



ETUDE STATISTIQUE

Dossier

Analyse d'une série temporelle

Agriculture - Production laitière bovine en France

Thibault GIBERT
Master II Ingénierie de Projets
Année universitaire 2010 - 2011





| | | |
|------|--|----|
| 1. | La filière lait de vache en France et dans le monde | 3 |
| 2. | Présentation de la série | 4 |
| 3. | L'Analyse de la saisonnalité | 6 |
| 3.1. | Série en Indice | 6 |
| 3.2. | Tableau de Buys-Ballot | 7 |
| 3.3. | Tableau de Buys-Ballot classé..... | 8 |
| 3.4. | Test de Fisher - Composante Saisonnière ? Tendancielle ? | 9 |
| 3.5. | La Fonction d'autocorrélation | 13 |
| 4. | La sélection du schéma | 15 |
| 4.1. | La procédure de la bande..... | 15 |
| 4.2. | Le test de Buys-Ballot | 16 |
| 4.3. | Le test de Buys-Ballot sur la période 93-08 | 17 |
| 5. | Désaisonnalisation | 18 |
| 5.1. | Comparaisons des coefficients saisonniers obtenus par les méthodes « Moyenne Mobile » et « CensusX12 » 19 | |
| 5.2. | Comparaison des Corrélogrammes..... | 21 |
| 5.3. | Désaisonnalisation 93-08 (Additif) | 22 |
| 6. | Prévision par les méthodes traditionnelles..... | 26 |
| 6.1. | Etude de la non-Stationnarité | 26 |
| 6.2. | Prévision de la chronique par régression | 29 |
| 7. | Algorithme de Box et Jenkins sur série désaisonnalisée | 35 |
| 6.3. | Identification du processus ARMA de notre série | 35 |
| 6.4. | Remarques | 39 |



1. La filière lait de vache en France et dans le monde

3

Le Lait

700 millions de tonnes de lait seront produits en 2010 dans le monde selon la FAO et l'OCDE, soit environ 22 220 litres par seconde.



Cette production est composée à 85% de lait de vache, le restant étant du lait de bufflonne (12%), de chèvre ($\approx 2\%$), de brebis (1%) et d'autres mammifères (0,2%).

Ces dernières années, la production mondiale de lait n'a cessé d'augmenter : Approximativement 12 milliards de litres par an (20% en dix ans). Cette croissance est due en grande partie à l'Asie (Inde et Chine), l'Océanie (Nouvelle Zélande) et à l'Amérique du sud (Brésil, Argentine) et du nord.

L'Asie est la première région productrice de lait au monde (34%) suivie de l'Union Européenne (27%). Avec 9% de sa production exportée cette dernière couvre près du tiers des échanges mondiaux.

La France occupe le second rang parmi les pays producteurs de lait de vache au sein de l'union européenne avec 18% derrière l'Allemagne 21% et le 5^{ème} au niveau mondial. Le premier reste l'Inde.



Les produits laitiers

Seulement 7% de la production mondiale de lait se retrouve sur le marché des produits laitiers, marché dominé par deux grandes zones : l'Océanie (40%) et l'UE (33%).

Au sein de l'UE, les importations de produits laitiers représentent moins de 2% de sa consommation intérieure et les exportations s'élèvent à 10% de la production.

Les fabrications de l'industrie laitière donnent lieu à peu d'échange au niveau mondial. Parallèlement, on assiste à une forte consommation locale.

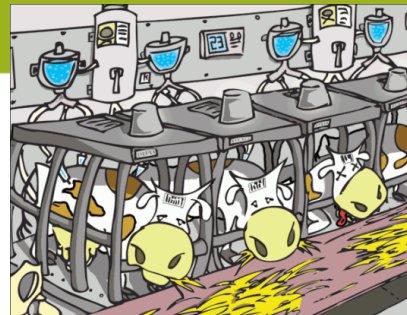
C'est en Europe que l'on constate la plus forte consommation de produits laitiers.

La France est le premier consommateur de beurre et de fromages au monde, soit environ 450 kg par an et par Français (en équivalent lait entier), ce qui la situe au dessus de la moyenne de l'UE (300kg).



Avec 3,586 millions de vaches laitières, 23 milliards de litres de lait sont collectés en France, transformés par 700 sites qui emploient 65 000 salariés.

Le secteur laitier occupe en France la première place dans le secteur de l'industrie alimentaire (25,6 milliards d'€, CNEIL 2008) avec actuellement environ 90 000 exploitants (ce nombre était de 370 000 au moment de la mise en œuvre des quotas laitiers en 1984¹).



2. Présentation de la série

Le sujet de notre étude est **la production laitière bovine en France** traité au travers de la série « **Collecte de lait de vache et produits obtenus** » exprimé en milliers de tonnes. Notre série est chronologique, les données y sont mensuelles et s'étalent du /10/1989 au /07/2010 en France. Les données proviennent de la base statistique Eurostat de l'Union Européenne.

Décomposition d'une chronique et Interrogation sur le profil temporel de notre série

La structure d'une chronique peut être décomposée, de façon standard, en 3 éléments :

- *La Tendance (Trend) notée T_t , correspondant au mouvement de long terme.*
La composante cyclique notée C_t .
Ces deux composantes sont regroupées en une seule appelée l'extra-saisonnier notée E_t

Quelle tendance pour la production laitière ?

Comment a varié au cours de ces deux dernières décennies la collecte de lait en France suite au progrès technique, la mécanisation, la concurrence mondiale et la politique de contrôle du secteur (PAC) ?

- *La composante saisonnière notée S_t , mouvement de court terme (période intra-annuelle) relativement régulier.*

Une consommation de produits laitiers stable sur l'année lisse t'elle la collecte de lait ?

Existe t'il un lien entre les quantités de lait collectées et les saisons qui structurent le climat français ?

- *La composante résiduelle notée R_t , représentant tout ce que les autres composantes n'ont pu expliquer du phénomène observé.*

Nous débuterons par une analyse classique de la série : Analyse de la saisonnalité, de la Tendance et des différentes méthodes traditionnelles de prévision.

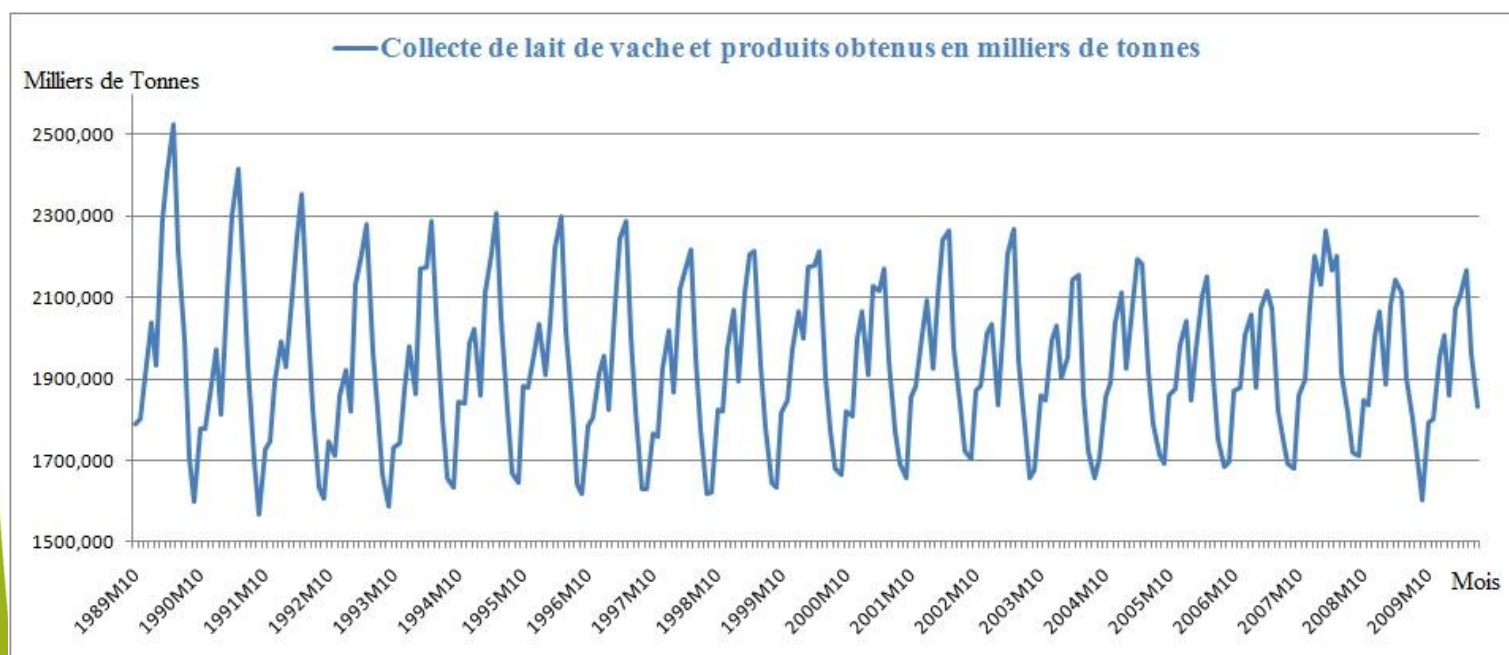
Par la suite nous réaliserons une analyse de notre chronique sous l'aspect Box et Jenkins : Méthodologie et Prévision par la technique de Box et Jenkins.

¹ La gestion administrée des Quotas laitiers mise en place a 1984 avait et a pour objectif d'équilibrer l'offre et la demande sur le marché du lait et des produits laitiers et de limiter le coût du secteur pour le Fond européen d'orientation et de garantie agricole (principal instrument de mise en œuvre de la PAC).

Ils consistent à maîtriser la production par l'attribution d'un quota à chaque exploitation.

Tableau N°1 : Collecte de lait de vache et produits obtenus en milliers de tonnes

| Années / Mois | Janvier | Février | Mars | Avril | Mai | Juin | Juillet | Août | Septembre | Octobre | Novembre | Décembre |
|---------------|----------|----------|----------|----------|----------|----------|----------|----------|-----------|----------|----------|----------|
| 1989 | | | | | | | | | | 1788,800 | 1801,300 | 1938,900 |
| 1990 | 2039,100 | 1934,900 | 2287,300 | 2409,900 | 2523,500 | 2211,800 | 1995,100 | 1706,300 | 1597,500 | 1777,000 | 1776,200 | 1887,600 |
| 1991 | 1973,100 | 1810,500 | 2117,700 | 2304,800 | 2417,700 | 2185,500 | 1936,200 | 1698,400 | 1568,900 | 1727,400 | 1746,200 | 1893,000 |
| 1992 | 1991,700 | 1929,900 | 2105,000 | 2247,500 | 2352,300 | 2040,100 | 1823,200 | 1632,100 | 1607,400 | 1746,400 | 1709,800 | 1860,400 |
| 1993 | 1921,500 | 1820,700 | 2133,300 | 2208,300 | 2281,200 | 1968,900 | 1824,000 | 1664,100 | 1586,300 | 1728,700 | 1741,500 | 1876,400 |
| 1994 | 1978,800 | 1863,200 | 2171,000 | 2172,700 | 2289,400 | 1988,400 | 1798,900 | 1656,600 | 1634,400 | 1844,300 | 1837,900 | 1988,000 |
| 1995 | 2022,441 | 1859,443 | 2113,948 | 2210,044 | 2307,724 | 2048,576 | 1825,376 | 1666,584 | 1644,228 | 1882,081 | 1879,052 | 1965,082 |
| 1996 | 2034,747 | 1908,828 | 2036,641 | 2221,766 | 2298,645 | 2016,762 | 1830,029 | 1639,531 | 1619,201 | 1784,000 | 1806,105 | 1912,650 |
| 1997 | 1957,186 | 1825,440 | 2077,130 | 2244,083 | 2288,323 | 2006,509 | 1820,171 | 1627,768 | 1627,671 | 1764,692 | 1758,042 | 1924,562 |
| 1998 | 2017,900 | 1866,500 | 2121,300 | 2162,000 | 2218,200 | 1952,800 | 1775,500 | 1616,700 | 1620,900 | 1823,300 | 1818,800 | 1975,400 |
| 1999 | 2069,749 | 1892,792 | 2111,261 | 2207,435 | 2214,606 | 1932,375 | 1780,362 | 1644,472 | 1633,599 | 1816,861 | 1846,884 | 1970,483 |
| 2000 | 2066,529 | 1999,666 | 2175,302 | 2178,463 | 2213,143 | 1907,740 | 1771,991 | 1679,093 | 1662,538 | 1820,109 | 1808,929 | 1998,627 |
| 2001 | 2066,495 | 1908,667 | 2126,341 | 2114,293 | 2171,776 | 1944,295 | 1771,106 | 1692,948 | 1655,822 | 1855,698 | 1883,089 | 2005,685 |
| 2002 | 2091,220 | 1925,779 | 2102,715 | 2242,688 | 2265,277 | 1977,018 | 1839,149 | 1723,371 | 1705,256 | 1870,539 | 1880,973 | 2010,179 |
| 2003 | 2033,473 | 1836,328 | 2011,375 | 2207,206 | 2269,669 | 1944,106 | 1782,180 | 1655,464 | 1677,832 | 1858,180 | 1846,701 | 1995,181 |
| 2004 | 2029,862 | 1903,918 | 1951,837 | 2142,731 | 2156,081 | 1858,636 | 1722,131 | 1655,683 | 1702,909 | 1853,537 | 1891,995 | 2038,741 |
| 2005 | 2112,720 | 1927,000 | 2082,140 | 2193,180 | 2182,840 | 1916,540 | 1787,320 | 1714,120 | 1691,600 | 1857,780 | 1874,910 | 1984,560 |
| 2006 | 2041,104 | 1848,520 | 1973,366 | 2106,220 | 2152,568 | 1901,494 | 1752,022 | 1684,041 | 1695,314 | 1869,618 | 1878,980 | 2005,291 |
| 2007 | 2059,440 | 1880,081 | 2072,471 | 2117,233 | 2072,051 | 1819,845 | 1755,570 | 1691,089 | 1678,220 | 1858,550 | 1898,068 | 2079,076 |
| 2008 | 2201,501 | 2131,541 | 2266,042 | 2168,414 | 2200,314 | 1914,645 | 1817,701 | 1720,166 | 1710,159 | 1847,078 | 1836,323 | 2001,510 |
| 2009 | 2067,105 | 1885,962 | 2081,064 | 2142,176 | 2113,554 | 1902,408 | 1799,622 | 1704,503 | 1602,158 | 1793,324 | 1798,900 | 1953,643 |
| 2010 | 2006,237 | 1859,018 | 2071,846 | 2104,995 | 2168,151 | 1959,215 | 1833,153 | | | | | |

