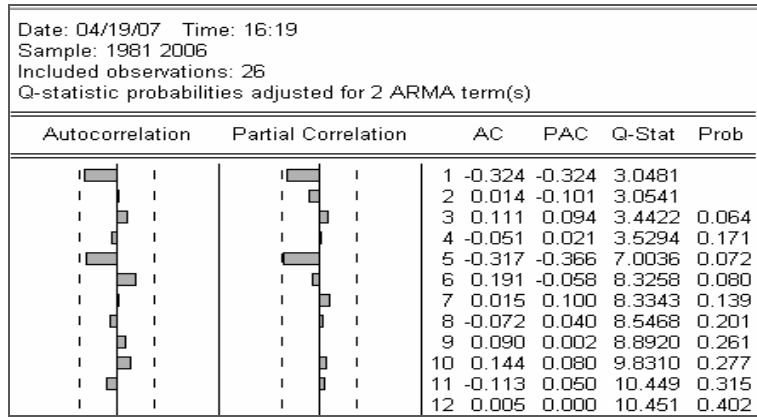


Figure 5.5 : Correlogramme des résidus du modèle ARIMA (1, 0, 5)

L'analyse du correlogramme des résidus, illustré par la *figure 5.5* montre que tous les termes sont à l'intérieur de l'intervalle de confiance, donc les résidus forment un bruit blanc, ceci est confirmé par la statistique de Ljung-Box (Q- stat) qui est inférieure à la valeur théorique de $\chi^2(h)$ quelque soit le retard h , et spécialement pour $h=12$ on a :

Q- stat (12) = 10,451 inférieure à $\chi^2(12) = 21.026$ Au seuil 0.05, donc les résidus forment un bruit blanc.

- **Test de normalité sur les résidus du modèle ARIMA (1, 0, 5) :**

L'histogramme de la distribution et les valeurs empiriques de Skewness et de Kurtosis et la statistique de Jarque- Bera sont données par la *figure 5.6* ci-dessous :

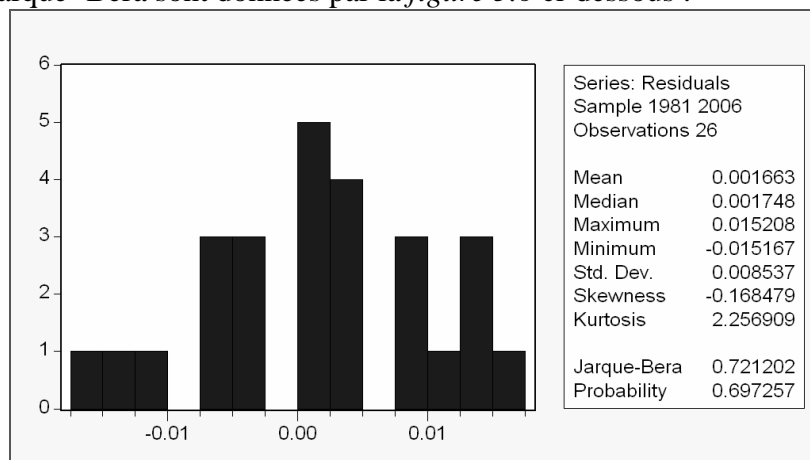


Figure 5.6 : Histogramme de la distribution des résidus du modèle ARIMA(1,0,5)

Le coefficient de Skewness (asymétrie) = -0,168479 .

Le coefficient de Kurtosis (aplatissement) = 2.256909

On construit alors les statistiques centrées réduites correspondantes que l'on compare à 1.96 qui représente la valeur de la loi normale au seuil 5%, on aura donc :

$$v_1 = (\sqrt{0,168479}) / (6 / 27) = 0,87073121 \text{ qui est inférieure à } 1,96.$$

$$v_2 = (3 - 2,256909) / (\sqrt{24 / 27}) = 0,788167 \text{ qui est inférieure à } 1,96.$$

L'hypothèse de normalité est donc acceptée, ce qui est confirmé par la statistique de Jarque-Berra : JB = 0.721202 qui est inférieure à $\chi^2(2) = 5.991$

Donc nous constatons que les résidus forment un bruit blanc gaussien.
Le modèle $ARIMA(1, 0, 5)$ est donc validé.

II. Etude de la série PIB :

1. Etude de la série brute PIB:

La série PIB représente l'évolution annuelle du PIB pour sur une période allant de 1980 à 2006.

1.1 Analyse préliminaire de la série PIB :

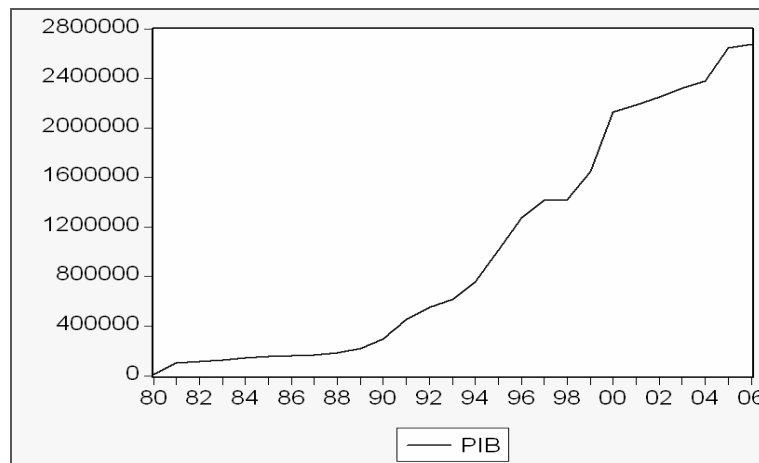


Figure 5. 7 : Représentation graphique de la série PIB

La représentation graphique de la série « PIB » *figure 5.7* nous montre que la série semble être affectée d'une tendance, il semble donc que la série PIB soit non stationnaire.

1.2 Examen du corrélogramme de la série PIB :

Le corrélogramme de la série PIB est représenté par la *figure 5.8*

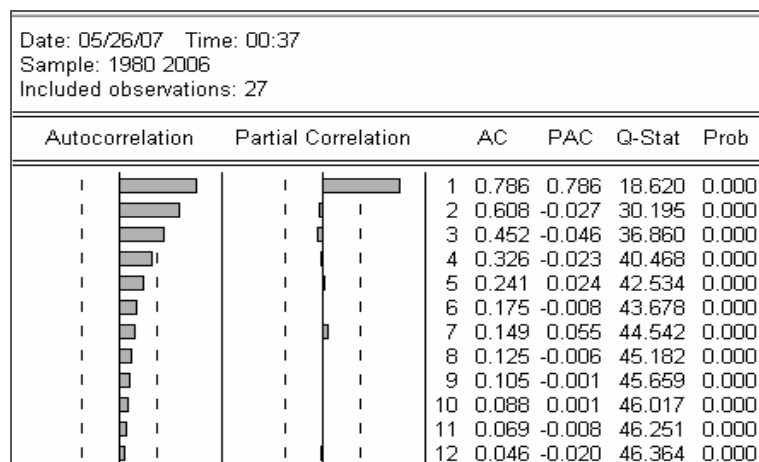


Figure 5. 8 : Corrélogramme de la série PIB

L'analyse du corrélogramme simple et partiel de la série brute nous indique une non stationnarité de cette dernière. En effet: la fonction d'autocorrélation simple et partielle diminue lentement vers zéro, de plus on remarque de nombreux pics sortant de l'intervalle de confiance, ceci est caractéristique d'une série non stationnaire.

Il est donc nécessaire de le vérifier en appliquant le test de Dickey-Fuller pour déterminer la stationnarisation du processus, il s'agit et par la suite de le rendre stationnaire

1.3 Test de Dickey- Fuller sur la série PIB :

Tableau 5. 5 : Test de Dickey Fuller sur PIB

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
Modèle 6	2.19	2.79	$2.19 \leq 2.79$	Tendance non significativement différente de zéro
Modèle 5	5.27	2.54	$5.27 \geq 2.54$	Constante significativement différente de zéro
	-2.63	-3.06	$-2.63 \geq -3.06$	Série non stationnaire de type DS

Commentaires :

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 2.19) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.

La tendance est donc non significativement différente de zéro, passons alors à l'estimation du modèle 5 pour tester la significativité de la constante.

- **Pour le Modèle 5 :** *modèle avec constante sans tendance déterministe :*

Nous remarquons que :

La statistique de Student (t- statistic = 5.27) est supérieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.

La constante est donc significativement différente de zéro. Passons alors au test d'existence de racine unitaire :

La statistique de Student (t – statistique = -2.63) est supérieure à la valeur critique (-3.06) lue dans la table de Dickey – Fuller au seuil 5%, on accepte alors l'hypothèse de non stationnarité H_0 . La série PIB est non stationnaire de type DS, la meilleure façon de la rendre stationnaire est de la différencier. La nouvelle série sera notée DPIB

2 Etude de la série différenciée « DPIB » :

2.1 Effet de la différenciation :

La représentation graphique de la série différenciée DPIB est donnée par la figure 5.9

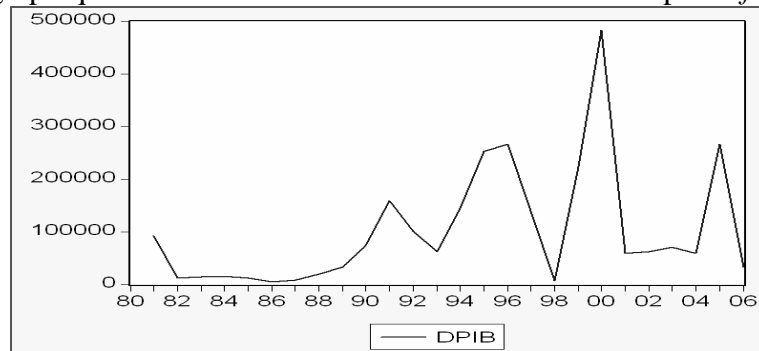


Figure 5. 9: Représentation graphique de la série DPIB

Nous remarquons que l'effet de la tendance a été plus au moins absorbé. Ceci étant, il est nécessaire d'appliquer le Test de Dickey- Fuller pour vérifier si la nouvelle série DPIB est stationnaire ou pas.

2.2 Test de Dickey- Fuller sur la série DPIB :

Nous allons appliquer le Test de Dickey – Fuller sur la série DPIB, le résultat est donné par le tableau 5.6 :

Tableau 5. 6: Test de Dickey- Fuller sur DPIB

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	1.45	2.79	$1.45 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	1.88	2.54	$1.88 \leq 2.54$	Constante non significativement différente de zéro
<i>Modèle 4</i>	-1.51	-1.96	$-1.52 \geq -1.96$	Série non stationnaire de type DS

- **Pour le Modèle 6 :** modèle avec constante et tendance déterministe

Nous remarquons que :

La statistique de Student (t- statistic = 1.45) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.

La tendance est donc non significativement différente de zéro, passons alors à l'estimation du modèle 5 pour tester la significativité de la constante

- **Pour le Modèle 5 :** modèle avec constante sans tendance

Nous remarquons que :

La statistique de Student (t- statistic = 1.88) est inférieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.

La constante est donc non significativement différente de zéro, passons à l'estimation du modèle 4 pour tester l'existence d'une racine unitaire

- **Pour le Modèle 4 :** *modèle sans constante, ni tendance*

La statistique de Student (t- statistique = -1.51) est supérieure à la valeur critique (-1.96) lue dans la table de Dickey – Fuller au seuil 5%, on accepte alors l'hypothèse de non stationnarité H_0 , la série DPIB est donc non stationnaire de type DS.

La meilleure façon de la rendre stationnaire est de la différencier, la nouvelle série sera notée DDPIB.