Graphique N°1 : Représentation de l'évolution temporelle de la série
Polygone des fréquences - Profil temporel de la chronique

Les données s'étalent du mois d'octobre de l'année 1989 au mois de décembre de l'année 2009.

Pour la suite de notre dossier nous travaillerons sur la période de 1990 à 2008 afin que la chronique ne contienne pas d'observations manquantes et afin de garder l'année 2009 comme horizon de la prévision dans le but de confronter nos prévisions à des valeurs observées.

L'observation de la représentation graphique de la série met en évidence que la collecte de lait a suivi ces deux dernières décennies une tendance linéaire que l'on peut visuellement supposer faiblement décroissante, voir plate (tendance stable plus visible avec un axe des abscisses plus étalé).

Des pics et des creux de période qui semble redondante et annuelle sont aussi mis en évidence. La série semble donc saisonnière et la forme de la courbe nous laisse à penser à une certaine rigidité de la saisonnalité.

Lorsqu'on étale le graphique en élargissant l'axe des abscisses, on peut identifier que la période autour du mois de Mai correspond aux pics et que celle autour du mois de septembre correspond aux creux.



3. L'Analyse de la saisonnalité

6

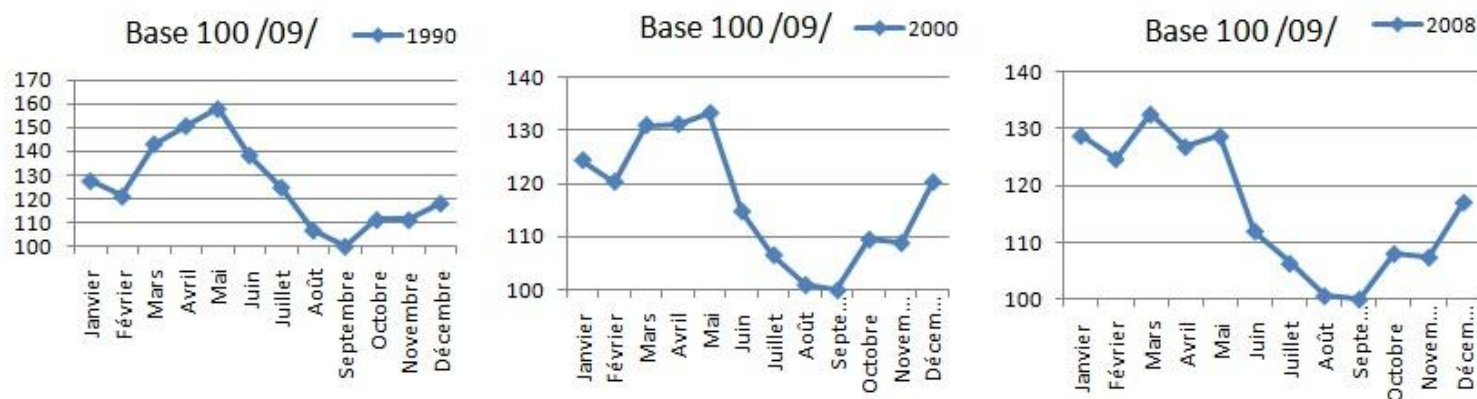
Identifier son existence et sa nature

3.1. Série en Indice

Nous allons observer le profil temporel des années 1990, 2000 et 2008 afin de représenter la saisonnalité de la production laitière pour chacune de ces années qui découpent et structurent (début, milieu, fin) notre série.

Ci-dessous les polygones des fréquences de ces 3 années dont nous avons préalablement transformé les données en indices avec comme base 100 le mois de l'année qui présente la plus faible valeur de collecte (Graphiques N°2).

Graphiques N°2 : Polygones des fréquences



Les graphiques ci-dessus confirment que les valeurs les plus fortes sont durant la période d'Avril-Mai et les plus basses durant la période d'Août-Septembre.

La représentation par année et en indice nous permet de mieux apprécier l'évolution des collectes et de la forme prise par leur saisonnalité.

On peut pressentir comme à la lecture du graphique de l'ensemble de la série (Graphique N°1) une certaine rigidité des variations saisonnières, mais à nuancer, les 3 graphiques n'ayant pas exactement le même profil.



3.2. Tableau de Buys-Ballot

L'analyse graphique a permis de mettre en évidence une saisonnalité qui pourrait être qualifiée de rigide. Le tableau de Buys-Ballot (*Tableau 1*) va permettre d'analyser plus précisément notre série.

Tableau N°2 : Buys-Ballot

| Années / Mois | Janvier | Février | Mars | Avril | Mai | Juin | Juillet | Août | Septembre | Octobre | Novembre | Décembre | Moyenne | Ecart type |
|---------------|------------|------------|----------|------------|-----------------|------------|------------|------------|-----------------|-----------|------------|------------|------------------|--------------------|
| 1990 | 2039,100 | 1934,900 | 2287,300 | 2409,900 | 2523,500 | 2211,800 | 1995,100 | 1706,300 | 1597,500 | 1777,000 | 1776,200 | 1887,600 | 2012,183 | 279,015465 |
| 1991 | 1973,100 | 1810,500 | 2117,700 | 2304,800 | 2417,700 | 2185,500 | 1936,200 | 1698,400 | 1568,900 | 1727,400 | 1746,200 | 1893,000 | 1948,283 | 250,2053784 |
| 1992 | 1991,700 | 1929,900 | 2105,000 | 2247,500 | 2352,300 | 2040,100 | 1823,200 | 1632,100 | 1607,400 | 1746,400 | 1709,800 | 1860,400 | 1920,483 | 226,5452892 |
| 1993 | 1921,500 | 1820,700 | 2133,300 | 2208,300 | 2281,200 | 1968,900 | 1824,000 | 1664,100 | 1586,300 | 1728,700 | 1741,500 | 1876,400 | 1896,242 | 208,3616786 |
| 1994 | 1978,800 | 1863,200 | 2171,000 | 2172,700 | 2289,400 | 1988,400 | 1798,900 | 1656,600 | 1634,400 | 1844,300 | 1837,900 | 1988,000 | 1935,300 | 194,631027 |
| 1995 | 2022,441 | 1859,443 | 2113,948 | 2210,044 | 2307,724 | 2048,576 | 1825,376 | 1666,584 | 1644,228 | 1882,081 | 1879,052 | 1965,082 | 1952,048 | 192,5291761 |
| 1996 | 2034,747 | 1908,828 | 2036,641 | 2221,766 | 2298,645 | 2016,762 | 1830,029 | 1639,531 | 1619,201 | 1784,000 | 1806,105 | 1912,650 | 1925,742 | 199,8037713 |
| 1997 | 1957,186 | 1825,440 | 2077,130 | 2244,083 | 2288,323 | 2006,509 | 1820,171 | 1627,768 | 1627,671 | 1764,692 | 1758,042 | 1924,562 | 1910,131 | 206,8039305 |
| 1998 | 2017,900 | 1866,500 | 2121,300 | 2162,000 | 2218,200 | 1952,800 | 1775,500 | 1616,700 | 1620,900 | 1823,300 | 1818,800 | 1975,400 | 1914,108 | 188,4383916 |
| 1999 | 2069,749 | 1892,792 | 2111,261 | 2207,435 | 2214,606 | 1932,375 | 1780,362 | 1644,472 | 1633,599 | 1816,861 | 1846,884 | 1970,483 | 1926,740 | 188,1350411 |
| 2000 | 2066,529 | 1999,666 | 2175,302 | 2178,463 | 2213,143 | 1907,740 | 1771,991 | 1679,093 | 1662,538 | 1820,109 | 1808,929 | 1998,627 | 1940,178 | 186,1280727 |
| 2001 | 2066,495 | 1908,667 | 2126,341 | 2114,293 | 2171,776 | 1944,295 | 1771,106 | 1692,948 | 1655,822 | 1855,698 | 1883,089 | 2005,685 | 1933,018 | 163,3540494 |
| 2002 | 2091,220 | 1925,779 | 2102,715 | 2242,688 | 2265,277 | 1977,018 | 1839,149 | 1723,371 | 1705,256 | 1870,539 | 1880,973 | 2010,179 | 1969,514 | 174,0034974 |
| 2003 | 2033,473 | 1836,328 | 2011,375 | 2207,206 | 2269,669 | 1944,106 | 1782,180 | 1655,464 | 1677,832 | 1858,180 | 1846,701 | 1995,181 | 1926,475 | 181,5502444 |
| 2004 | 2029,862 | 1903,918 | 1951,837 | 2142,731 | 2156,081 | 1858,636 | 1722,131 | 1655,683 | 1702,909 | 1853,537 | 1891,995 | 2038,741 | 1909,005 | 157,174335 |
| 2005 | 2112,720 | 1927,000 | 2082,140 | 2193,180 | 2182,840 | 1916,540 | 1787,320 | 1714,120 | 1691,600 | 1857,780 | 1874,910 | 1984,560 | 1943,726 | 163,9812216 |
| 2006 | 2041,104 | 1848,520 | 1973,366 | 2106,220 | 2152,568 | 1901,494 | 1752,022 | 1684,041 | 1695,314 | 1869,618 | 1878,980 | 2005,291 | 1909,045 | 146,1061668 |
| 2007 | 2059,440 | 1880,081 | 2072,471 | 2117,233 | 2072,051 | 1819,845 | 1755,570 | 1691,089 | 1678,220 | 1858,550 | 1898,068 | 2079,076 | 1915,141 | 153,7069349 |
| 2008 | 2201,501 | 2131,541 | 2266,042 | 2168,414 | 2200,314 | 1914,645 | 1817,701 | 1720,166 | 1710,159 | 1847,078 | 1836,323 | 2001,510 | 1984,616 | 193,0768224 |
| 2009 | 2067,105 | 1885,962 | 2081,064 | 2142,176 | 2113,554 | 1902,408 | 1799,622 | 1704,503 | 1602,158 | 1793,324 | 1798,900 | 1953,643 | 1903,702 | 164,9455591 |
| Moyenne | 2038,784 | 1897,983 | 2105,862 | 2200,057 | 2249,444 | 1971,922 | 1810,382 | 1673,652 | 1646,095 | 1818,957 | 1825,968 | 1966,304 | Moyenne Générale | Ecart type Général |
| Ecart type | 59,0737329 | 70,1664356 | 79,81683 | 68,9655626 | 102,229019 | 94,0430175 | 60,3308175 | 31,5389141 | 41,8611873 | 48,065644 | 55,0178473 | 55,9606781 | 1933,784 | 195,6044831 |

Le tableau de Buys-Ballot nous indique que la moyenne des valeurs correspondant au mois de Mai est la plus élevée sur la période de 1990 à 2009 en comparaison avec les autres mois.

En moyenne la collecte de lait de vache atteint donc son maximum au mois de Mai.

A l'inverse c'est au mois de Septembre que la moyenne est la plus faible.

Le couple Avril-Mai apparait comme la période aux moyennes les plus élevées, et le couple Août-Septembre comme celle aux moyennes les plus faibles comme nous le montraient les graphiques des valeurs indicés.

Ces résultats confirment ce qui avait été observé sur la représentation temporelle de la série (Graphique N°1).



3.3. Tableau de Buys-Ballot classé

Classement des mois en fonction de leurs valeurs

A chaque mois est attribuée une couleur :

| | | | | | | | | | | | |
|---------|---------|------|-------|-----|------|---------|------|-----------|---------|----------|----------|
| Janvier | Février | Mars | Avril | Mai | Juin | Juillet | Août | Septembre | Octobre | Novembre | Décembre |
|---------|---------|------|-------|-----|------|---------|------|-----------|---------|----------|----------|

Les valeurs du tableau sont ensuite classées par ordre décroissant.

Les couleurs d'arrière-plan des cellules suivent les valeurs et nous permettent de rapidement voir la position de chaque mois ; mais aussi d'avoir une vision d'ensemble sur les mouvements durant les 20 années.

Tableau N°3 : Buys-Ballot classé

| | | | | | | | | | | | | |
|------|---------|----------|---------|----------|---------|--------|---------|----------|----------|----------|----------|----------|
| 1990 | 2523,5 | 2409,9 | 2287,3 | 2211,8 | 2039,1 | 1995,1 | 1934,9 | 1887,6 | 1777 | 1776,2 | 1706,3 | 1597,5 |
| 1991 | 2417,7 | 2304,8 | 2185,5 | 2117,7 | 1973,1 | 1936,2 | 1893 | 1810,5 | 1746,2 | 1727,4 | 1698,4 | 1568,9 |
| 1992 | 2352,3 | 2247,5 | 2105 | 2040,1 | 1991,7 | 1929,9 | 1860,4 | 1823,2 | 1746,4 | 1709,8 | 1632,1 | 1607,4 |
| 1993 | 2281,2 | 2208,3 | 2133,3 | 1968,9 | 1921,5 | 1876,4 | 1824 | 1820,7 | 1741,5 | 1728,7 | 1664,1 | 1586,3 |
| 1994 | 2289,4 | 2172,7 | 2171 | 1988,4 | 1988 | 1978,8 | 1863,2 | 1844,3 | 1837,9 | 1798,9 | 1656,6 | 1634,4 |
| 1995 | 2307,72 | 2210,044 | 2113,95 | 2048,576 | 2022,44 | 1965,1 | 1882,08 | 1879,052 | 1859,443 | 1825,376 | 1666,584 | 1644,228 |
| 1996 | 2298,65 | 2221,766 | 2036,64 | 2034,747 | 2016,76 | 1912,7 | 1908,83 | 1830,029 | 1806,105 | 1784 | 1639,531 | 1619,201 |
| 1997 | 2288,32 | 2244,083 | 2077,13 | 2006,509 | 1957,19 | 1924,6 | 1825,44 | 1820,171 | 1764,692 | 1758,042 | 1627,768 | 1627,671 |
| 1998 | 2218,2 | 2162 | 2121,3 | 2017,9 | 1975,4 | 1952,8 | 1866,5 | 1823,3 | 1818,8 | 1775,5 | 1620,9 | 1616,7 |
| 1999 | 2214,61 | 2207,435 | 2111,26 | 2069,749 | 1970,48 | 1932,4 | 1892,79 | 1846,884 | 1816,861 | 1780,362 | 1644,472 | 1633,599 |
| 2000 | 2213,14 | 2178,463 | 2175,3 | 2066,529 | 1999,67 | 1998,6 | 1907,74 | 1820,109 | 1808,929 | 1771,991 | 1679,093 | 1662,538 |
| 2001 | 2171,78 | 2126,341 | 2114,29 | 2066,495 | 2005,69 | 1944,3 | 1908,67 | 1883,089 | 1855,698 | 1771,106 | 1692,948 | 1655,822 |
| 2002 | 2265,28 | 2242,688 | 2102,72 | 2091,22 | 2010,18 | 1977 | 1925,78 | 1880,973 | 1870,539 | 1839,149 | 1723,371 | 1705,256 |
| 2003 | 2269,67 | 2207,206 | 2033,47 | 2011,375 | 1995,18 | 1944,1 | 1858,18 | 1846,701 | 1836,328 | 1782,18 | 1677,832 | 1655,464 |
| 2004 | 2156,08 | 2142,731 | 2038,74 | 2029,862 | 1951,84 | 1903,9 | 1892 | 1858,636 | 1853,537 | 1722,131 | 1702,909 | 1655,683 |
| 2005 | 2193,18 | 2182,84 | 2112,72 | 2082,14 | 1984,56 | 1927 | 1916,54 | 1874,91 | 1857,78 | 1787,32 | 1714,12 | 1691,6 |
| 2006 | 2152,57 | 2106,22 | 2041,1 | 2005,291 | 1973,37 | 1901,5 | 1878,98 | 1869,618 | 1848,52 | 1752,022 | 1695,314 | 1684,041 |
| 2007 | 2117,23 | 2079,076 | 2072,47 | 2072,051 | 2059,44 | 1898,1 | 1880,08 | 1858,55 | 1819,845 | 1755,57 | 1691,089 | 1678,22 |
| 2008 | 2266,04 | 2201,501 | 2200,31 | 2168,414 | 2131,54 | 2001,5 | 1914,65 | 1847,078 | 1836,323 | 1817,701 | 1720,166 | 1710,159 |
| 2009 | 2142,18 | 2113,554 | 2081,06 | 2067,105 | 1953,64 | 1902,4 | 1885,96 | 1799,622 | 1798,9 | 1793,324 | 1704,503 | 1602,158 |

En ce qui concerne les trois mois les plus forts, par ordre décroissant : Mai, Avril, Mars et les trois mois les plus faibles, par ordre croissant : Septembre, Août, Juillet ; la redondance est évidente.

On remarque que le mois de Mai revient en premier position de 1990 à 2004, et que le mois de septembre occupe lui la dernière place dans la majorité des années.

A contrario on observe de nombreux mouvements à l'intérieur du tableau pour les 6 colonnes du centre.

La question sur la nature de la saisonnalité vient à se poser : Rigide, Déterministe ou Souple, Aléatoire ?

Au vu des conclusions précédentes et du tableau ci-dessus nous pouvons émettre l'hypothèse d'être en présence d'une saisonnalité rigide. Nous testerons cette hypothèse en effectuant deux méthodes de dessaisonnalisation :

Une avec l'emploi de coefficients saisonniers identiques sur toute la période, adaptée à une saisonnalité rigide

Une avec l'emploi des techniques de filtrage par moyennes mobiles (Census), adaptée à une saisonnalité souple.

Si les coefficients saisonniers obtenus avec les deux méthodes (moyenne pour Census) se superposent, l'hypothèse d'une saisonnalité rigide sera acceptée sinon la saisonnalité sera qualifiée de souple.



3.4. Test de Fisher - Composante Saisonnière ? Tendancielle ?

(A partir de cette page pour les raisons stipulées précédemment la série traitée ne contient que la période de 1990 à 2008)

Analyse de la variance : Détection d'une saisonnalité et/ou d'une tendance.

Ce test suppose que la chronique est sans tendance ou encore sans extra-saisonnalité. Dans le cas contraire on éliminera cette composante par la procédure du filtrage aux différences premières.

Nous allons décomposer la somme totale des carrés de la série en trois sommes de carrés correspondant respectivement à :

Aux colonnes notée S_p pour périodes ici le mois, aux lignes notée S_a pour les années et enfin aux résidus notée S_r .

| Somme des carrés | Degrés de liberté | Désignation | Variance |
|---|---------------------|------------------|-----------------------------|
| $S_p = N \sum_j (x_{.j} - x_{..})^2$ | p-1 | Variance période | $V_p = S_p / (p-1)$ |
| $S_a = p \sum_i (x_{i.} - x_{..})^2$ | N-1 | Variance année | $V_a = S_a / (N-1)$ |
| $S_r = \sum_i \sum_j (x_{ij} - x_{i.} - x_{.j} + x_{..})^2$ | (p-1)*(N-1) | Variance résidu | $V_r = S_r / [(p-1)*(N-1)]$ |
| $S_T = S_a + S_p + S_r$ | N*p-1 | Variance totale | $V_t = S_T / N*(p-1)$ |

N : Le nombre d'années = 19

p : La périodicité = 12

Deux effets seront testés :

- **L'effet période** soit l'effet mois, s'il est significatif la série est saisonnière ;
- **L'effet année**, s'il est significatif la série contient une tendance ou a été mal transformée.



Tableau N°4 : Analyse de la variance

| Désignation | Somme des carrés (S) | Degré de liberté | Variance (V) |
|---------------|----------------------|------------------|--------------|
| Période (p) | 7805421,844 | 11 | 709583,804 |
| Année (a) | 179039,6004 | 18 | 9946,644468 |
| Résidu (r) | 860290,4528 | 198 | 4344,901277 |
| Total (t) | 8844751,897 | 227 | 38963,66475 |

A partir de ce tableau nous allons construire les tests d'hypothèse :

Test de l'influence du facteur colonne, la période : ici le mois

On va tester l'hypothèse H_0 : Pas d'influence du facteur colonne.

Règles de décision :

Si le Fisher empirique $F_c = V_p / V_r >$ au Fisher lu dans la table (F_t), alors on rejette l'hypothèse H_0 , la série est saisonnière.

Si le Fisher empirique $F_c <$ au Fisher lu dans la table (F_t), alors on accepte l'hypothèse H_0 , pas d'influence du facteur colonne, pas d'effet mois, la chronique est non saisonnière.

Test au seuil de 5% :

La valeur du Fisher théorique (à $v_1=11$ et $v_2=198$ degrés de liberté) est de 1,84 ($F_t=1,84$).

La valeur du Fisher calculé est de 163,3141374 ($F_c \approx 163,314$).

$F_c = 163,314 > F_t = 1,84$; Le Fisher calculé est plus grand que le Fisher théorique.

On rejette l'hypothèse H_0 , la série est donc saisonnière.

Test de l'influence du facteur ligne, la tendance

On va tester l'hypothèse H_0 : Pas d'influence du facteur année.

Règles de décision :

Si Fisher empirique $F_c = V_a / V_r >$ valeur du Fisher théorique (F_t), alors on rejette l'hypothèse H_0 , la série est donc affectée d'une tendance.

Si Fisher empirique $F_c <$ valeur du Fisher théorique (F_t), alors l'hypothèse H_0 est acceptée, pas d'influence du facteur année, la série n'est pas affectée d'une tendance.

Test au seuil de 5% :

La valeur du Fisher théorique (à $v_3=18$ et $v_2=198$ degrés de liberté) est de 1,66 ($F_t=1,66$)

La valeur du Fisher calculé est de 2,289268233 ($F_c \approx 2,289$)

$F_c \approx 2,289 > F_t = 1,66$; Le Fisher calculé est plus grand que le Fisher théorique.

On rejette l'hypothèse H_0 , la série est donc affectée d'une tendance (à la limite de la significativité).

Comme stipulé plus haut pour que le test de l'analyse de la variance soit significatif il faut que la série soit sans tendance ou encore sans extra-saisonnalité, or d'après le test précédent notre série est affectée d'une tendance.

Ainsi nous allons réitérer le test sur la chronique transformée par les différences premières ce qui doit normalement éliminer la présence de l'extra-saisonnier.

NB :

Avant les années 1980 il n'existait pas de tests permettant de détecter la non-stationnarité, en conséquence les transformations utilisées étaient les filtres aux différences et la formule de Box Cox.