3 Etude de la série différenciée DDPIB :

3.1 Effet de la différenciation :

La représentation graphique de la série différenciée DDPIB est donnée par la *figure 5.10*

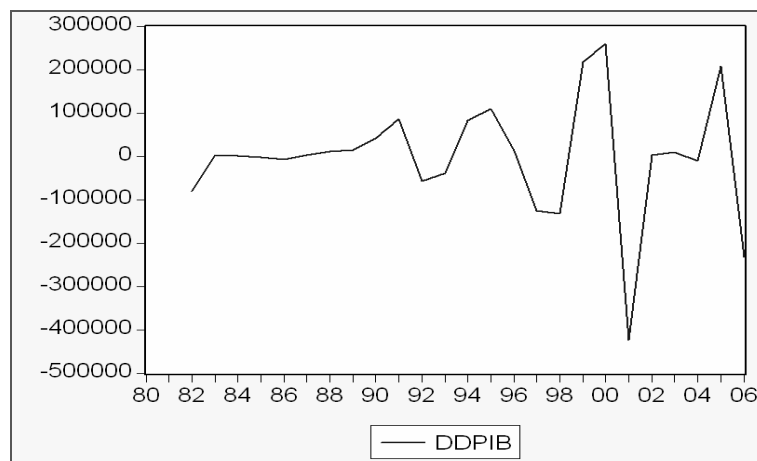


Figure 5. 10 : Représentation graphique de la série DDPIB

Nous remarquons que l'effet de la tendance a été absorbé. Il convient d'appliquer le Test de Dickey- Fuller pour vérifier la stationnarité de la nouvelle série DDPIB.

3.2 Test de Dickey- Fuller sur la série ' DDPIB' :

Nous allons appliquer le Test de Dickey – Fuller sur la série 'DDPIB', le résultat est donné par le *tableau 5.7*

Tableau 5. 7: Test de Dickey- Fuller sur DDPIB

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
Modèle 6	1.48	2.79	$1.48 \leq 2.79$	Tendance non significativement différente de zéro
Modèle 5	0.51	2.54	$0.55 \leq 2.54$	Constante non significativement différente de zéro
Modèle 4	-5.76	-1.95	$-5.76 \leq -1.96$	Série non stationnaire de type DS

Commentaires :

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 1.48) est inférieure à la valeur critique (= 2.79) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la tendance est non significativement différente de zéro

On passe à l'estimation du modèle 5 pour tester la significativité de la constante

- **Pour le Modèle 5 :** *modèle avec constante sans tendance*

Nous remarquons que :

La statistique de Student (t- statistic = 0.51) est inférieure à la valeur critique (2.54) lue dans la table de Dickey Fuller au seuil 5%.

Donc la constante est non significativement différente de zéro.

Passons à l'estimation du modèle 4 pour tester l'existence d'une racine unitaire

- **Pour le Modèle 4 :** *modèle sans constante, ni tendance*

La statistique de Student (t- statistique = -5.76) est inférieure à la valeur critique (-1.95) lue dans la table de Dickey – Fuller au seuil 5%, on rejette alors l'hypothèse de non stationnarité H_0 , la série « DDPIB » est donc stationnaire.

La série étant stationnarisée , nous pouvons passer maintenant à la modélisation du processus par la méthode de Box and Jenkins.

4 Mise en oeuvre de la méthode de Box and Jenkins sur DDBIP :

4.1 Spécification du modèle :

Nous allons examiner le corrélogramme de la série DDPIB *figure 5.11* :

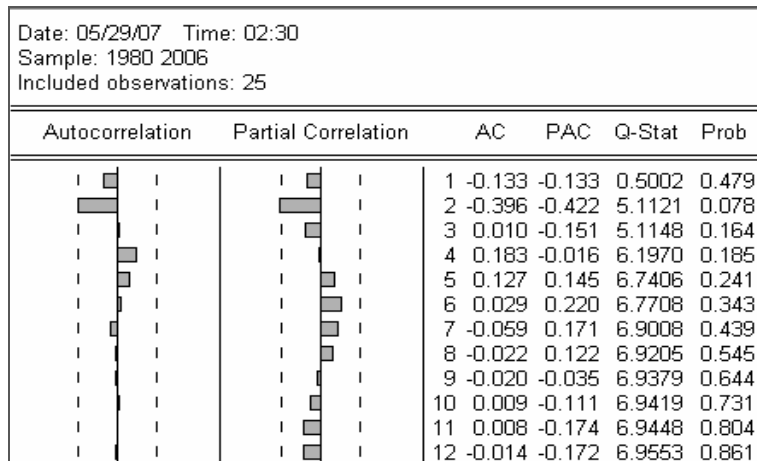


Figure 5.11: Corrélogramme de la série DPIB

Nous remarquons que :

- Le corrélogramme simple a un pic plus ou moins significatif au décalage : 2
- Le corrélogramme partiel a un pic plus ou moins significatif au décalage : 2

Nous avons plusieurs modèles candidats pour représenter la série « DDPIB », parmi lesquels nous avons retenu les modèles suivants :

$ARIMA(2, 2, 2)$, $ARIMA(2, 2, 0)$, $ARIMA(2, 2, 4)$

4.2 Estimation des paramètres des modèles :

L'estimation des paramètres est effectuée en utilisant le logiciel *Eviews*. Les résultats pour chaque modèle sont donnés dans les tableaux (5.8), (5.9) et (5.10):

- Estimation des paramètres de $ARIMA(2, 2, 2)$:

Tableau 5.8 : Estimation du modèle $ARIMA(2, 2, 2)$ du PIB

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(2)	0.048144	0.231212	0.208222	0.8371
MA(2)	-0.908777	0.127568	-7.123881	0.0000
R-squared	0.386979	Mean dependent var		826.6004
Adjusted R-squared	0.357788	S.D. dependent var		144391.0
S.E. of regression	115712.2	Akaike info criterion		26.23854
Sum squared resid	2.81E+11	Schwarz criterion		26.33728
Log likelihood	-299.7432	Durbin-Watson stat		2.698857
Inverted AR Roots	.22	-.22		
Inverted MA Roots	.95	-.95		

Le coefficient de détermination $R^2 = 0.386979$.

• **Estimation des paramètres de ARIMA (2, 2, 0) :**

Tableau 5. 9 : Estimation du modèle ARIMA(2,2,0) du PIB

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(2)	-0.489593	0.213947	-2.288386	0.0321
R-squared	0.192239	Mean dependent var		826.6004
Adjusted R-squared	0.192239	S.D. dependent var		144391.0
S.E. of regression	129772.2	Akaike info criterion		26.42745
Sum squared resid	3.70E+11	Schwarz criterion		26.47682
Log likelihood	-302.9157	Durbin-Watson stat		2.666232
	.95	.00-.95i		

Le coefficient de détermination $R^2 = 0.192239$.

• **Estimation des paramètres de ARIMA (2, 2, 4) :**

Tableau 5. 10 : Estimation du modèle (2, 2, 4) du PIB

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(2)	-0.799779	0.151335	-5.284839	0.0000
MA(4)	-0.819750	0.108336	-7.566764	0.0000
R-squared	0.404895	Mean dependent var		826.6004
Adjusted R-squared	0.376556	S.D. dependent var		144391.0
S.E. of regression	114008.9	Akaike info criterion		26.20888
Sum squared resid	2.73E+11	Schwarz criterion		26.30762
Log likelihood	-299.4021	Durbin-Watson stat		2.577912
Inverted MA Roots	.95	.00-.95i		

Le coefficient de détermination $R^2 = 0.404895$.

4.3 Adéquation des modèles :

• **Test concernant les paramètres du modèle ARIMA (2, 2, 2) :**

- Le rapport de Student du coefficient du modèle autorégressif (0.208222) est en valeur absolue inférieur à 1.96, il est donc non significativement différents de zéro. Ceci est confirmé par la probabilité de nullité de ce coefficient (0.8371) qui est supérieur à 0.05, on rejette donc ce modèle.

- **Test concernant les paramètres du modèle ARIMA (2, 2, 0) :**
 - Le rapport de Student du coefficient du modèle autorégressif (2.288386) est en valeur absolue supérieur à 1.96, il est donc significativement différent de zéro, ceci est confirmé par la probabilité de nullité de ce coefficient (0.0321) qui est inférieure à 0.05.
 - Les racines du modèle autorégressif sont en module supérieures à 1 (leurs inverses fournis par *Eviews* sont inférieurs à 1) donc les conditions de causalité et d'inversibilité sont vérifiées, on accepte donc le modèle.

- **Test concernant les paramètres du modèle ARIMA (2, 2, 4) :**
 - Les rapports de Student des coefficients du modèle (5.284839 et 7.566764) sont en valeur absolue supérieurs à 1.96, ils sont donc significativement différents de zéro ceci est confirmé par les probabilités de nullité de ces coefficients (0.0000 et 0.0000) qui sont inférieures à 0.05.
 - Les racines du polynômes autorégressif et moyenne mobile sont en module supérieures à 1 (leurs inverses fournis par *Eviews* sont inférieurs à 1) donc les conditions de causalité et d'inversibilité sont vérifiées), on accepte donc le modèle.

4.4 Choix du modèle :

Tableau 5. 11 : Choix du meilleur modèle pour le PIB

<i>Le modèle</i>	R^2	\bar{R}^2	<i>BIC</i>	<i>AIC</i>
ARIMA (2, 2, 4)	0.404895	0.376556	26.30762	26.20888
ARIMA (2, 2, 0)	0.192239	0.192239	26.47682	26.42745

Suivant le critère du pouvoir prédictif et les critères d'informations mentionnés dans le tableau ci-dessus, le modèle *ARIMA* (2, 2, 4) sera choisi pour représenter le processus.

En effet :

- Les coefficients R^2 et \bar{R}^2 du modèle *ARIMA* (2, 2, 4) sont supérieurs à ceux du modèle *ARIMA* (2, 2, 0).
- Les critères d'information AIC et BIC du modèle *ARIMA* (2, 2, 4) leur sont inférieurs.

Donc on retient le modèle : **ARIMA (2, 2, 4)** pour représenter la chronique.

4.5 Test sur les résidus :

A partir de la représentation graphique la série résiduelle, actuelle et estimée *figure 5.12*

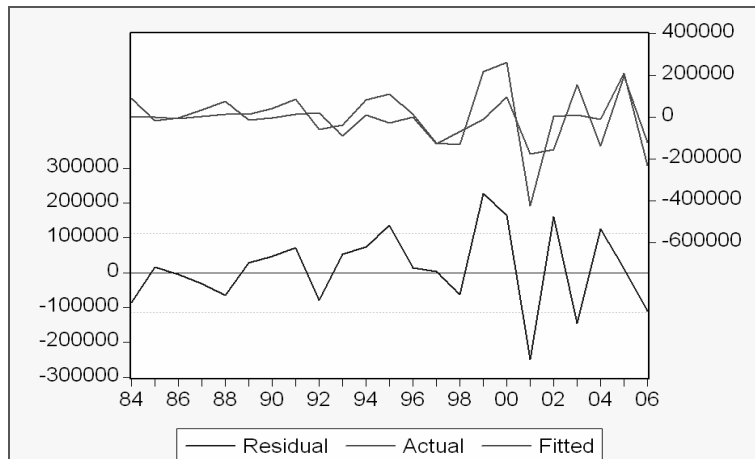


Figure 5. 12: Représentation graphique de la série résiduelle, actuelle et estimée de PIB

Nous constatons que le graphe de la série estimée est semblable à celui de la série réelle à quelques pics près, donc le modèle explique bien à 40.4895 % la variation totale de la variable dépendante ($R^2 = 0.404895$).

- **Corrélogramme des résidus du modèle ARIMA (2, 2, 4) :**

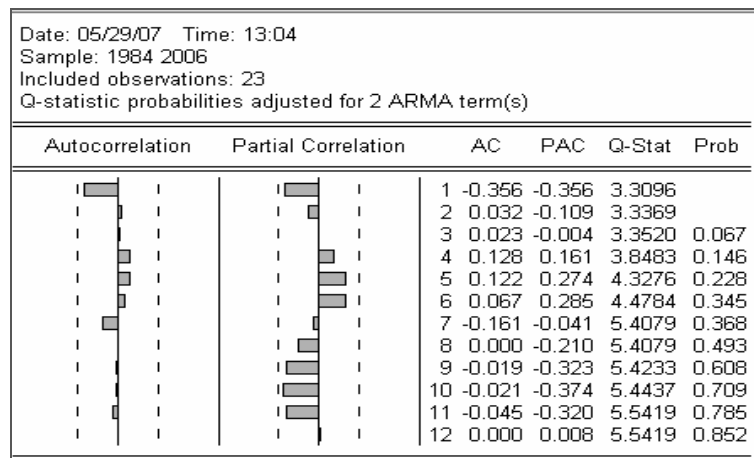


Figure 5. 13 : Corrélogramme des résidus du modèle ARIMA (2, 2, 4)

L'analyse du corrélogramme des résidus montre que tous les termes sont à l'intérieur de l'intervalle de confiance, donc les résidus forment un bruit blanc.

Ceci est confirmé par la statistique de Ljung-Box (Q- stat) qui est inférieure à la valeur théorique $\chi^2(h)$ quelque soit le retard h, et spécialement pour h= 12, en effet on a :

Q- stat (12) = 5.5419 inférieure à $\chi^2(12) = 21.026$ lue au seuil 0.05, donc les résidus forment un bruit blanc.

- **Test de normalité sur les résidus de ARIMA (2, 2, 4) :**

L'histogramme de la distribution et les valeurs empiriques de Skewness et de Kurtosis et la statistique de Jarque- Bera sont donnés par la figure 5.14

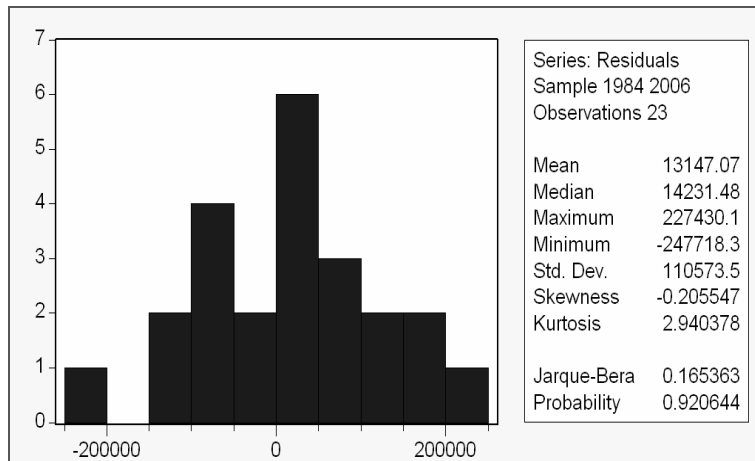


Figure 5. 14 : Distribution des résidus du modèle ARIMA (2, 2, 4)

On a :

Le coefficient de Skewness (asymétrie) = $-0,205547$

Le coefficient de Kurtosis (aplatissement) = $2,940378$

On construit alors les statistiques centrées réduites correspondantes, que l'on compare à 1,96 qui représente la valeur de la loi normale au seuil 5%, on aura alors :

$$v_1 = (\sqrt{(0,205547)}) / (6 / 27) = 0,96 \text{ inférieure à } 1,96$$

$$v_2 = (3 - 2,940378) / (\sqrt{24 / 27}) = 0,063 \text{ inférieure à } 1,96$$

L'hypothèse de normalité est donc acceptée, ce qui est confirmé par la statistique de Jarque-Berra, en effet : $JB = 0.165363$ inférieure à $\chi^2(2) = 5.991$

Nous concluons donc que les résidus forment un bruit blanc gaussien.

Le modèle ARIMA (2, 2, 4) est donc validé.

III. Etude de la série Prix essence P_{ess} :

La série P_{ess} représente l'évolution annuelle du prix de l'essence sur une période allant de 1980 à 2006.

1. Etude de la série :

1.1 Analyse préliminaire de la série brute:

La première étape consiste à l'analyse du graphe série, ceci permettra d'avoir une idée sur la nature et les caractéristiques de la série (tendance, saisonnalité.....etc.)

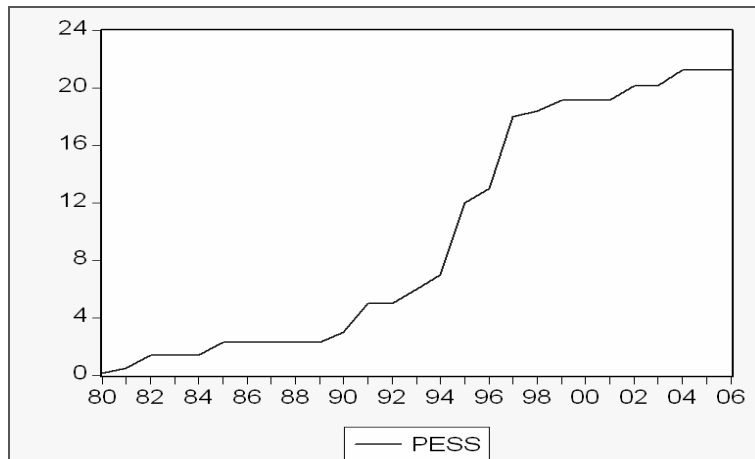


Figure 5.15 : Représentation graphique de la série Pess

La représentation graphique de la série P_{ess} *figure 5.15* nous montre que la série est affectée d'une tendance, il semble donc que la série P_{ess} soit non stationnaire.

1.2 Examen du corrélogramme de la série « Pess » :

Le corrélogramme de la série P_{ess} est représenté par la *figure 5.16*

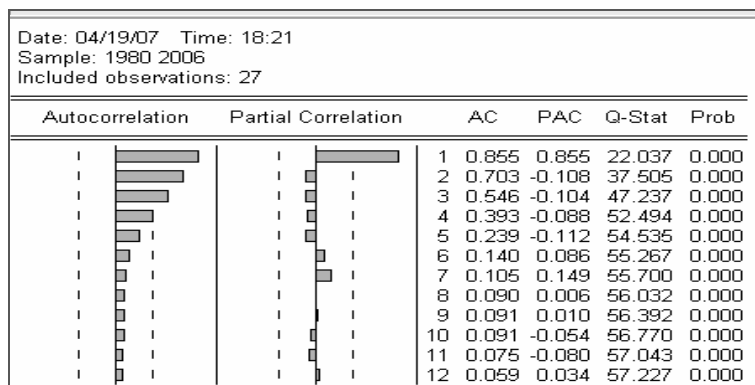


Figure 5.16 : Corrélogramme de la série Pess

L'analyse du corrélogramme simple et partiel de la série brute nous indique une non stationnarité de la série. En effet, la fonction d'autocorrélation simple et partielle diminue lentement vers zéro, on remarque aussi de nombreux pics sortant de la bande de confiance, ceci est caractéristique d'une série non stationnaire.

Il devient donc nécessaire de vérifier ceci en appliquant le Test de Dickey- Fuller pour déterminer de quel processus non stationnaire il s'agit et par la suite de le rendre stationnaire

1.3 Test de Dickey- Fuller sur la série Pess :

Nous allons appliquer le Test de Dickey – Fuller sur la série P_{ess} . Les résultats sont donnés par le *tableau 5.12*

Tableau 5. 12 : Test de Dickey Fuller sur Pess

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	2.49	2.79	$2.49 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	2.02	2.54	$2.02 \leq 2.54$	Constante non significativement différente de zéro
<i>Modèle 4</i>	0.022	-1.95	$0.022 \geq -1.95$	Série non stationnaire de type DS

Commentaires :

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 2.49) est inférieure à la valeur critique (2.79) lue dans la tables de Dickey- Fuller au seuil 5%, donc la tendance est non significativement différente de zéro.

On passe donc à l'estimation du modèle 5 pour tester la significativité de la constante.

- **Pour le Modèle 5 :** *modèle avec constante sans tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 2.02) est inférieure à la valeur critique (2.54) lue dans la tables de Dickey- Fuller au seuil 5%.

Donc la constante est non significativement différente de zéro.

On passe alors à l'estimation du modèle 4 pour vérifier l'existence ou non d'une racine unitaire:

- **Pour le Modèle 4 :** *modèle sans constante, ni tendance*

La statistique de Student (t- statistique = 0.022) est supérieure à la valeur critiques lue dans la table de Dickey – Fuller (= 1.956406) au seuil 5%, on accepte alors l'hypothèse H_0 , il existe une racine unitaire La série ' P_{ess} ' est donc non stationnaire, de type : DS.

La meilleure façon de la rendre stationnaire est de la différencier d'ordre 1.

La nouvelle série sera notée DP_{ess} .

2. Etude de la série différenciée « DPess » :

2.1 Effet de la différenciation :

La représentation graphique de la série DP_{ess} est donnée par la *figure 5.17*

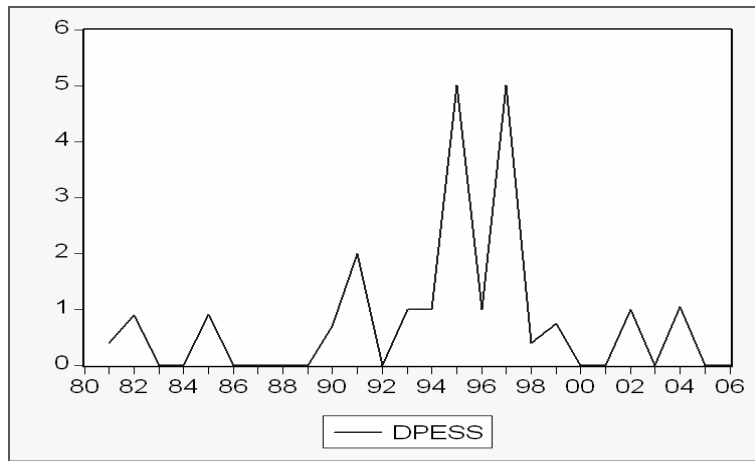


Figure 5.17: Représentation graphique de la série DP_{ess}

Nous remarquons que l'effet de la tendance a été absorbé. Il devient nécessaire d'appliquer le Test de Dickey-Fuller pour vérifier la stationnarité de la nouvelle série DP_{ess} .

2.2 Test de Dickey- Fuller sur la série DP_{ess} :

Nous allons appliquer le Test de Dickey – Fuller sur la série DP_{ess} . Les résultats sont donnés par le *tableau 4.13*

Tableau 5.13 : Test de Dickey Fuller sur DP_{ess}

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	0.088	2.79	$0.088 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	1.05	2.54	$1.05 \leq 2.54$	Constante non significativement différente de zéro
<i>Modèle 4</i>	-1.46	-1.95	$-1.46 \geq -1.95$	Série non stationnaire de type DS

Commentaires :

- **Pour le Modèle 6 :** *modèle avec constante et tendance*

Nous remarquons que :

La statistique de Student (t- statistic = 0.088) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.

La tendance est donc non significativement différente de zéro, passons alors à l'estimation du modèle 5 afin de tester la significativité de la constante

- **Pour le Modèle 5 :** *modèle avec constante sans tendance*

Nous remarquons que :

La statistique de Student (t- statistic = 1.05) est inférieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la constante est non significativement différente de zéro.

On passe alors à l'estimation du modèle 4 pour tester l'hypothèse H_0 : « existence d'une racine unitaire »

- **Pour le Modèle 4 : modèle sans constante ni tendance**

La statistique de Student (t- statistic = -1.46) est supérieure à la valeur critique lue dans la table de Dickey – Fuller (-1.95) lue au seuil 5%.

D'après les résultats, nous acceptons l'hypothèse nulle de racine unitaire H_0 alors la série DP_{ess} n'est pas stationnaire de type DS.

La meilleure façon de la rendre stationnaire est de la différencier d'ordre 2.