Ces transformations peuvent être mal adaptées aux caractéristiques de la non-stationnarité, ce qui a pour effet d'introduire au sein de notre série des mouvements parasites.

Depuis, les cas de non-stationnarité sont analysés à partir de deux types de processus :

- Processus TS : Non stationnarité de type déterministe ;
- Processus DS : Non stationnarité de type aléatoire.

A chaque processus correspond une bonne méthode de « stationnarisation ».

Pour « stationnariser » un processus TS on utilisera celle des moindres carrés ordinaires, pour un processus DS on emploiera le filtre aux différences.

Afin de savoir si notre série correspond à un processus TS ou DS nous utiliserons par la suite les tests de Dickey et Fuller augmentés, ces tests s'effectuent sur la série Corrigée des Variations Saisonnières (notée CVS) or nous n'avons pas encore désaisonné notre série, c'est pourquoi nous utilisons la chronique $x_t - x_{t-1}$ pour remettre en œuvre le test de Fisher.

Nous allons donc tenter d'éliminer la tendance par un passage aux différences première et reconstruire le tableau de Buys-Ballot de la série supposée sans extra-saisonnier afin d'observer une éventuelle saisonnalité.

La mise en œuvre du test de Fisher sur la série transformée infirmera ou confirmera la présence d'une saisonnalité et devrait confirmer que la série est désormais dénuée de tendance.

Tableau N°5 : Analyse de la variance pour la chronique $x_t - x_{t-1}$

| Désignation | Somme carré (S) | Degré de liberté | Variance (V) |
|---------------|-----------------|------------------|--------------|
| Période (p) | 4849179,476 | 11 | 440834,498 |
| Année (a) | 3338,61277 | 18 | 185,478487 |
| Résidu (r) | 494818,0613 | 198 | 2499,08112 |
| Total (t) | 5347336,15 | 227 | 23556,5469 |

A partir de ce tableau nous allons reconstruire les tests d'hypothèse :

Test de l'influence du facteur colonne, la période : ici le mois

Nous allons tester l'hypothèse H_0 : Pas d'influence du facteur colonne.

Règles de décision :

Si le Fisher empirique $F_c = V_p/V_r >$ au Fisher lu dans la table (F_t), alors on rejette l'hypothèse H_0 , la série est saisonnière.

Si le Fisher empirique $F_c <$ au Fisher lu dans la table (F_t), alors on accepte l'hypothèse H_0 , pas d'influence du facteur colonne, pas d'effet mois, la chronique est non saisonnière.

Test au seuil de 5% :

La valeur du Fisher théorique (à $v_1=11$ et $v_2=198$ degrés de liberté) est de 1,84 ($F_t=1,84$).

La valeur du Fisher calculé est de 176,398635 ($F_c \approx 176,399$).

$F_c \approx 176,399 > F_t = 1,84$; Le Fisher calculé est plus grand que le Fisher théorique.

On rejette l'hypothèse H_0 , la série est donc saisonnière.

Test de l'influence du facteur ligne, la tendance

On va tester l'hypothèse H_0 : Pas d'influence du facteur année.

Règles de décision :

Si Fisher empirique $F_c = V_a/V_r >$ valeur du Fisher théorique (F_t), alors on rejette l'hypothèse H_0 , la série est donc affectée d'une tendance.

Si Fisher empirique $F_c <$ valeur du Fisher théorique (F_t), alors l'hypothèse H_0 est acceptée, pas d'influence du facteur année, la série n'est pas affectée d'une tendance.

Test au seuil de 5% :

La valeur du Fisher théorique (à $v_3=18$ et $v_2=198$ degrés de liberté) est de 1,66 ($F_t=1,66$)

La valeur du Fisher calculé est de 0,074218674 ($F_c \approx 0,074$)

$F_c \approx 0,074 < F_t = 1,66$; Le Fisher calculé est plus petit que le Fisher théorique.

On accepte l'hypothèse H_0 , la série n'est donc pas affectée d'une tendance.

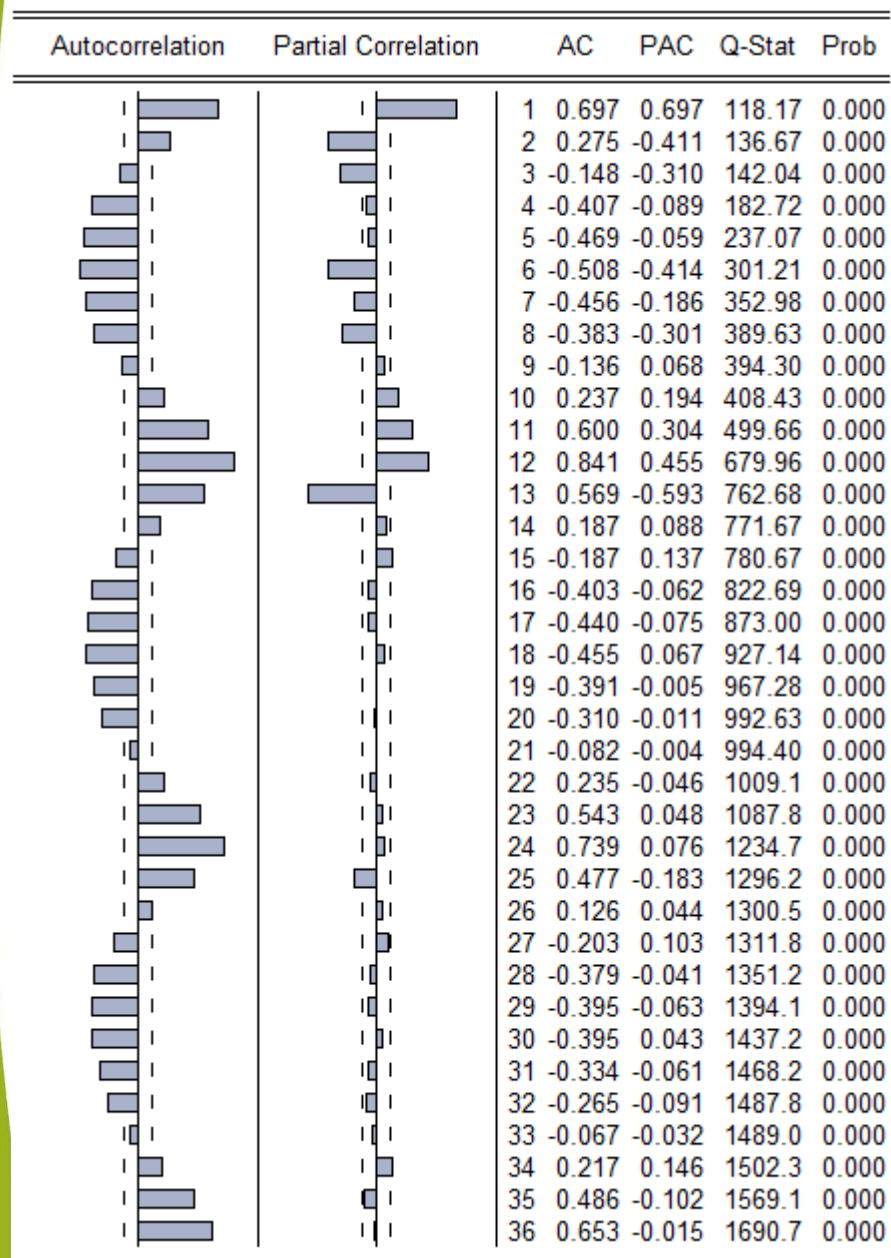
Vu qu'il s'agit de la série filtrée, le test ne fait que confirmer que la transformation a éliminé la tendance.

Ce deuxième test confirme donc les résultats précédents à savoir que notre série est saisonnière.

3.5. La Fonction d'autocorrélation

Etude de la fonction d'autocorrélation (notée FAC) par lecture de sa représentation graphique : Le corrélogramme.

Graphique N° 3 : Corrélogramme de la Série brute

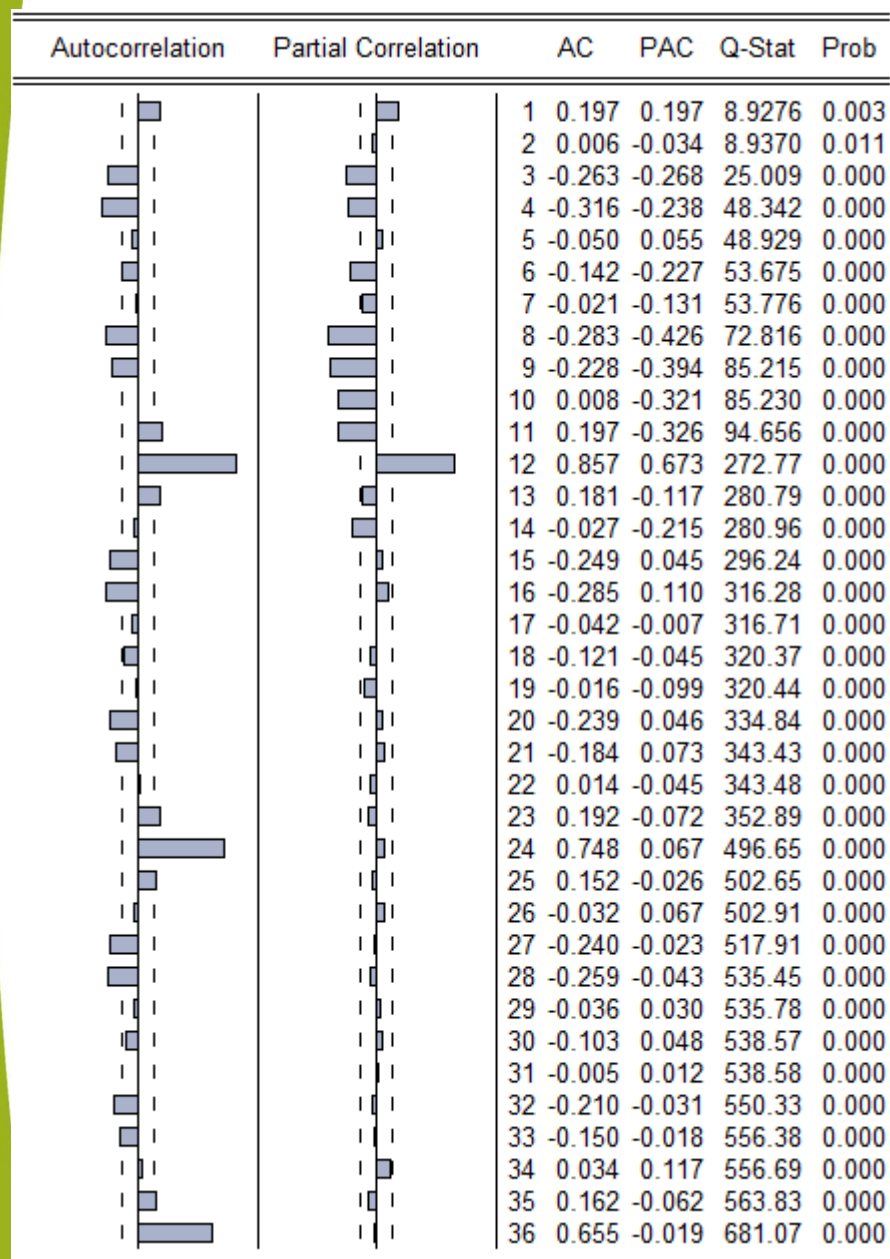


Le corrélogramme met en évidence l'existence d'un élément cyclique au sein de la série.

On remarque que les pics se répètent tous les 12 retards.

Nos pics semblent être constant au cours des retards ce qui peut amener à penser à une tendance assez stable et à l'existence d'un cycle au sein de notre série.

Graphique n°4 : Corrélogramme de la série transformée par les différences premières



Décomposition temporelle de la chronique :

Le corrélogramme contient les caractéristiques d'une variable aléatoire ainsi que d'une saisonnalité.

On remarque que la transformation a atténué le cycle des variations au sein de la série.

La modélisation de la série selon la technique de décomposition – reconstruction repose sur un modèle appelé « Schéma de décomposition » et suppose le bon choix du type de schéma de décomposition.

En effet les 3 grandes composantes d'une chronique que sont l'extra-saisonnier, la saisonnalité et le résidu peuvent s'influencer. La saisonnalité d'une chronique peut être parfois influencée par l'extra-saisonnier et/ou le résidu.