La nouvelle série sera notée DDP_{ess}

3. Etude de la série différenciée deux fois DDP_{ess} :

3.1 Effet de la différenciation :

La représentation graphique de la série DDP_{ess} est donnée par la *figure 5.18*:

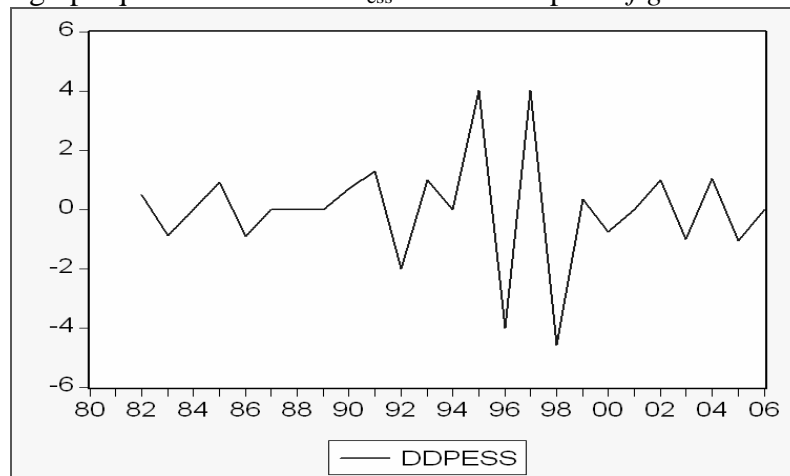


Figure 5. 18 : Représentation graphique de la série DDP_{ess}

Nous remarquons bien que l'effet de la tendance a été absorbé, nous allons appliquer le test de Dickey - Fuller pour vérifier que la nouvelle série DDP_{ess} est bien stationnaire.

3.2 Test de Dickey- Fuller sur la série DDP_{ess}

On refait les mêmes étapes déjà effectuées sur la série P_{ess} et DP_{ess} . Les résultats du test sont donnés par le *tableau 5.14*

Tableau 5. 14: Test de Dickey Fuller sur DDP_{ess}

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	0.45	2.79	$0.45 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	0.17	2.54	$0.17 \leq 2.54$	Constante non significativement différente de zéro
<i>Modèle 4</i>	-11.90	-1.95	$-11.90 \leq -1.95$	Série stationnaire

Commentaires :

- **Pour le Modèle 6 :** *modèle avec constante et tendance*

Nous remarquons que :

La statistique de Student (t- statistic = 0.45) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.

La tendance est donc non significativement différente de zéro, nous passons alors à l'estimation du modèle 5 pour tester la significativité de la constante.

- **Pour le Modèle 5 :** *modèle avec constante sans tendance*

Nous remarquons que :

La statistique de Student (t- statistic = 0.17) est inférieure à la valeur critique (2.54,) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la constante est non significativement différente de zéro.

Passons alors à l'estimation du modèle 4 pour tester H_0 : « existence d'une racine unitaire

- **Pour le Modèle 4 :** *modèle sans constante, ni tendance*

La statistique de Student (t- statistic = t = -11.90) est inférieure à la valeur critique lue dans la table de Dickey – Fuller (-1.95) au seuil 5%.

D'après les résultats, nous rejetons l'hypothèse nulle de racine unitaire H_0 alors nous acceptons l'hypothèse de stationnarité.

La série DDP_{ess} étant stationnaire, alors nous pouvons passer à la modélisation du processus en utilisant la méthode de Box & Jenkins.

4. Mise en œuvre de la méthode de Box and Jenkins :

4.1 Spécification du modèle :

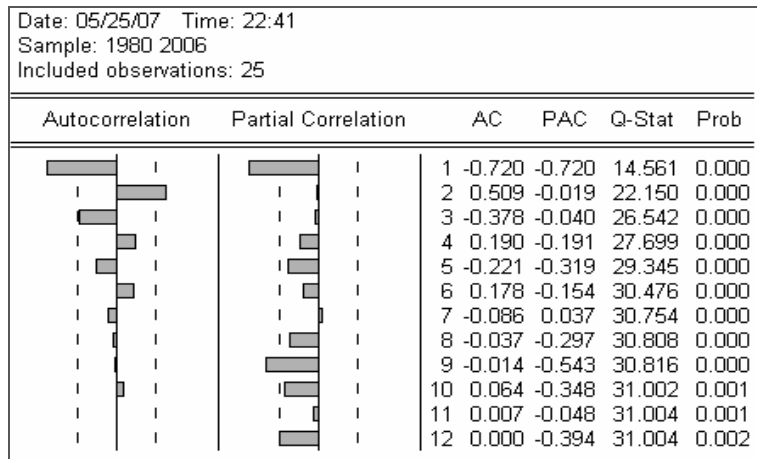


Figure 5. 19 : Corrélogramme de la série DDP_{ess}

L'examen du corrélogramme de la série DDP_{ess} *figure 5.19* montre que :

- Le corrélogramme simple a trois pics significatifs aux décalages 1, 2, 3 ;
- Le corrélogramme partiel a trois pics significatifs au décalage 1, 9 et 12

Nous avons plusieurs modèles candidats qui peuvent représenter la série DDP_{ess} , parmi lesquels nous avons retenu les modèles suivants :

$ARIMA(1, 2, 1)$, $ARIMA(1, 2, 8)$, $ARIMA(2, 2, 11)$, $ARIMA(9, 2, 3)$

4.2 Estimation des paramètres :

On utilise le logiciel *Eviews* pour effectuer l'estimation des paramètres de chaque modèle. Les résultats pour chaque modèle sont donnés dans les tableaux (5.15), (5.16), (5.17) et (5.18).

- *Estimation des paramètres de $ARIMA(1, 2, 1)$:*

Tableau 5. 15 : Estimation du modèle $ARIMA(1,2,1)$ du DDP_{ess}

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.730362	0.202548	-3.605868	0.0016
MA(1)	0.022880	0.299528	0.076385	0.9398
R-squared	0.519096	Mean dependent var		-0.037083
Adjusted R-squared	0.497237	S.D. dependent var		1.907027
S.E. of regression	1.352192	Akaike info criterion		3.520987
Sum squared resid	40.22533	Schwarz criterion		3.619158
Log likelihood	-40.25184	Durbin-Watson stat		1.979625
Inverted AR Roots	-.73			
Inverted MA Roots	-.02			

$R^2 = 0.519096$

• *Estimation des paramètres de ARIMA (1, 2, 8) :*

Tableau 5. 16 : Estimation du modèle ARIMA (1, 2, 8) du DDP_{ess}

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.708808	0.148868	-4.761311	0.0001
MA(8)	-0.860583	0.054046	-15.92324	0.0000
R-squared	0.679015	Mean dependent var		-0.037083
Adjusted R-squared	0.664425	S.D. dependent var		1.907027
S.E. of regression	1.104719	Akaike info criterion		3.116715
Sum squared resid	26.84890	Schwarz criterion		3.214886
Log likelihood	-35.40058	Durbin-Watson stat		1.810096
Inverted AR Roots	-.71			
Inverted MA Roots	.98	.69-.69i	.69+.69i	-.00-.98i
		-.69+.69i	-.69+.69i	-.98

$R^2 = 0.679015$

• *Estimation des paramètres de ARIMA (2, 2, 11) :*

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(2)	0.615151	0.175181	3.511519	0.0021
MA(11)	-0.866935	0.053272	-16.27379	0.0000
R-squared	0.626857	Mean dependent var		0.000000
Adjusted R-squared	0.609088	S.D. dependent var		1.941019
S.E. of regression	1.213582	Akaike info criterion		3.307972
Sum squared resid	30.92842	Schwarz criterion		3.406710
Log likelihood	-36.04167	Durbin-Watson stat		2.844022
Inverted AR Roots	.78	-.78		
Inverted MA Roots	.99	.83-.53i	.83+.53i	.41-.90i
		-.14-.98i	-.14+.98i	-.65-.75i
		-.65+.75i	-.95-.28i	-.95+.28i

Tableau 5. 17 : Estimation du Modèle ARIMA(2,2,11) du DDP_{ess}

$R^2 = 0.626857$

• **Estimation des paramètres de ARIMA (9, 2, 3) :**

L'estimation du modèle ARIMA (9, 2, 3) est donnée par le tableau ci-dessous :

Tableau 5. 18 : Estimation du modèle ARIMA (9, 2, 3) du DDP_{ess}

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(9)	0.121088	0.163673	0.739814	0.4716
MA(3)	-0.978036	0.068703	-14.23569	0.0000
R-squared	0.497733	Mean dependent var		-0.043750
Adjusted R-squared	0.461857	S.D. dependent var		2.319617
S.E. of regression	1.701631	Akaike info criterion		4.017520
Sum squared resid	40.53767	Schwarz criterion		4.114094
Log likelihood	-30.14016	Durbin-Watson stat		2.954012
Inverted AR Roots	.79	.61+.51i	.61-.51i	.14+.78i
	.14-.78i	-.40-.68i	-.40+.68i	-.74+.27i
	-.74-.27i			
Inverted MA Roots	.99	-.50+.86i	-.50-.86i	

$R^2 = 0.497733$

4.3 Adéquation des modèles :

• **Test concernant les paramètres du modèle ARIMA (1, 2, 1) :**

- Le rapport de Student du coefficient du modèle moyenne mobile est en valeur absolue (0.076385) inférieur à 1.96, il est donc significativement différent de zéro chose confirmée par la probabilité de nullité de ce coefficient qui est supérieure à 0.05 (0.9398), le modèle est donc rejeté.

• **Test concernant les paramètres du modèle ARIMA (1, 2, 8) :**

- Les rapports de Student des coefficients du modèle sont en valeur absolue (4.761311 et 15.92324) supérieur à 1.96, ils sont donc significativement différents de zéro chose confirmée par les probabilités de nullité de ces coefficients qui sont inférieures à 0.05 (0.0001 et 0.0000).
- Les racines du polynômes autorégressif et moyenne mobile sont en module supérieures à 1 (leurs inverses fournis par *Eviews* sont inférieurs à 1) donc les conditions de causalité et d'inversibilité sont vérifiées.

Le modèle ARIMA (1, 2, 8) est donc accepté.

• **Test concernant les paramètres du modèle ARIMA (2, 2, 11) :**

- Les rapports de Student des coefficients du modèle sont en valeur absolue (3.511519 et 16.27379) supérieurs à 1.96, ils sont donc significativement différents de zéro chose confirmée par la probabilité de nullité de ces coefficients qui est inférieur à 0.05 (0.0021 et 0.0000).
- Les racines du polynômes autorégressif et moyenne mobile sont en module supérieures à 1 (leurs inverses fournis par *Eviews* sont inférieurs à 1) donc les conditions de causalité et d'inversibilité sont vérifiées

Le modèle ARIMA (2, 2, 11) est donc accepté.

• **Test concernant les paramètres du modèle ARIMA (9, 2, 3) :**

- Le rapport de Student du coefficient du modèle autorégressif est en valeur absolue (0.739814) inférieur à 1.96, il est donc significativement différent de zéro chose confirmée par la probabilité de nullité de ce coefficient qui est supérieure à 0.05 (0.4716), le modèle est donc rejeté.

4.4 Choix du modèle :

Tableau 5. 19 : Choix du meilleur modèle de DDP_{ess}

<i>Le modèle</i>	R^2	\bar{R}^2	<i>BIC</i>	<i>AIC</i>
ARIMA (1, 2, 8)	0.679015	0.664425	3.214886	3.116715
ARIMA (2, 2, 11)	0.626857	0.609088	3.406710	3.307972

Suivant le critère du pouvoir prédictif et les critères d'information mentionnées dans le tableau (4.19), le modèle ARIMA (1, 2, 8) sera choisi pour représenter le processus.

En effet :

- Les coefficients R^2 et \bar{R}^2 du modèle ARIMA (1, 2, 8) sont supérieurs à ceux des modèles ARIMA (2, 2, 11) et ARIMA (1, 2, 1)
- Les critères AIC et BIC du modèle ARIMA (1, 2, 8) leur sont inférieurs.

4.5 Test sur les résidus :

La représentation graphique de la série résiduelle, actuelle et estimée est représentée par le graphique (5.20)

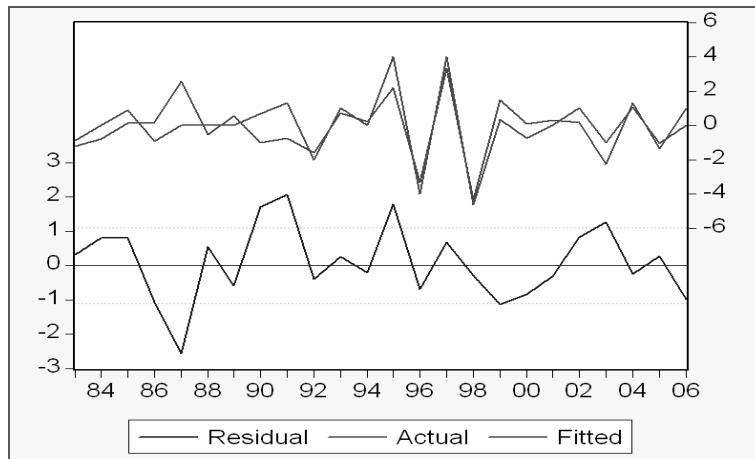


Figure 5. 20 : Représentation graphique de la série résiduelle, actuelle et estimée de P_{ess}

Nous constatons que le graphe de la série estimée est presque semblable à celui de la série réelle à quelques pics près, donc le modèle explique à 63.56 % la variation total de la variable dépendante ($R^2 = 0.635660$)

- **Corrélogramme des résidus du modèle ARIMA (1, 2, 8) :**

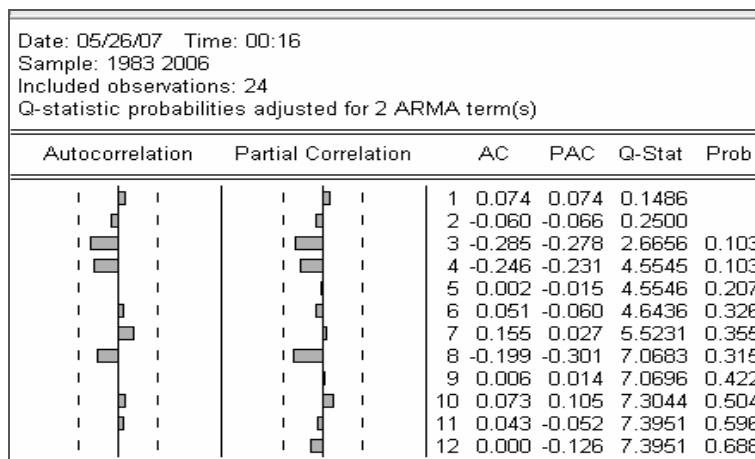


Figure 5. 21 : Corrélogramme des résidus du modèle ARIMA(1, 2, 8)

L'analyse du corrélogramme des résidus, illustré par la *figure 5.21* montre que tous les termes sont à l'intérieur de l'intervalle de confiance, donc les résidus forment un bruit blanc, ceci est confirmé par la statistique de Ljung-Box (Q -stat) qui est inférieure à la valeur théorique de $\chi^2(h)$ quelque soit le retard h , et spécialement pour $h=12$ on a :

Q -stat (12) = 7.3951 inférieure à $\chi^2(12) = 21.026$ Au seuil 0.05, donc les résidus forment un bruit blanc.

- **Test de normalité sur les résidus de ARIMA (1, 2, 8) :**

L'histogramme de la distribution et les valeurs empiriques de Skewness et de Kurtosis et la statistique de Jarque-Bera sont données par la *figure 5.22*

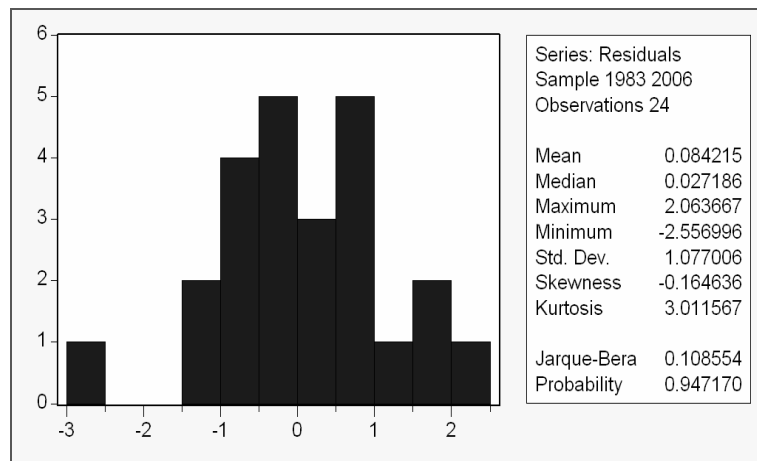


Figure 5. 22 : Histogramme de la distribution des résidus du modèle ARIMA (1, 2, 8)

Le coefficient de Skewness (asymétrie) = $-0,164636$

Le coefficient de Kurtosis (aplatissement) = $3,011567$

On construit alors les statistiques centrées réduites correspondantes que l'on compare à 1,96 qui représente la valeur de la loi normale au seuil 5%.

$$v_1 = (\sqrt{0,164636}) / (6 / 27) = 0,86 \text{ qui est inférieure à } 1,96$$

$$v_2 = (3,011567 - 3) / (\sqrt{(24 / 27)}) = 0,0122 \text{ qui est à inférieure à } 1,96$$

On accepte donc l'hypothèse de symétrie et l'hypothèse de distribution plate, l'hypothèse de normalité est donc acceptée, ce qui est confirmé par la statistique de Jarque- Berra, en effet : $JB = 0.108554$ inférieure à $\chi^2(2) = 5.991$

Donc on en conclut que les résidus forment un bruit blanc gaussien.

Le modèle $ARIMA(1, 2, 8)$ est donc validé, nous pouvons passer alors à la dernière phase qui est la prévision.

Annexe 6 : Stationnarisation des séries consommations

1. Test de stationnarité de C_{GPLc} :

a) Test de Dickey- Fuller sur la série C_{GPLc} :

Nous allons appliquer le Test de Dickey – Fuller sur la série C_{GPLc}

Nous utilisons le Test programmé par le logiciel *EViews*. Le résultat est donné dans le tableau ci-dessous :

Tableau 6. 1 : Test de Dickey- Fuller sur C_{GPLc}

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	2.16	2.79	$2.16 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	1.30	2.54	$1.30 \leq 2.54$	Constante non significativement différente de zéro
<i>Modèle 4</i>	-0.05	-1.95	$-0.05 \geq -1.95$	Série non stationnaire de type DS

Commentaires

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 2.16) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la tendance est non significativement différente de zéro.

Passons donc à l'estimation du modèle 5 pour tester la significativité de la constante.

- **Pour le Modèle 5 :** *modèle avec constante sans tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 1.30) est inférieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la constante est non significativement différente de zéro.

Passons alors à l'estimation du modèle 4 pour vérifier l'existence ou non d'une racine unitaire.

- **Pour le Modèle 4 :** *modèle sans constante, ni tendance*

On teste l'hypothèse nulle H_0 : « Existence d'une racine unitaire » :

La statistique de Student (t- statistique = -0.05) est supérieure à la valeur critique lue dans la table de Dickey – Fuller (-1.95) au seuil 5%, on accepte alors l'hypothèse de non stationnarité H_0 , il existe donc au moins une racine unitaire, la série C_{GPLc} est donc non stationnaire de type : DS.

La meilleure façon de la rendre stationnaire est de la différencier une fois.

La nouvelle série sera notée DC_{GPLc}

b) Test de Dickey- Fuller sur la série DC_{GPLc} :

Nous allons appliquer le Test de Dickey – Fuller sur la série DC_{GPLc} .

Tableau 6. 2 : Test de Dickey- Fuller sur DC_{GPLc}

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	0.96	2.79	$2.16 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	1.28	2.54	$1.28 \leq 2.54$	Constante non significativement différente de zéro
<i>Modèle 4</i>	-1.24	-1.95	$-1.24 \geq -1.95$	Série non stationnaire de type DS

Commentaires

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 0.96) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la tendance est non significativement différente de zéro.

Nous passons donc à l'estimation du modèle 5 pour tester la significativité de la constante.

- **Pour le Modèle 5 :** *modèle avec constante sans tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 1.28) est inférieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la constante est non significativement différente de zéro.

Nous passons alors à l'estimation du modèle 4 pour vérifier l'existence ou non d'une racine unitaire.

- **Pour le Modèle 4 :** *modèle sans constante, ni tendance*

On teste l'hypothèse nulle H_0 : « Existence d'une racine unitaire » :

La statistique de Student (t- statistique = -1.24) est supérieure à la valeur critique lue dans la table de Dickey – Fuller (-1.95) au seuil 5%, on accepte alors l'hypothèse de non stationnarité H_0 , il existe donc au moins une racine unitaire, la série DC_{GPLc} est donc non stationnaire de type : DS.

La meilleure façon de la rendre stationnaire est de la différencier.

La nouvelle série sera notée DDC_{GPLc}

c) *Test de Dickey- Fuller (DF) sur la série DDC_{GPLc} :*

Nous allons appliquer le Test de Dickey – Fuller sur la série DDC_{GPLc} .

Tableau 6. 3 : Test de Dickey- Fuller sur DDC_{GPLc}

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	0.81	2.79	$2.16 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	0.16	2.54	$0.16 \leq 2.54$	Constante non significativement différente de zéro
<i>Modèle 4</i>	-5.83	-1.95	$-5.83 \leq -1.95$	Série stationnaire

Commentaire :

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 0.81) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la tendance est non significativement différente de zéro.

Nous passons donc à l'estimation du modèle 5 pour tester la significativité de la constante.

- **Pour le Modèle 5 :** *modèle avec constante sans tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 0.16) est inférieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la constante est non significativement différente de zéro.

On passe alors à l'estimation du modèle 4 pour vérifier l'existence ou non d'une racine unitaire.

- **Pour le Modèle 4 :** *modèle sans constante, ni tendance*

On teste l'hypothèse nulle H_0 : « Existence d'une racine unitaire » :

La statistique de Student (t- statistique = -5.83) est inférieure à la valeur critique lue dans la table de Dickey – Fuller (-1.95) au seuil 5%, on rejette alors l'hypothèse de non stationnarité H_0 , et donc la série DDC_{GPLc} est stationnaire.

2. Test de stationnarité de C_{gas} :

a) Test de Dickey- Fuller sur la série C_{gas} :

Nous utilisons le Test programmé par le logiciel *EVIIEWS*. Le résultat est donné dans le tableau ci-dessous :

Tableau 6. 4 : Test de Dickey- Fuller sur C_{gas}

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	2.20	2.79	$2.20 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	0.25	2.54	$0.25 \leq 2.54$	Constante non significativement différente de zéro
<i>Modèle 4</i>	0.99	-1.95	$0.99 \geq -1.95$	Série non stationnaire de type DS

Commentaires :

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 2.20) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la tendance est non significativement différente de zéro.

On passe donc à l'estimation du modèle 5 pour tester la significativité de la constante.

- **Pour le Modèle 5 :** *modèle avec constante sans tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 0.25) est inférieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la constante est non significativement différente de zéro.

Nous passons à l'estimation du modèle 4 pour vérifier l'existence ou non d'une racine unitaire.

- **Pour le Modèle 4 :** *modèle sans constante, ni tendance*

On teste l'hypothèse nulle H_0 : « Existence d'une racine unitaire » :

La statistique de Student (t- statistique = 0.99) est supérieure à la valeur critique lue dans la table de Dickey – Fuller (-1.95) au seuil 5%, on accepte alors l'hypothèse de non stationnarité H_0 , il existe donc au moins une racine unitaire, la série C_{gas} est donc non stationnaire de type : DS.

La meilleure façon de la rendre stationnaire est de la différencier une fois.

La nouvelle série sera notée DC_{gas}

b) Test de Dickey- Fuller sur la série DC_{gas} :

Nous allons appliquer le Test de Dickey – Fuller sur la série DC_{gas} .

Tableau 6. 5 : Test de Dickey- Fuller sur DC_{gas}

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
Modèle 6	1.03	2.79	$1.03 \leq 2.79$	Tendance non significativement différente de zéro
Modèle 5	1.07	2.54	$1.07 \leq 2.54$	Constante non significativement différente de zéro
Modèle 4	-0.71	-1.95	$-0.71 \geq -1.95$	Série non stationnaire de type DS

Commentaires

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 1.03) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la tendance est non significativement différente de zéro.