Afin de modéliser ces interactions il existe trois grands types de schéma :

- Le schéma additif : $X_t = Et + St + Rt$ où les composantes sont indépendantes. Dans ce schéma la saisonnalité est rigide en amplitude et en période.
- Le schéma multiplicatif : $X_t = St * Et + Rt$ où la composante saisonnière est liée à l'extra-saisonnier. Dans ce schéma la saisonnalité est souple avec variation de l'amplitude au cours du temps.
- Le schéma multiplicatif complet : $X_t = St * Et * Rt$ où les 3 composantes interagissent.

Nous allons mettre en œuvre deux techniques pour sélectionner le type de schéma à retenir :
La procédure de la Bande et Le test de Buys-Ballot.

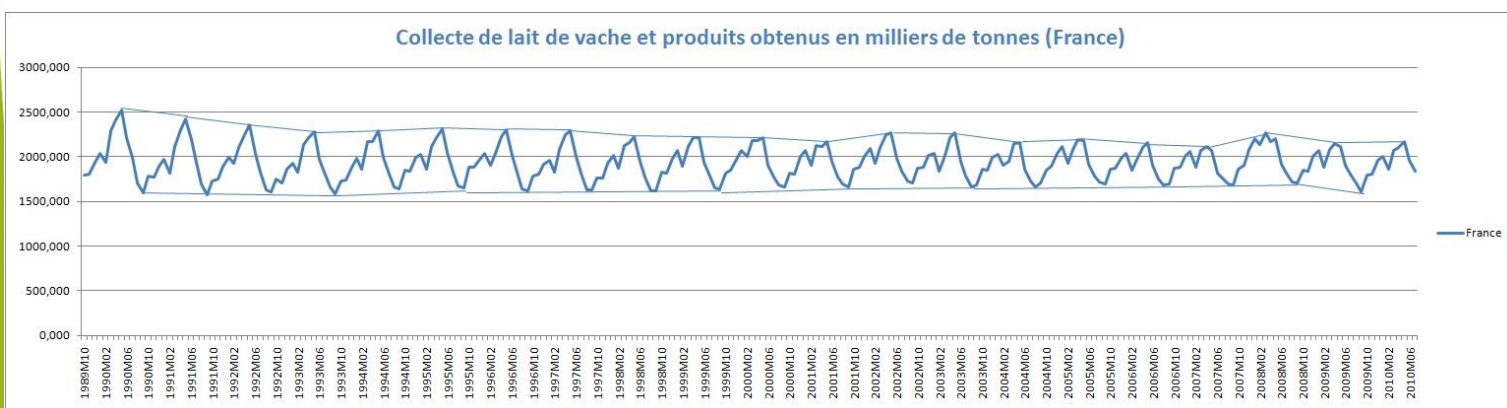
4.1. La procédure de la bande

Elle consiste à relier par une ligne brisée toutes les valeurs hautes et toutes les valeurs basses dans le graphique représentant le profil temporel de la chronique.

Règles de décision :

- Si les deux lignes sont parallèles on opte pour un schéma additif.
- Dans le cas contraire le schéma multiplicatif est plus adapté.

Graphique N°5 : Evolution temporelle de la série & procédure de la bande



L'examen visuel (délicat) du graphique ci-dessus :

- Un tracé grossier amènera à conclure sur un modèle additif (les deux lignes inférieures et supérieures parallèles)
- Un tracé plus précis met en évidence un parallélisme plus qu'imparfait et amène à conclure à un modèle multiplicatif.

Nb : On remarque aussi que les années de 1990 à 1993 comportent des variations d'amplitudes plus importantes, sûrement du au ajustement de l'offre suite a la mise en œuvre des quotas laitiers en 1984 dont le choc se résorbe.



4.2. Le test de Buys-Ballot

Il se base sur le tableau de Buys-Ballot (Tableau 1) et consiste à tester si l'écart type et la moyenne sont indépendants.

Règles de décision :

Le schéma est, par définition, additif si l'écart type et la moyenne sont indépendants.

Le schéma est multiplicatif si l'écart type et la moyenne sont dépendants.

On estime par la méthode des MCO les paramètres a_1 et a_0 de l'équation : $\text{Ecart-type} = a_1 * \text{moyenne} + a_0 + \epsilon_i$

Tableau N°6 : MCO /90/ - /08/

| Dependent Variable: EI | | | | |
|----------------------------|-------------|-----------------------|-------------|--------|
| Method: Least Squares | | | | |
| Date: 11/08/10 Time: 15:04 | | | | |
| Sample: 1990 2008 | | | | |
| Included observations: 19 | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| C | -869.7736 | 472.2780 | -1.841656 | 0.0830 |
| MOYENNE | 0.548767 | 0.243999 | 2.249051 | 0.0381 |
| R-squared | 0.229313 | Mean dependent var | 192.2921 | |
| Adjusted R-squared | 0.183978 | S.D. dependent var | 32.99300 | |
| S.E. of regression | 29.80387 | Akaike info criterion | 9.726454 | |
| Sum squared resid | 15100.60 | Schwarz criterion | 9.825869 | |
| Log likelihood | -90.40132 | F-statistic | 5.058228 | |
| Durbin-Watson stat | 0.303164 | Prob(F-statistic) | 0.038055 | |

Test de significativité du paramètre estimé « a_1 » au seuil de 5% : Test de Student

Hypothèse H_0 : le coefficient n'est pas significativement différent de 0.

Règles de décision :

Si la valeur du Student empirique de la moyenne (T_c) : « t-Statistic » < valeur du Student théorique (T_t) alors on accepte H_0 .

Si la valeur du Student empirique de la moyenne : « t-Statistic » > valeur du Student théorique alors on rejette H_0 .

Si le coefficient n'est pas significativement différent de 0 alors on accepte l'hypothèse d'un schéma additif.

Si le coefficient est significativement différent de 0 alors on accepte l'hypothèse d'un schéma multiplicatif.

Pour un risque de première espèce α de 5% et à $v=n-2=17$ degrés de liberté,

la valeur du Student théorique (T_t) est de 2,110

La valeur du Student calculé (T_c) est de 2,249051

$T_c > T_t$ (à la limite de la significativité) on rejette H_0 , le coefficient est significativement différent de 0, alors on accepte l'hypothèse d'un schéma multiplicatif.



4.3. Le test de Buys-Ballot sur la période 93-08

17

En effet à la lecture du graphique nous avons remarqué que les années de 1990 à 1992 comportaient des variations d'amplitude étrangement plus fortes que les années suivantes, ce qui peut avoir faussé le test.

Règles de décision :

- Le schéma est, par définition, additif si l'écart type et la moyenne sont indépendants.
- Le schéma est multiplicatif si l'écart type et la moyenne sont dépendants.

On estime par la méthode des MCO les paramètres a_1 et a_0 de l'équation : $\text{Ecart-type} = a_1 * \text{moyenne} + a_0 + \varepsilon_i$

Tableau N°7 : MCO /93/ - /08/

| Dependent Variable: ET Method: Least Squares Date: 11/28/10 Time: 15:21 Sample: 1993 2008 Included observations: 16 | | | | |
|---|-------------|-----------------------|-------------|--------|
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| C | 76.00457 | 421.3544 | 0.180382 | 0.8594 |
| MOYENNE | 0.054440 | 0.218225 | 0.249468 | 0.8066 |
| R-squared | 0.004426 | Mean dependent var | 181.1115 | |
| Adjusted R-squared | -0.066687 | S.D. dependent var | 19.22745 | |
| S.E. of regression | 19.85822 | Akaike info criterion | 8.931581 | |
| Sum squared resid | 5520.883 | Schwarz criterion | 9.028155 | |
| Log likelihood | -69.45265 | F-statistic | 0.062234 | |
| Durbin-Watson stat | 0.647258 | Prob(F-statistic) | 0.806622 | |

Test de significativité du paramètre estimé « a_1 » au seuil de 5% : Test de Student

Hypothèse H_0 : le coefficient n'est pas significativement différent de 0,

Règles de décision :

Si la valeur du Student empirique de la moyenne (T_c) : « t-Statistic » < valeur du Student théorique (T_t) alors on accepte H_0 .

Si la valeur du Student empirique de la moyenne : « t-Statistic » > valeur du Student théorique alors on rejette H_0 .

Si le coefficient n'est pas significativement différent de 0 alors on accepte l'hypothèse d'un schéma additif.

Si le coefficient est significativement différent de 0 alors on accepte l'hypothèse d'un schéma multiplicatif.

Pour un risque de première espèce α de 5% et à $v=n-2=14$ degrés de liberté,
la valeur du Student théorique (T_t) est de 2,145

La valeur du Student calculé (T_c) est de 0,249468

$T_c < T_t$ alors on accepte H_0 , le coefficient n'est pas significativement différent de 0, alors on accepte l'hypothèse d'un schéma additif.

Les conclusions sont différentes pour le test de Buys-Ballot, en effet si l'on retire de notre étude les années 1990-91 et 92 le schéma retenu est un schéma additif.



5. Désaisonnalisation

Nous savons donc désormais que notre série chronologique est structurée par une saisonnalité de nature rigide.

L'étape suivante consiste à la désaisonnaliser, sans modifier les autres composantes de la chronique.

Le choix de la technique la mieux appropriée dépend de la nature déterministe ou aléatoire de la saisonnalité.

Ce qui semble logique, le choix d'une méthode qui donne des coefficients saisonniers identiques pour toutes les années ne peut être approprié qu'à une série dont la saisonnalité aurait la même allure sur l'ensemble de ces dites années.

A l'inverse une série dont la saisonnalité varie d'une année sur l'autre ne peut se voir modélisée et prévue avec des coefficients saisonniers fixes par mois, trimestre...

Pour la désaisonnalisation, nous allons scinder notre étude en deux parties : La première couvrant la période 1990-2008, la seconde allant de 1993 à 2008.

- Dans un premier temps, nous travaillerons sur la série de 1990 à 2008.

Sur cette durée la série présente une composante saisonnière de nature plutôt rigide et le schéma de décomposition sélectionné est le schéma multiplicatif.

Nous allons pour confirmer la rigidité, appliquer deux méthodes de désaisonnalisation :

- La méthode CensusX12 qui est appropriée à une saisonnalité souple ;
- La méthode Moyenne Mobile appropriée à une saisonnalité rigide.

- Dans un second temps, nous travaillerons sur la série de 1993 à 2008 :

La nature de la saisonnalité est comme précédemment plutôt rigide.

Par contre le schéma retenu est ici un schéma additif, en accord avec la saisonnalité rigide en amplitude et en période, nature de la saisonnalité sur laquelle nous avaient orienté le tableau de Buys-Ballot et l'analyse graphique.