Nous passons donc à l'estimation du modèle 5 pour tester la significativité de la constante.

- **Pour le Modèle 5 :** *modèle avec constante sans tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 1.07) est inférieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la constante est non significativement différente de zéro.

Nous passons alors à l'estimation du modèle 4 pour vérifier l'existence ou non d'une racine unitaire.

- **Pour le Modèle 4 :** *modèle sans constante, ni tendance*

On teste l'hypothèse nulle H_0 : « Existence d'une racine unitaire » :

La statistique de Student (t- statistique = -0.71) est supérieure à la valeur critique lue dans la table de Dickey – Fuller (-1.95) au seuil 5%, on accepte alors l'hypothèse de non stationnarité H_0 , il existe donc au moins une racine unitaire et la série DC_{gas} est donc non stationnaire de type : DS.

La meilleure façon de la rendre stationnaire est de la différencier une fois.

La nouvelle série sera notée DDC_{gas}

c) *Test de Dickey- Fuller sur la série DDC_{gas} :*

Nous allons appliquer le Test de Dickey – Fuller sur la série DDC_{gas} .

Tableau 6. 6 : Test de Dickey- Fuller sur DDC_{gas}

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	0.73	2.79	$0.73 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	0.38	2.54	$0.38 \leq 2.54$	Constante non significativement différente de zéro
<i>Modèle 4</i>	-9.69	-1.95	$-9.69 \geq -1.95$	Série stationnaire

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 0.73) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la tendance est non significativement différente de zéro.

Nous passons donc à l'estimation du modèle 5 pour tester la significativité de la constante.

- **Pour le Modèle 5 :** *modèle avec constante sans tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 0.38) est inférieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la constante est non significativement différente de zéro.

Passons alors à l'estimation du modèle 4 pour vérifier l'existence ou non d'une racine unitaire.

- **Pour le Modèle 4 :** *modèle sans constante, ni tendance*

On teste l'hypothèse nulle H_0 : « Existence d'une racine unitaire » :

La statistique de Student (t- statistique = -9.69) est inférieure à la valeur critique lue dans la table de Dickey – Fuller (-1.95) au seuil 5%, on rejette alors l'hypothèse de non stationnarité H_0 , la série DDC_{gas} est donc stationnaire.

3. Test de stationnarité de C_{ess} :

a) Test de Dickey- Fuller sur la série C_{ess} :

Nous allons appliquer le Test de Dickey – Fuller sur la série C_{ess} .

Nous utilisons le Test programmé par le logiciel *EViews*. Le résultat est donné dans le tableau ci-dessous :

Tableau 6. 7 : Test de Dickey- Fuller sur C_{ess}

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	2.41	2.79	$2.41 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	3.47	2.54	$3.47 \geq 2.54$	Constante significativement différente de zéro
	-3.25	-3.71	$-3.25 \geq -3.71$	Série non stationnaire de type DS

On a :

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 2.41) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la tendance est non significativement différente de zéro.

Passons donc à l'estimation du modèle 5 pour tester la significativité de la constante.

- **Pour le Modèle 5 :** *modèle avec constante sans tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 3.47) est supérieure à la valeur critique (= 2.54) lue dans la table de Dickey- Fuller au seuil 5%.

La constante est donc significativement différente de zéro.

On passe alors au test de l'hypothèse nulle H_0 : « Existence d'une racine unitaire » :

On a :

La statistique de Student (t- statistique -3.25) est supérieure à la valeur critique lue dans la table de Dickey – Fuller (-3.71) au seuil 5%, on accepte alors l'hypothèse de non stationnarité H_0 . Il existe donc au moins une racine unitaire, la série C_{ess} est donc non stationnaire de type : DS.

La meilleure façon de la rendre stationnaire est de la différencier une fois.

La nouvelle série sera notée DC_{ess}

b) Test de Dickey- Fuller (DF) sur la série 'DCess' :

Nous allons appliquer le Test de Dickey – Fuller sur la série 'DCess'.

Nous utilisons le Test programmé sur le logiciel Eviews. Le résultat est donné par le tableau ci-dessous :

Tableau 6. 8 : Test de Dickey- Fuller sur DCess

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
Modèle 6	2.16	2.79	$2.16 \leq 2.79$	Tendance non significativement différente de zéro
Modèle 5	0.61	2.54	$0.61 \leq 2.54$	Constante non significativement différente de zéro
Modèle 4	-3.12	-1.95	$-3.12 \leq -1.95$	Série stationnaire

Commentaires:

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 2.16) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la tendance est non significativement différente de zéro.

On passe donc à l'estimation du modèle 5 pour tester la significativité de la constante.

- **Pour le Modèle 5 :** *modèle avec constante sans tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 0.61) est inférieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la constante est non significativement différente de zéro.

Passons alors à l'estimation du modèle 4 pour vérifier l'existence ou non d'une racine unitaire.

- **Pour le Modèle 4 :** *modèle sans constante, ni tendance*

On teste l'hypothèse nulle H_0 : « Existence d'une racine unitaire » :

La statistique de Student (t- statistique = -3.12) est inférieure à la valeur critique lue dans la table de Dickey – Fuller (-1.95) au seuil 5%, on rejette alors l'hypothèse de non stationnarité H_0 , et donc la série DC_{ess} est stationnaire.

4. Test de stationnarité de P_{gas} :

a) Test de Dickey- Fuller sur la série P_{gas} :

Nous allons appliquer le Test de Dickey – Fuller sur la série C_{gas} .

Nous utilisons le Test programmé sur le logiciel Eviews. Le résultat est donné par le tableau ci-dessous :

Tableau 6. 9 : Test de Dickey- Fuller sur P_{gas}

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
Modèle 6	2.49	2.79	$2.49 \leq 2.79$	Tendance non significativement différente de zéro
Modèle 5	1.70	2.54	$1.70 \leq 2.54$	Constante non significativement différente de zéro
Modèle 4	2.07	-1.95	$2.07 \geq -1.95$	Série non stationnaire de type DS

Commentaires

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 2.49) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la tendance est non significativement différente de zéro.

Passons donc à l'estimation du modèle 5 pour tester la significativité de la constante.

- **Pour le Modèle 5 :** *modèle avec constante sans tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 1.70) est inférieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la constante est non significativement différente de zéro.

On passe alors à l'estimation du modèle 4 pour vérifier l'existence ou non d'une racine unitaire.

- **Pour le Modèle 4 :** *modèle sans constante, ni tendance*

On teste l'hypothèse nulle H_0 : « Existence d'une racine unitaire » :

La statistique de Student (t- statistique = 2.07) est supérieure à la valeur critique lue dans la table de Dickey – Fuller (-1.95) au seuil 5%, on accepte alors l'hypothèse de non stationnarité H_0 , il existe donc au moins une racine unitaire, la série P_{gas} est donc non stationnaire de type : DS.

La meilleure façon de la rendre stationnaire est de la différencier une fois.

La nouvelle série sera notée DP_{gas}

b) Test de Dickey- Fuller (DF) sur la série 'DPgas' :

Nous allons appliquer le Test de Dickey – Fuller sur la série 'DCgas'.

Nous utilisons le Test programmé sur le logiciel Eviews. Le résultat est donné par le tableau ci-dessous :

Tableau 6. 10 : Test de Dickey- Fuller sur DPgas

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	0.69	2.79	$0.69 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	2.03	2.54	$2.03 \leq 2.54$	Constante non significativement différente de zéro
<i>Modèle 4</i>	-1.16	-1.95	$-1.16 \geq -1.95$	Série non stationnaire de type DS

Commentaires :

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 0.69) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la tendance est non significativement différente de zéro.

On passe donc à l'estimation du modèle 5 pour tester la significativité de la constante.

- **Pour le Modèle 5 :** *modèle avec constante sans tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 2.03) est inférieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la constante est non significativement différente de zéro.

Passons alors à l'estimation du modèle 4 pour vérifier l'existence ou non d'une racine unitaire.

- **Pour le Modèle 4 :** *modèle sans constante, ni tendance*

On teste l'hypothèse nulle H_0 : « Existence d'une racine unitaire » :

La statistique de Student (t- statistique = -1.16) est supérieure à la valeur critique lue dans la table de Dickey – Fuller (-1.95) au seuil 5%, on accepte alors l'hypothèse de non stationnarité H_0 , il existe donc au moins une racine unitaire et donc la série « DPgas» est non stationnaire de type : DS.

La meilleure façon de la rendre stationnaire est de la différencier une fois.

La nouvelle série sera notée « DDP_{gas} »

c) *Test de Dickey- Fuller sur la série DDPgas :*

Nous allons appliquer le Test de Dickey – Fuller sur la série ‘DDPgas’.
 Nous utilisons le Test programmé sur le logiciel Eviews. Le résultat est donné par le tableau ci-dessous :

Tableau 6. 11 : Test de Dickey- Fuller sur DDPgas

Modèle	t- statistic	Valeur critique de DF au seuil 5%	Comparaison entre les deux valeurs	Conclusion
<i>Modèle 6</i>	0.08	2.79	$0.08 \leq 2.79$	Tendance non significativement différente de zéro
<i>Modèle 5</i>	0.42	2.54	$0.42 \leq 2.54$	Constante non significativement différente de zéro
<i>Modèle 4</i>	-10.26	-1.95	$-10.26 \leq -1.95$	Série stationnaire

On a:

- **Pour le Modèle 6 :** *modèle avec constante et tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 0.08) est inférieure à la valeur critique (2.79) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la tendance est non significativement différente de zéro.

On passe donc à l’estimation du modèle 5 pour tester la significativité de la constante.

- **Pour le Modèle 5 :** *modèle avec constante sans tendance déterministe*

Nous remarquons que :

La statistique de Student (t- statistic = 0.42) est inférieure à la valeur critique (2.54) lue dans la table de Dickey- Fuller au seuil 5%.

Donc la constante est non significativement différente de zéro.

On passe alors à l’estimation du modèle 4 pour vérifier l’existence ou non d’une racine unitaire.

- **Pour le Modèle 4 :** *modèle sans constante, ni tendance*

On teste l’hypothèse nulle H_0 : « Existence d’une racine unitaire » :

La statistique de Student (t- statistique -10.26) est inférieure à la valeur critique lue dans la table de Dickey – Fuller (-1.95) au seuil 5%, on rejette alors l’hypothèse de non stationnarité H_0 , et donc la série « DDP_{gas}» est stationnaire.

Annexe 7 : Le Questionnaire

Dans le cadre de notre projet de modélisation et de prévision de la consommation nationale du GPL carburant, une étude de sondage, traitant le comportement du consommateur algérien pour le segment véhicule léger, a été élaborée.

Le questionnaire :



Ecole Nationale Polytechnique
Département du Génie Industriel

Dans le cadre de notre projet de fin d'étude intitulé : « modélisation et prévision de la consommation nationale du GPL carburant », nous effectuons un sondage sur le comportement du consommateur Algérien vis-à-vis du choix de son carburant

Veillez nous accorder un peu de votre temps pour répondre à ce questionnaire en cochant dans les cases appropriées :

1. Votre sexe : Femme Homme
2. Quel est votre statut social : Etudiant(e) Salarié(e) Retraité(e)
 Profession libérale Autre.....
3. Quel est votre age ? Moins de 25 ans Entre 25 ans et 35 ans
 Entre 35 et 55 ans Plus de 55 ans
4. Utilisez vous votre véhicule : Systématiquement Par nécessité
 Pour le travail Pour le loisir Autre.....
5. Quel carburant consommez-vous ? Essence super ou normal Essence sans plomb
 GPL carburant (GPL c) Gasoil
- Pourquoi ce choix ? Economique Ecologique Disponibilité du carburant
 Performant Autre (Précisez).....
6. Etes vous satisfaits du carburant que vous utilisez actuellement? (Justifiez votre réponse)
 Oui Non
.....
7. Etes vous prêts à changer de carburant si le prix de ce dernier aurait tendance à augmenter considérablement ?
 Oui Lequel Essence super ou normal Essence sans plomb Gasoil
 GPL carburant
- Pourquoi ce choix ?
- Non Pourquoi ?
8. Quel est le kilométrage effectué en moyenne par an ?
 Moins de 9000 km Entre 9000 km et 25000 km
 Entre 25000 et 35000 km Plus de 35000 km

9. Quelle est votre consommation moyenne de carburant (Litre /100 Km):

- 4 -5 6-7 8-9 10 et plus

10. Changez vous de voiture tous les :

- 0 ans – 2ans 3 ans - 5ans 6 ans – 10 ans Plus de 10 ans

11. Lors de l'achat d'un véhicule, quels sont les critères les plus dominants dans le choix de votre véhicule ?

Critère	Faible	Moyen	Fort
Le prix d'acquisition	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Consommation de carburant (type de carburant)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Puissance	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Marque	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Age du véhicule	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Caractéristiques techniques et robustesse	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Options de sécurité (Abs)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Caractéristiques environnementales (pot catalytique, émission de CO ₂)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Taille, ligne, couleur, agencement intérieur	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Service après vente	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

12. Pensez-vous à utiliser le transport urbain pour « économiser le véhicule », si le service de ce dernier s'améliorait ?

- Oui Non

13. Quel est votre moyen de transport en commun urbain préféré ? :

- Le taxi Le bus Le train

Pourquoi ?

- Economique Sécurisant Confortable Efficace(délais) Autre(précisez).....

14. Quel est d'après vous la classification des carburants par niveau de pollution ?

Carburant	Peu polluant	Moyennement polluant	Très polluant
Essence super	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Essence normale	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Essence sans plomb	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Gasoil	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
GPL carburant	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

15. Pourquoi d'après vous le GPL carburant n'est pas très convoité ? Est-ce un problème:

- de sécurité de disponibilité au niveau des stations services
 de méconnaissance du GPL carburant. de la conversion du véhicule en GPL c
 d'espace (encombrant au niveau de la malle) je ne sais pas
 d'usure du moteur Autre(précisez).....

16. Si vous ne consommez pas du GPL carburant et si on vous disait qu'il est *sécurisant, disponible, pas encombrant, peu polluant, le moins coûteux* sur le marché et que *l'installation*

du kit GPL est subventionnée par l'état, seriez vous prêts à le consommer ? (Justifiez votre réponse) Oui Non

.....
Fin du sondage.

Le sondage a été pratiqué sur une population variée (hétérogène) et dans des lieux différents et notamment : l'Ecole Polytechnique, Sonatrach DG, Sonatrach TRC, Hôpital Mustapha, Stations de service (Naftal et privées), Salle de sport, et site des anciens élèves de l'ENP.

Nous avons procédé par ailleurs à une segmentation par sexe, statut social et age à fin d'avoir une meilleure appréciation du comportement du consommateur.

Pour les trois segments d'étude comportementale, relative à l'appréciation de l'Essence (super, sans plomb et normal), le Gasoil et le GPL c, nous avons axé notre questionnaire sur la pratique de substitution et de l'appréciation des qualités de ces carburants.

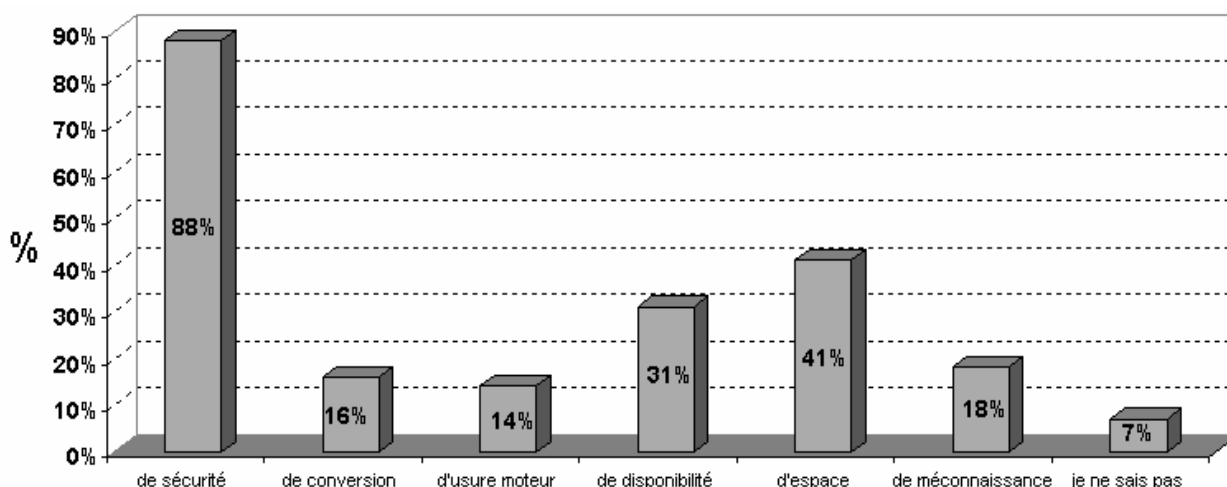


Figure 7. 1 : Graphique relatif à la non convoitise du GPLc (question 15)

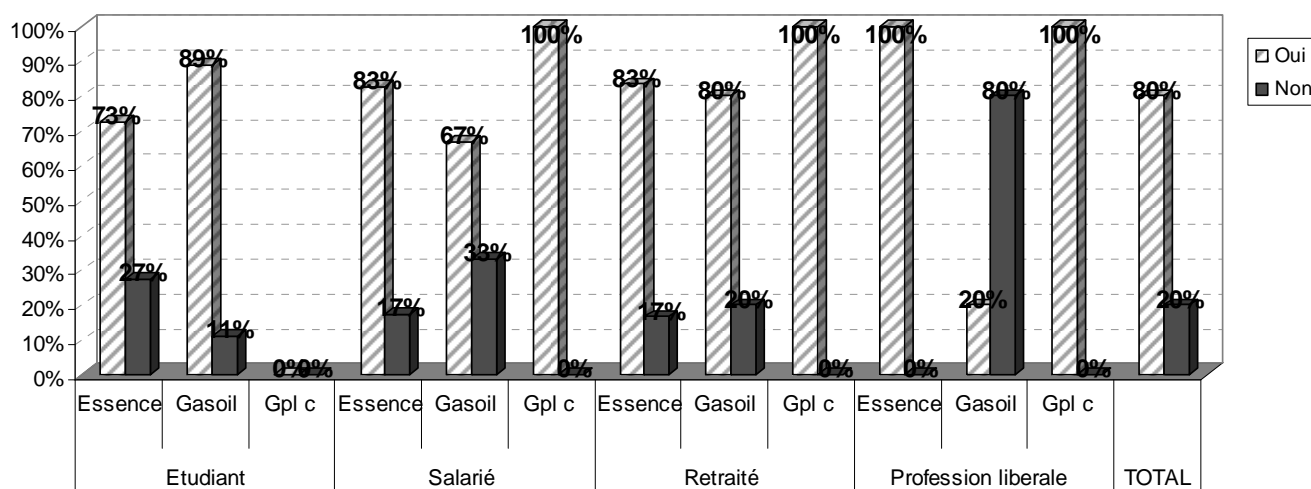


Figure 7. 2 : Graphique relatif aux taux de substitution d'un carburant par le GPLc (question 16).

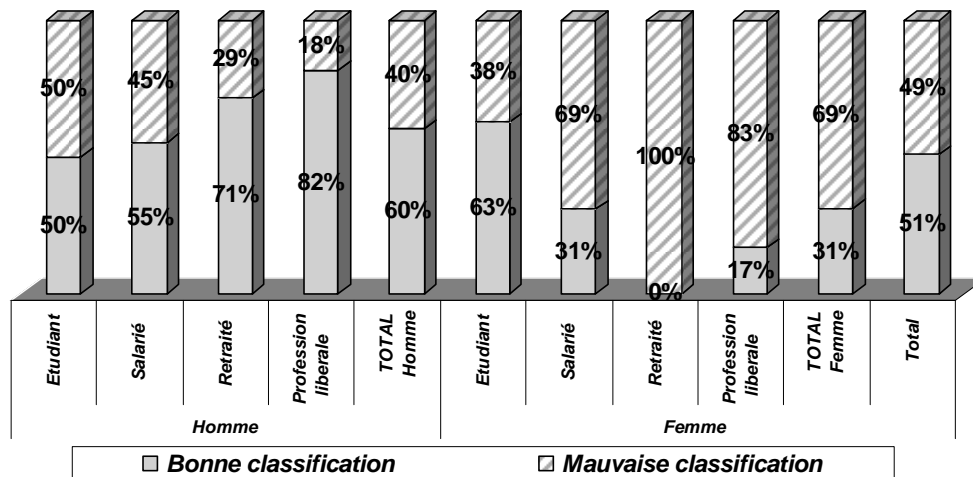


Figure 7.3 : Graphique relatif à la perception de la pollution du GPLc (question 14)

Graphique de substitution des carburants lors d'une augmentation des prix

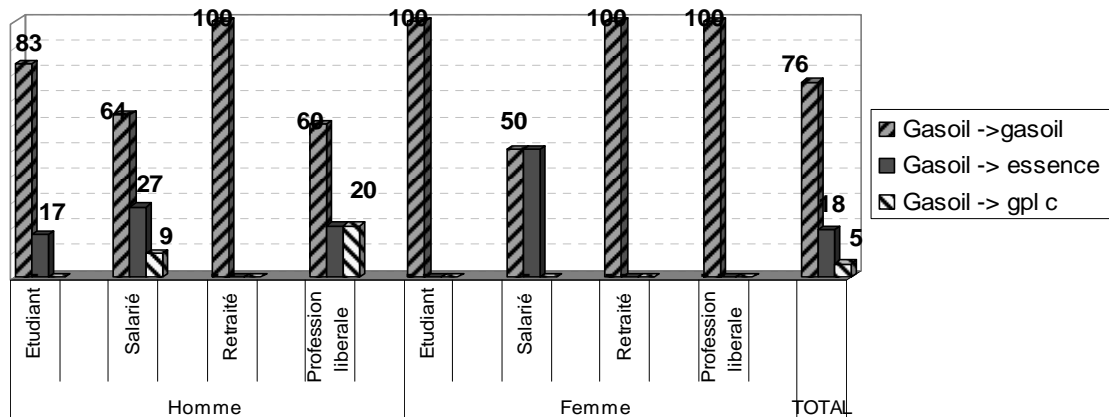


Figure 7.4 : Impact de l'augmentation du prix du gasoil

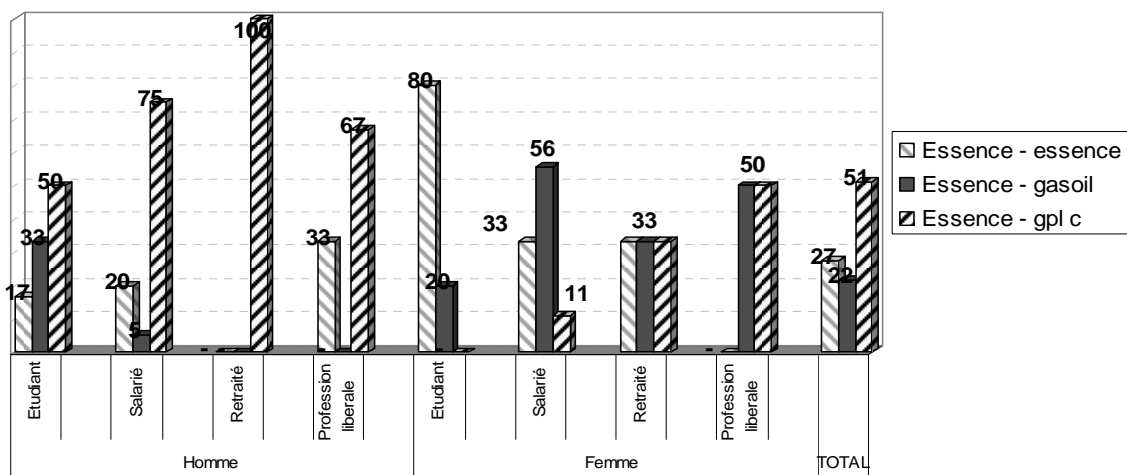


Figure 7.5 : Impact de l'augmentation du prix de l'essence

BIBLIOGRAPHIE

Les références

- **Ouvrages et documents**

[AIK 03] AIKAPACCHI B., 2003, Cours sur la régression multiple , Centre Interuniversitaire de calcul de Grenoble, France.