La page suivante présente donc un tableau comparatif des deux méthodes

Tableau N°8 : Coefficients saisonniers « Moyenne Mobile » et « CensusX12 »

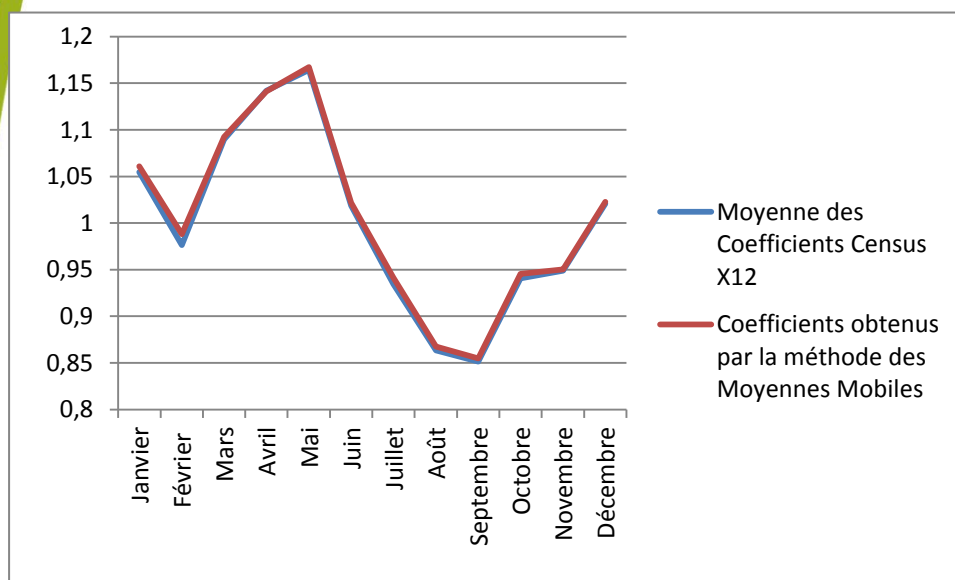
| Années/Mois | Coefficients obtenus par la méthode CensusX12 | | | | | | | | | | | |
|--|---|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-------------|
| | Janvier | Février | Mars | Avril | Mai | Juin | Juillet | Août | Septembre | Octobre | Novembre | Décembre |
| 1990 | 1,016762567 | 0,948431673 | 1,10288307 | 1,17718806 | 1,23180261 | 1,07511283 | 0,98103766 | 0,85985597 | 0,81650918 | 0,90270548 | 0,9066365 | 0,97914396 |
| 1991 | 1,020620193 | 0,952768015 | 1,10195261 | 1,17654353 | 1,22891021 | 1,06948089 | 0,97539314 | 0,86078662 | 0,82110923 | 0,90462555 | 0,90712486 | 0,981709439 |
| 1992 | 1,026483887 | 0,961025611 | 1,10187996 | 1,17121051 | 1,22201096 | 1,06025716 | 0,96401045 | 0,86045574 | 0,82895336 | 0,9120082 | 0,91151202 | 0,988134302 |
| 1993 | 1,031335623 | 0,96700039 | 1,10214813 | 1,16329071 | 1,21134398 | 1,05047383 | 0,95328955 | 0,86048525 | 0,83462554 | 0,92023259 | 0,92213486 | 0,995165944 |
| 1994 | 1,035675655 | 0,971148214 | 1,10107809 | 1,15230637 | 1,20092885 | 1,04580694 | 0,94416626 | 0,8569424 | 0,83856318 | 0,93019154 | 0,9351027 | 1,002510714 |
| 1995 | 1,037186126 | 0,969237649 | 1,09716537 | 1,14602923 | 1,19293525 | 1,04461377 | 0,94171929 | 0,85359344 | 0,84033901 | 0,9327605 | 0,94502873 | 1,005597045 |
| 1996 | 1,040843498 | 0,970752023 | 1,09487498 | 1,1424111 | 1,186497 | 1,04178661 | 0,94053936 | 0,84964843 | 0,84288085 | 0,93414386 | 0,94759111 | 1,009755014 |
| 1997 | 1,046849382 | 0,972354906 | 1,09484313 | 1,14296134 | 1,17915505 | 1,03500306 | 0,93861055 | 0,84803039 | 0,84485118 | 0,93397282 | 0,94771783 | 1,013189022 |
| 1998 | 1,057120507 | 0,978432409 | 1,09679007 | 1,14128242 | 1,16836985 | 1,02465635 | 0,93270722 | 0,85012878 | 0,8472109 | 0,9379654 | 0,94792615 | 1,019739428 |
| 1999 | 1,065771069 | 0,982615719 | 1,09625626 | 1,13954296 | 1,1555749 | 1,01518616 | 0,92585763 | 0,85640066 | 0,85007995 | 0,94202551 | 0,95247193 | 1,024395951 |
| 2000 | 1,070237293 | 0,986230361 | 1,09246931 | 1,13569116 | 1,14511939 | 1,00832148 | 0,92107458 | 0,86388361 | 0,8543766 | 0,94649464 | 0,95701581 | 1,029430777 |
| 2001 | 1,068866714 | 0,986220679 | 1,08577921 | 1,13508337 | 1,13826316 | 1,00422675 | 0,91977694 | 0,86786656 | 0,85840523 | 0,95150306 | 0,96176407 | 1,032556935 |
| 2002 | 1,065734713 | 0,98595093 | 1,07782083 | 1,13464934 | 1,13736158 | 0,99942048 | 0,91911581 | 0,8680067 | 0,8631201 | 0,95618586 | 0,96469156 | 1,03718019 |
| 2003 | 1,064149896 | 0,985779456 | 1,07224649 | 1,13400243 | 1,13464413 | 0,99430568 | 0,91748032 | 0,86763176 | 0,86676643 | 0,96054972 | 0,96822045 | 1,0406135 |
| 2004 | 1,067386748 | 0,985996108 | 1,06958521 | 1,1284422 | 1,13214028 | 0,9885626 | 0,91424831 | 0,86963204 | 0,87011683 | 0,96234059 | 0,97203807 | 1,043627304 |
| 2005 | 1,073089875 | 0,985227568 | 1,07194623 | 1,1220992 | 1,12410064 | 0,98350023 | 0,91356635 | 0,87368672 | 0,87173163 | 0,96275797 | 0,97376882 | 1,04456093 |
| 2006 | 1,079387987 | 0,985165012 | 1,07728038 | 1,11723374 | 1,11668433 | 0,97771449 | 0,91441793 | 0,87640029 | 0,87357276 | 0,96179836 | 0,97198731 | 1,045973675 |
| 2007 | 1,083661287 | 0,98650363 | 1,08331182 | 1,11705536 | 1,10834525 | 0,97124872 | 0,9174959 | 0,87856299 | 0,87470651 | 0,96013572 | 0,96810625 | 1,047497337 |
| 2008 | 1,086282561 | 0,988478822 | 1,08698736 | 1,11816512 | 1,10418222 | 0,96692677 | 0,9190295 | 0,87931617 | 0,87598103 | 0,95944871 | 0,96559364 | 1,049132676 |
| Moyenne des Coefficients Census X12 | 1,054602399 | 0,976279957 | 1,08985782 | 1,14185201 | 1,16412472 | 1,01876867 | 0,93439667 | 0,86322708 | 0,85125787 | 0,94062348 | 0,94875961 | 1,020521797 |
| Coefficients obtenus par la méthode des Moyennes Mobiles | 1,060983 | 0,987899 | 1,092315 | 1,141369 | 1,167304 | 1,021617 | 0,941314 | 0,867662 | 0,854917 | 0,94548 | 0,95026 | 1,02289 |

Les coefficients saisonniers obtenus sont quasiment identiques.

Nous les superposons au sein d'un même graphique pour obtenir un meilleur aperçu.



Graphique N°6 : Représentation graphique des coefficients saisonniers



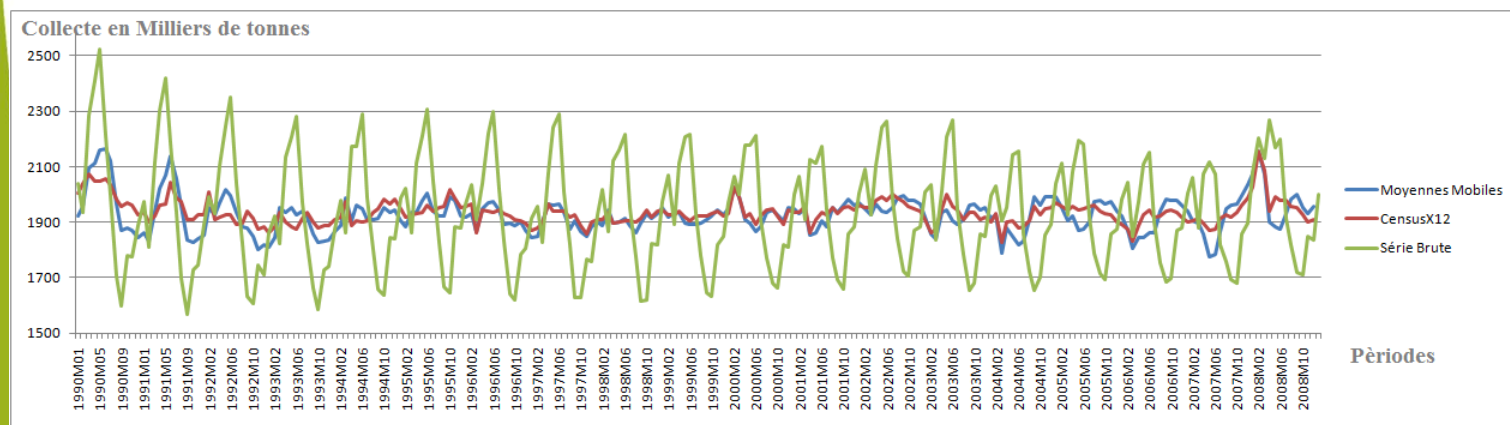
La superposition des moyennes des coefficients saisonniers obtenus respectivement par la méthode Census et celle des Moyennes Mobiles confirme que la nature de la composante saisonnière est Rigide.

Graphique N°7 : Représentation graphique de la chronique désaisonnalisée par les deux méthodes



A la lecture du graphique nous constatons une similitude entre la forme des deux courbes, de fortes variations persistent probablement attribuables au résidu.

Graphique N°8 : Profil temporel de la chronique Brute et Désaisonnalisée par Census et Moyennes Mobiles

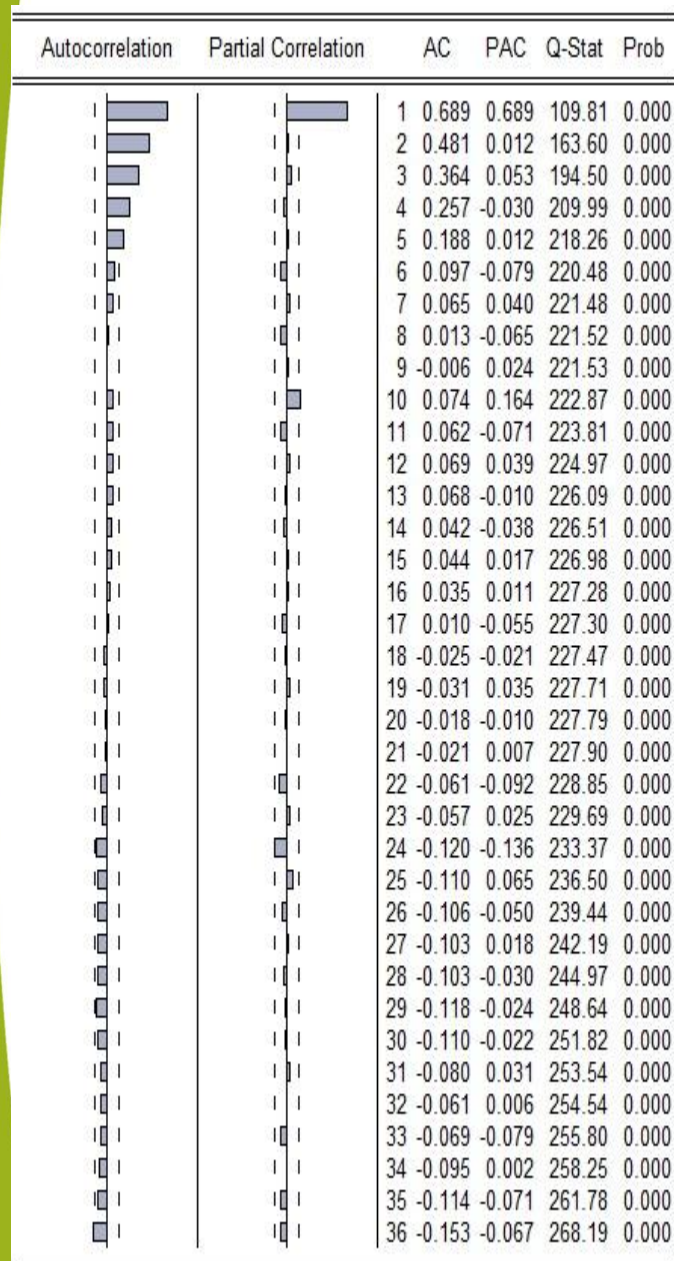


Le graphique ci-dessus montre que la désaisonnalisation par les deux méthodes a lissé notre série. Mais certains pics persistent aux mêmes dates que sur la série brute à nuancer l'efficacité de ces techniques.