[ALV 03] ALVES D. B., 2003, La modélisation de gaz au Brésil à court et moyen terme, Institut Français du pétrole (IFP), Paris.

[AMB 05] AMBAPOUR S. et MASSAMBA C., 2005, Croissance économique et consommation d'énergie au Congo : Une analyse de terme de causalité, CNCEE.

[BOU 98] BOURBONNAIS R., 1998, Prévision des ventes, Edition DUNOD, Paris.

[BOU 02] BOURBONNAIS R., 2002, Econométrie, Edition DUNOD, Paris.

[BOU 06] BOUKABOUS A., 2006, Cours d'économétrie, Département du Génie Industriel Ecole Nationale Polytechnique (ENP), Alger.

[CHA 03] CHARPENTIER A., 2003, Cours de séries temporelles : théories et applications , Université Paris Dauphine, Paris.

[DAU 05a] Cours sur les séries temporelles, 2001, Paris Dauphine (France).

[FIO 99] FIORDALISO A., 1999, Systèmes flous et prévision de séries temporelles, Edition HERMES SCIENCE, Paris.

[FOU 04] FOUCART T., 2004, Colinéarité et régression linéaire, Université Pierre et Marie Curie, Paris.

[JON 99] JOHNSTON J. et DINARDO J., 1999, Méthodes Econométriques, Edition ECONOMICA, Paris.

[KAL 02] KALIKA M., HELFER J. et ORSONI J., 2002, Management stratégie et organisation, 4^{ième} édition VUIBERT, Paris.

[KHE 05] KHEMIRI M. H., 2005, La contribution des investissements en stocks à la volatilité du PIB réel au Canada, Mémoire présenté comme exigence partielle de la maîtrise en économie, Université du Québec., Montréal.

[LAD 06] LADJOUZ S. et ZAIR Y., 2006, Modélisation et prévision de l'activité d'exploration pétrolière en Algérie, Mémoire de fin d'études d'ingénieur, Faculté de Mathématique, Département de recherche opérationnelle, Université des Sciences et de la Technologie Houari Boumédiène (USTHB), Alger.

[LAN 03] LANTZ F., 2003 , La modélisation des marchés de pétrole et des produits pétroliers, Distribuer par IFP (Institut Français du Pétrole), Paris.

[LAN 05] LANTZ F., 2005, La modélisation de la demande d'énergie dans le monde, Distribuer par IFP (Institut français du Pétrole), Paris.

[RAU 02] RAULT C. et L'HORTY Y. , 2002, « Inflation , salaire et SMIC, Quelles relations ? (*Analyse impulsionnelle*), Document de recherche centre d'Etude des Politiques Economiques (EPEE) de l'université d'Evry, Val d'Essonne (France).

[SH 05] Sonatrach, Rapport annuel 2005, 2006, Edition Sonatrach.

[TOT 04] TOTAL , 2004, Les chemins du développement durable- le GPL carburant, édition TOTAL.

[UPM 04] Université de Pierre et Marie Curie (UPM), 2004, Cours sur la régression, Paris.

- **Sitographie**

[ANTAGAZ] Comment fonctionne un moteur GPL c
http://www.antargaz.fr/part_gpl_carburant.html

[CFBP] *Comité Français du Butane et Propane.* www.cfbp.com

[EVE] *EViews* www.eviews.com

[MEM] *Ministère de l'Energie et des Mines* <http://www.mem-algeria.org/fr/index.htm>

[NAFTAL] www.naftal.dz

[NAFTEC] www.naftec-dz.com

[ONS] *Office Nationale des Statistiques* www.ons.dz

[OUTF] *Installation GPL carburant* <http://www.outrefranc.com/pratique/gpl.html>

[PRIMAGAZ] *GPL carburant*
<http://www.primagaz.fr/majic/pageServer/010100001z/fr/F.A.Q..html>

[SFC] *Société Française de Chimie* <http://www.sfc.fr/donnees/orga/gaznat/texgnat.htm#GPL>

[SH] *Sonatrach* www.sonatrach-dz.com

[UFIP] *Union française des industries pétrolières* <http://www.ufip.fr/index.php>

Les consultations

- **Ouvrages et documents**

[AKO 04] AKOSSOU A.Y. J. et PALM R. , 2004, Conséquences de la sélection de variable sur l'interprétation des résultats en régression linéaire multiple, Unité de statistique et d'informatique, Faculté universitaire des sciences agronomiques de Gembloux, Belgique.

[ARO 03] AROURI M.E. , 2003, Intégration financière et diversification internationale de portefeuilles : Une analyse multi variée , Université Paris X , Nanterre.

[AZI 07] AZIZA M., Quotidien d'Oran (édition du 5 Mars 2007) lundi 5), Et si l'Algérie venait à importer du Diesel ?, Algérie.

[AZZ 03] AZZOUG S., 2004, Présentation de la distribution des produits pétroliers en Algérie,Naftal , Le 22 Avril 2003, Séminaire sur l'énergie, Hôtel Aurassi Alger.

[BES 03] BESSE P. 2003 , Pratique de la modélisation statistique, Publications du laboratoire de statique et probabilités, Université Paul Sabatier, Toulouse III.

[BOY 04] BOYER E., 2004, Fondamentaux de la chaîne gazière, Edition IFP TRAINING, Paris.

[BOU 03] BOUKHLIFA K. ,2003 , Présentation des secteurs des mines, des hydrocarbures et e l'électricité, Ministère de l'Energie et des Mines, Le 22 Avril 2003, Séminaire sur l'énergie, Hôtel Aurassi Alger.

[DAUP 01b] Cours sur la modélisation VAR, 2001, Paris Dauphine (France).

[DOZ 05] DOZ M., Cours « séries temporelles linéaires », deuxième année 2004-2005, Ecole Nationale de la Statistique et de l'Administration Economique (ENSAE), Paris.

[DRA 07] DRAMANI L., 2007, Multi country model in CFA zone, University Cheikh Anta Diop of Dakar, Edition MPRA (Munich Personal RePEc).

[GAL 02] LE GALL P. , Les représentations du monde et les pensées analogiques des économètres , un siècle de modélisation en perspective, Revue des Sciences Humaines 2002 /1, N° 6, P. 39-64.

[IFP 04] Institut Français du Pétrole (IFP), 2004, Cours d'économétrie, Paris.

[IFP 05] IFP MOTEURS ET CARBURANTS, IFP moteurs et carburants un avantage concurrentiel, Edition IFP direction de la communication.

[LAB 06] LABRIET M. , 2006, Quelle contribution des nouvelles technologies de transport aux stratégies climatiques à long terme ? Modélisation et analyse technico-économique, Colloque : Transport, énergie et gaz à effet de serre : vers un rationnement ?, Lyon (France).

[LAN 04] LANTZ F., 2004, Analysis of external energy consumption in the car industry, 19th World Energy Congress, September 5-9; 2004, Sydney, Australia.

[LUB 07] LUBRANO M., 2007, Modèles VAR, Modèles VAR structurels à équations simultanées, Université de Renne 1, France.

[MEG 05] MEGHOUFEL M., MILOUDI A., 2005, Conception d'un modèle de prévision basé sur les réseaux neurones artificiels –Application consommations des carburants (essence-gasoil), Mémoire de projet de fin d'études d'ingénieur, Département du Génie Industriel, Ecole Natioanle Polytechnique, Alger.

[MEM 99] Ministère de l'Energie et des mines, évolution rétrospective du bilan énergétique national algérien de 1980 à 1999, édition ministère de l'énergie et des mines.

[MEM 05] Ministère de l'Energie et des Mines, bilan énergétique national de l'année 2005, 2006, édition ministère de l'énergie et des mines.

[MEM 06] Direction des systèmes d'information, des analyse économiques et de la documentation Ministère de l'Energie et des Mines, 2006, Résultats du secteur d'énergie et des mines pour l'année 2005, édition Ministère de L'Energie et des Mines.

[MOG 06] MOGLIANI M., 2006, La dollarisation et la stabilité macroéconomique en Equateur : analyse VAR de la volatilité, Mémoire Master recherche en Analyse et Politique économiques, L'Ecole des Hautes Etudes en Science Sociales, Paris.

[OUI 04] OUIS L. , 2004, Les tendances actuelles du marché du pétrole et leurs impacts sur la stratégie de développement de l'activité transport, Mémoire de MBA, Management Développement International Business School, Alger.

[PAG 03] PAGE L., 2006, une introduction générale aux outils économétriques, Université Paris 1.

[PAR 00] PARENT A. et RAULT C., 2000, La contribution des facteurs financiers à la stabilisation du régime bi-métallique : une approche VAR, Université Paris VIII Panthéon–Sorbonne et Université Paris I (France).

[PED 06] PEDESSAC J. , 2006, Le GPL carburant, Débats BIP ENEPRESS « les véhicules propres à la croisée des chemins » , Le mardi 20 juin 2006.

[PMT 04] Naftal, Plan à Moyen Terme 2005 – 2009, Décembre 2004, Edition Naftal

[RAH 06] RAHAL H.A., juin 2006, Etude comparative des modèles de prévision ARMA et des réseaux de neurones artificiels, mémoire de projet de fin d'études d'ingénieur, département du Génie Industriel, Ecole Nationale Polytechnique, Alger.

[REB O3] REBAH M. , 2003, Le GPL c pour 1000 Taxis, Journal la Dépêche (édition du 14 octobre 2003) ,Algérie.

[RID 05] RIDING A. , 2005, Prévisions des coûts liés aux prêts de financement par emprunt consentis en vertu de la LPPE et de la LFPEC , édition EQUINOX Management Consultants.

[SAL 94] SALMAT G. , 1994, Le retournement conjoncturel de 1992 et 1993 : Une modélisation VAR des postes de la demande, Université Paris 1, Paris.

[SH 04a] Sonatrach activité commerciale, décembre 2004, Commercialisation gaz et développement à l'international, 4^{ème} édition Sonatrach.

[SH 04b] Sonatrach activité commerciale, décembre 2004, Sahara blend et produits dérivés, édition Sonatrach.

[STU 06] STUDER E. , L'Algérie pourrait manquer d'essence , Le 13 Mars 2006, Blog Finance.

[TAV 04] TAVERA C., 2004, La modélisation VAR stationnaire « standard », 2004, Université de Renne 1, Renne.

[TOT 05] TOTAL , 2005, Maîtriser la chaîne gazière, édition Total.

[TRA 95] TRAORE F., 1995, L'impact des subventions américaines sur le prix mondial du coton : une approche par les modèles vectoriels autorégressifs Bayésiens, Centre d'Études et de Recherches sur le Développement International, Clermont-Ferrand (France).

- **Sitographie**

[BUR] BURTSCHY B. et MENENDIAN C., 1980, À propos de la prévision à court terme de la production industrielle, Revue de statistique appliquée, tome 28, n° 2, p 5-24. www.numdam.org/item?id=RSA_1980_28_2_5_0

[CON] Gpl, diesel ou essence ?
<http://conseils-auto.lacentrale.fr/conseils/pratique/gpl-diesel-ou-essence.html>

[GSEG] Guide en Sciences Economiques et de Gestion.
<http://www.ext.upmc.fr/urfist/Ecoline/statistiques/indexmonde1.html>

[HOR] Horizon GPL <http://www.horizongpl.com/>

[IND] Energie et matières premières <http://www.industrie.gouv.fr/energie/sommaire.htm>

[PRE] La presse algérienne www.aai-online.com/liens/press.php

[S&D] Science & décision www.science&décision.com