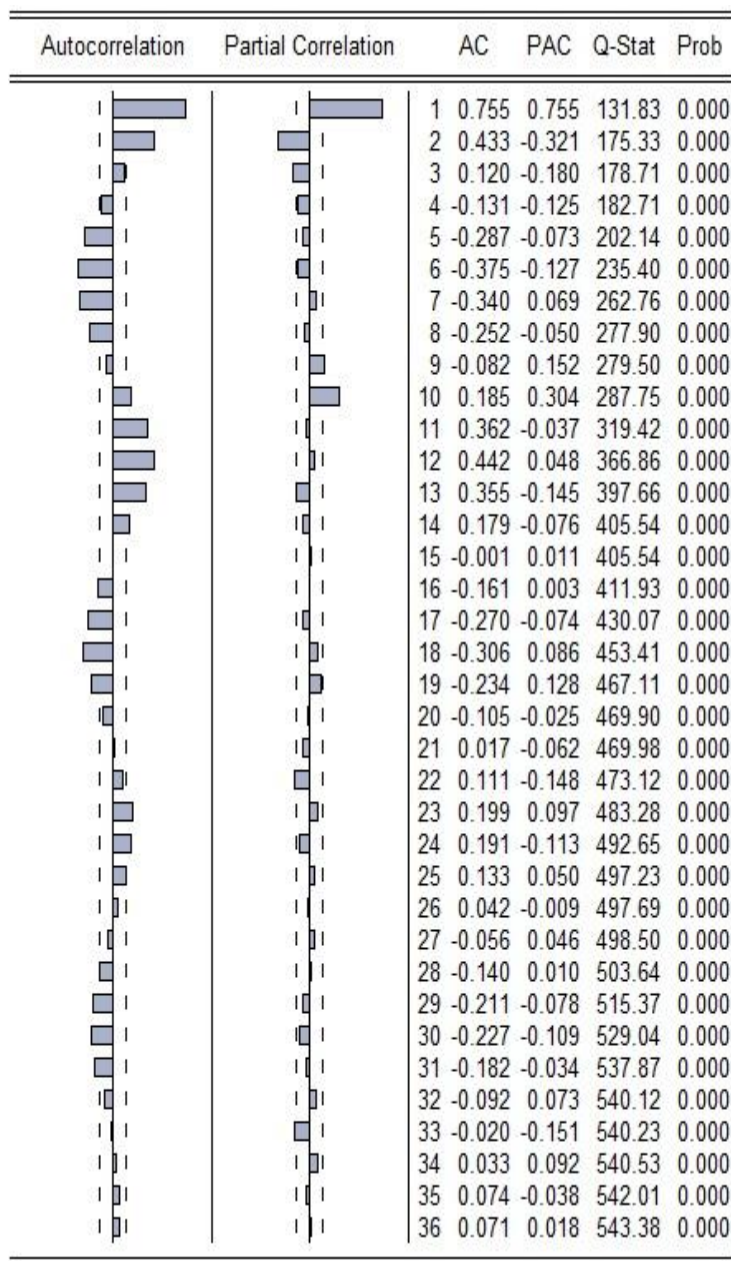
5.2. Comparaison des Corrélogrammes

Graphique N°9 : Corrélogramme de la Série
désaisonnalisée par la méthode Census



Le corrélogramme révèle l'absence de saisonnalité et révèle l'absence de saisonnalité et une tendance légèrement décroissante.

Graphique N°10 : Corrélogramme de la Série
désaisonnalisée par méthode des moyennes mobiles



L'absence de pic nous fait pressentir qu'il s'agit de la meilleure méthode de désaisonnalisation
Détérioration de la composante cyclique ?

Des pics dépassent la limite de la significativité.



5.3. Désaisonnalisation 93-08 (Additif)

22

Nous étudions ici la série de 1993 à 2008

Nous pouvons nous servir des résultats émis tout au long du dossier pour établir que la chronique contient une composante saisonnière de nature rigide (Graphique N°1, Buys-Ballot normal et classé, Superposition des coefficients saisonniers).

Par contre le schéma de décomposition retenu pour la période 1993-2008 est différent de celui retenu antérieurement, il s'agit d'un schéma additif.

Nous allons donc utiliser la Méthode de désaisonnalisation par régression selon un schéma additif.

Tableau N°9 : Estimation de la tendance à partir de la série brute : $\hat{E}t = \hat{a}0 + \hat{a}1 * t$

Dependent Variable: COLLECTE
Method: Least Squares
Date: 11/28/10 Time: 17:07
Sample: 1993M01 2008M12
Included observations: 192

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| C | 1929.497 | 26.72647 | 72.19425 | 0.0000 |
| T | 0.012353 | 0.240164 | 0.051437 | 0.9590 |
| R-squared | 0.000014 | Mean dependent var | 1930.689 | |
| Adjusted R-squared | -0.005249 | S.D. dependent var | 183.9614 | |
| S.E. of regression | 184.4435 | Akaike info criterion | 13.28293 | |
| Sum squared resid | 6463690. | Schwarz criterion | 13.31686 | |
| Log likelihood | -1273.161 | F-statistic | 0.002646 | |
| Durbin-Watson stat | 0.645735 | Prob(F-statistic) | 0.959031 | |

Les coefficients de régression sur le temps sont égaux à :
 $\hat{a}1 = 0,012353$
 $\hat{a}0 = 1929,497$

Tableau N°10 : Coefficients saisonniers

| | Janvier | Février | Mars | avril | Mai | Juin | Juillet | Août | Septembre | Octobre | Novembre | Décembre |
|---------------|------------|-------------|------------|------------|------------|-------------|-------------|-------------|-------------|-------------|-------------|------------|
| 1993 | -8,009 | -108,821706 | 203,765941 | 278,753588 | 351,641235 | 39,328882 | -105,583471 | -265,495824 | -343,308177 | -200,92053 | -188,132883 | -53,245236 |
| 1994 | 49,142411 | -66,469942 | 241,317705 | 243,005352 | 359,692999 | 58,680646 | -130,831707 | -273,14406 | -295,356413 | -85,468766 | -91,881119 | 58,206528 |
| 1995 | 92,635175 | -70,375178 | 184,117469 | 280,201116 | 377,868763 | 118,70841 | -104,503943 | -263,308296 | -285,676649 | -47,836002 | -50,877355 | 35,140292 |
| 1996 | 104,792939 | -21,138414 | 106,662233 | 291,77488 | 368,641527 | 86,746174 | -99,999179 | -290,509532 | -310,851885 | -146,065238 | -123,972591 | -17,439944 |
| 1997 | 27,083703 | -104,67465 | 147,002997 | 313,943644 | 358,171291 | 76,344938 | -110,005415 | -302,420768 | -302,530121 | -165,521474 | -172,183827 | -5,67618 |
| 1998 | 87,649467 | -63,762886 | 191,024761 | 231,712408 | 287,900055 | 22,487702 | -154,824651 | -313,637004 | -309,449357 | -107,06171 | -111,574063 | 45,013584 |
| 1999 | 139,350231 | -37,619122 | 180,837525 | 276,999172 | 284,157819 | 1,914466 | -150,110887 | -286,01324 | -296,898593 | -113,648946 | -83,638299 | 39,948348 |
| 2000 | 135,981995 | 69,106642 | 244,730289 | 247,878936 | 282,546583 | -22,86877 | -158,630123 | -251,540476 | -268,107829 | -110,549182 | -121,741535 | 67,944112 |
| 2001 | 135,799759 | -22,040594 | 195,621053 | 183,5607 | 241,031347 | 13,537994 | -159,663359 | -237,833712 | -274,972065 | -75,108418 | -47,729771 | 74,853876 |
| 2002 | 160,376523 | -5,07683 | 171,846817 | 311,807464 | 334,384111 | 46,112758 | -91,768595 | -207,558948 | -225,686301 | -60,415654 | -49,994007 | 79,19964 |
| 2003 | 102,481287 | -94,676066 | 80,358581 | 276,177228 | 338,627875 | 13,052522 | -148,885831 | -275,614184 | -253,258537 | -72,92289 | -84,414243 | 64,053404 |
| 2004 | 98,722051 | -27,234302 | 20,672345 | 211,553992 | 224,891639 | -72,565714 | -209,083067 | -275,54342 | -228,329773 | -77,714126 | -39,268479 | 107,465168 |
| 2005 | 181,431815 | -4,300538 | 150,827109 | 261,854756 | 251,502403 | -14,80995 | -144,042303 | -217,254656 | -239,787009 | -73,619362 | -56,501715 | 53,135932 |
| 2006 | 109,667579 | -82,928774 | 41,904873 | 174,74652 | 221,082167 | -30,004186 | -179,488539 | -247,481892 | -236,221245 | -61,929598 | -52,579951 | 73,718696 |
| 2007 | 127,855343 | -51,51601 | 140,861637 | 185,611284 | 140,416931 | -111,801422 | -176,088775 | -240,582128 | -253,463481 | -73,145834 | -33,640187 | 147,35546 |
| 2008 | 269,768107 | 199,795754 | 334,284401 | 236,644048 | 268,531695 | -17,149658 | -114,106011 | -211,653364 | -221,672717 | -84,76607 | -95,533423 | 69,641224 |
| Moyenne Ecart | 113,421 | -30,733 | 164,740 | 250,389 | 293,193 | 12,982 | -139,851 | -259,974 | -271,598 | -97,293 | -87,729 | 52,457 |
| Somme | 0,041 | | | | | | | | | | | |

Les coefficients saisonniers provisoires sont donnés par les écarts entre la série observée et la tendance :

Les écarts relatifs aux mêmes mois sont rassemblés afin de calculer leurs moyennes qui correspondent aux coefficients saisonniers des mois considérés (bleu foncé).

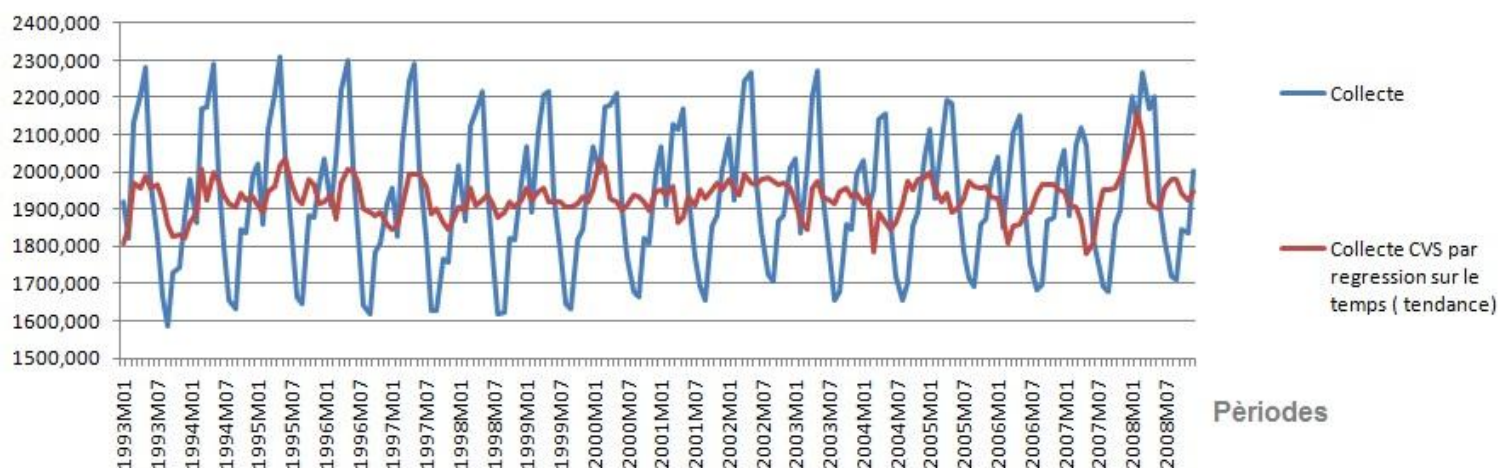
Le principe de conservation des aires est vérifié, la normalisation des coefficients saisonniers ne s'impose pas.

Enfin nous pouvons calculer la série **corrigée des variations saisonnières (CVS)** par différence entre la série brute et le coefficient saisonnier du mois considéré.











































































Graphique N°11 : Série CVS par régression sur le temps

Milliers de tonnes



La désaisonnalisation par régression a retiré les variations saisonnières de la série et fait apparaître la tendance et les variations résiduelles.

| Autocorrelation | Partial Correlation | AC | PAC | Q-Stat | Prob | |
|---|---|----|--------|--------|--------|-------|
|  |  | 1 | 0.663 | 0.663 | 85.804 | 0.000 |
|  |  | 2 | 0.334 | -0.189 | 107.67 | 0.000 |
|  |  | 3 | 0.073 | -0.120 | 108.73 | 0.000 |
|  |  | 4 | -0.100 | -0.084 | 110.70 | 0.000 |
|  |  | 5 | -0.187 | -0.057 | 117.68 | 0.000 |
|  |  | 6 | -0.274 | -0.167 | 132.74 | 0.000 |
|  |  | 7 | -0.257 | 0.020 | 146.03 | 0.000 |
|  |  | 8 | -0.239 | -0.108 | 157.57 | 0.000 |
|  |  | 9 | -0.134 | 0.078 | 161.23 | 0.000 |
|  |  | 10 | 0.107 | 0.264 | 163.56 | 0.000 |
|  |  | 11 | 0.225 | -0.044 | 174.00 | 0.000 |
|  |  | 12 | 0.283 | 0.053 | 190.53 | 0.000 |
|  |  | 13 | 0.168 | -0.161 | 196.44 | 0.000 |
|  |  | 14 | 0.039 | -0.047 | 196.76 | 0.000 |
|  |  | 15 | -0.057 | -0.030 | 197.44 | 0.000 |
|  |  | 16 | -0.136 | -0.010 | 201.35 | 0.000 |
|  |  | 17 | -0.178 | -0.066 | 208.13 | 0.000 |
|  |  | 18 | -0.188 | 0.067 | 215.68 | 0.000 |
|  |  | 19 | -0.129 | 0.066 | 219.26 | 0.000 |
|  |  | 20 | -0.046 | -0.009 | 219.72 | 0.000 |
|  |  | 21 | 0.010 | -0.051 | 219.74 | 0.000 |
|  |  | 22 | 0.052 | -0.109 | 220.34 | 0.000 |
|  |  | 23 | 0.116 | 0.112 | 223.31 | 0.000 |
|  |  | 24 | 0.096 | -0.097 | 225.36 | 0.000 |
|  |  | 25 | 0.052 | 0.035 | 225.97 | 0.000 |
|  |  | 26 | -0.001 | -0.013 | 225.97 | 0.000 |
|  |  | 27 | -0.058 | -0.001 | 226.72 | 0.000 |
|  |  | 28 | -0.095 | -0.006 | 228.78 | 0.000 |
|  |  | 29 | -0.123 | -0.029 | 232.23 | 0.000 |
|  |  | 30 | -0.094 | -0.030 | 234.26 | 0.000 |
|  |  | 31 | -0.051 | -0.015 | 234.85 | 0.000 |
|  |  | 32 | 0.017 | 0.101 | 234.92 | 0.000 |
|  |  | 33 | 0.039 | -0.124 | 235.27 | 0.000 |
|  |  | 34 | 0.068 | 0.123 | 236.36 | 0.000 |
|  |  | 35 | 0.092 | -0.055 | 238.35 | 0.000 |
|  |  | 36 | 0.053 | -0.028 | 239.03 | 0.000 |

Graphique N°12 : Le Corrélogramme nous montre que la saisonnalité a bien été éliminée

Dans notre cas, nous avons vu que la saisonnalité est plutôt rigide, les coefficients des méthodes Census et moyennes mobiles sont quasiment identiques.

Au vu de la complexité de notre extra-saisonnier, des différents graphiques et corrélogramme nous allons conclure que pour notre série dans sa totalité la meilleure méthode de désaisonnalisation est celle de CENSUSX12 méthode habituellement employée pour les saisonnalités souple mais qui peut s'appliquer à une saisonnalité rigide celle-ci étant un cas particulier de la saisonnalité souple.

Ainsi nous utiliserons donc la série désaisonnalisée par la méthode CensusX12 dans la suite de notre dossier.

Nous allons donc maintenant réitérer l'analyse de la variance et le test de Fisher sur la série corrigée des variations saisonnières par la méthode CensusX12.

Ce test devrait confirmer que la série ne contient plus de composante saisonnière et nous renseigner sur la présence d'une extra-saisonnalité.

N : Le nombre d'années = 19

p : La périodicité = 12

Deux effets seront testés :

- **L'effet période** soit l'effet mois, si il est significatif la série est saisonnière ;
- **L'effet année**, s'il est significatif la série contient une tendance ou a été mal transformé.

Tableau N°11 : Analyse de la variance pour notre série

| Désignation | Somme des carrés (S) | Degré de liberté | Variance (V) |
|---------------|----------------------|------------------|--------------|
| Période (p) | 6065,506031 | 11 | 551,4096392 |
| Année (a) | 170576,6167 | 18 | 9476,478708 |
| Résidu (r) | 252326,8334 | 198 | 1274,377946 |
| Total (t) | 428968,9561 | 227 | 1889,731084 |

A partir de ce tableau nous allons construire les tests d'hypothèse :

Test de l'influence du facteur colonne, la période : ici le mois

On va tester l'hypothèse H_0 : Pas d'influence du facteur colonne.

Règles de décision :

Si le Fisher empirique $F_c = V_p/V_r >$ au Fisher lu dans la table (F_t), alors on rejette l'hypothèse H_0 , la série est saisonnière.

Si le Fisher empirique $F_c <$ au Fisher lu dans la table (F_t), alors on accepte l'hypothèse H_0 , pas d'influence du facteur colonne, pas d'effet mois, la chronique est non saisonnière.

Test au seuil de 5% :

La valeur du Fisher théorique (à $v_1=11$ et $v_2=198$ degrés de liberté) est de 1,84 ($F_t=1,84$).

La valeur du Fisher calculé est de 0,432689251 ($F_c \approx 0,433$).

$F_c = 0,433 < F_t = 1,84$; Le Fisher calculé est plus petit que le Fisher théorique.

On accepte l'hypothèse H_0 , pas d'influence du facteur colonne, la série n'est pas affectée d'une saisonnalité.

Test de l'influence du facteur ligne, la tendance

On va tester l'hypothèse H_0 : Pas d'influence du facteur année.

Règles de décision :

Si Fisher empirique $F_c = V_a/V_r >$ valeur du Fisher théorique (F_t), alors on rejette l'hypothèse H_0 , la série est donc affectée d'une tendance.

Si Fisher empirique $F_c <$ valeur du Fisher théorique (F_t), alors l'hypothèse H_0 est acceptée, pas d'influence du facteur année, la série n'est pas affectée d'une tendance.

Test au seuil de 5% :

La valeur du Fisher théorique (à $v_3=18$ et $v_2=198$ degrés de liberté) est de 1,66 ($F_t=1,66$)

La valeur du Fisher calculé est de 7,436160313 ($F_c \approx 7,436$)

$F_c \approx 7,436 > F_t = 1,66$; Le Fisher calculé est plus grand que le Fisher théorique.

On rejette l'hypothèse H_0 , la série est donc affectée d'une tendance (plus à la limite de la significativité comme c'était le cas avec la série brute)



6. Prédiction par les méthodes traditionnelles

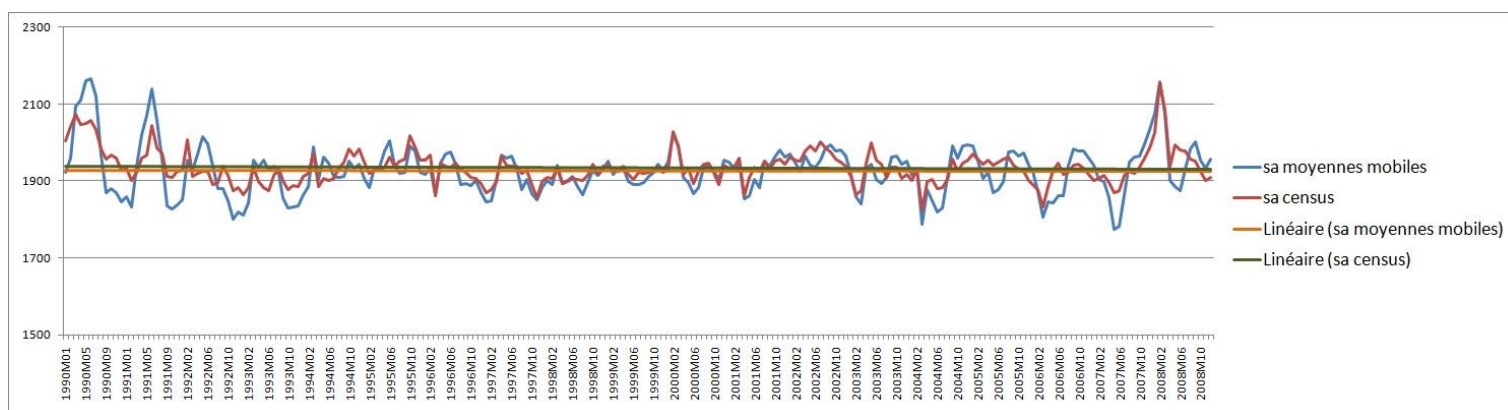
26

Afin de mieux sélectionner la méthode traditionnelle de prévision il est préférable de connaître la nature de notre tendance.

L'analyse des corrélogrammes après désaisonnalisation, la représentation graphique du profil temporel de la série, et le graphique ci-dessous nous ont donné une idée de la forme prise par la tendance, qui peut être supposée linéaire.

6.1. Etude de la non-Stationnarité

Graphique N° 13 : Séries désaisonnalisées et Courbes de tendance linéaire



Un extra-saisonnier a été décelé durant l'analyse de la saisonnalité.

On peut supposer au regard du graphique ci-dessus que la tendance est linéaire.

Il existe des tests permettant de déceler la non-stationnarité et ses caractéristiques, leurs conclusions serviront à pouvoir choisir la transformation la plus adaptée.

Les cas de non stationnarité les plus fréquents sont analysés à partir de deux types de processus :

- Processus TS : Non stationnarité de type déterministe ;
- Processus DS : Non stationnarité de type aléatoire.

Test de racine unitaire et identification du processus :

Nous allons appliquer les tests de Dickey-Fuller sur notre série désaisonnalisée par la méthode Census X12.

Test de l'hypothèse H_0 : Présence d'une racine unitaire.

Nous suivrons la stratégie des test de Dickey-Fuller simples et tester le modèle 3 :

$$X_t = c + bt + \Phi_1 x_{t-1} + a_t$$



Tableau N°12 : Dickey et Fuller Augmenté

Null Hypothesis: COLLECTE_SA has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic based on SIC, MAXLAG=14)

| | t-Statistic | Prob.* |
|--|-------------|--------|
| Augmented Dickey-Fuller test statistic | -6.441883 | 0.0000 |
| Test critical values: 1% level | -3.998997 | |
| 5% level | -3.429745 | |
| 10% level | -3.138397 | |

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

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(COLLECTE_SA)
Method: Least Squares
Date: 11/29/10 Time: 01:29
Sample (adjusted): 1990M02 2008M12
Included observations: 227 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| COLLECTE_SA(-1) | -0.307887 | 0.047795 | -6.441883 | 0.0000 |
| C | 595.9936 | 92.76977 | 6.424437 | 0.0000 |
| @TREND(1990M01) | -0.006554 | 0.031680 | -0.206894 | 0.8363 |
| R-squared | 0.156386 | Mean dependent var | -0.430428 | |
| Adjusted R-squared | 0.148854 | S.D. dependent var | 33.84599 | |
| S.E. of regression | 31.22549 | Akaike info criterion | 9.733474 | |
| Sum squared resid | 218407.0 | Schwarz criterion | 9.778738 | |
| Log likelihood | -1101.749 | F-statistic | 20.76219 | |
| Durbin-Watson stat | 2.030222 | Prob(F-statistic) | 0.000000 | |

Test individuel d'hypothèse H0 :

H0 : $\Phi_1 = 1$ ou $\Phi_1 - 1 = 0$, $\rho = \Phi_1 - 1$, il existe une racine unitaire.

H1 : $|\Phi_1| < 1$, il n'y a pas de racine unitaire et le processus est stationnaire.

Règles de décision :

Si t-Statistic Augmented Dickey-Fuller (Tadf) < t-Statistic (Tt) à 0,05% alors H0 est rejetée.

Si t-Statistic Augmented Dickey-Fuller (Tadf) > t-Statistic (Tt) à 0,05% alors H0 est acceptée.

Tadf = -6,441883 < Tt = -3,429745

H0 est rejetée, notre modèle ne possède pas de racine unitaire.

Test de Student $b = 0$

On teste l'hypothèse $H_0 : b$ n'est pas significativement différent de 0.

Règles de décision :

Si le Student calculé $>$ Student tabulé alors on rejette H_0

Si le Student calculé $<$ Student tabulé alors on accepte H_0

Student calculé = $-0,206894 <$ Student tabulé = $1,972$ donc on accepte H_0 , b n'est pas significativement différent de 0.

La tendance du modèle : $X_t = c + bt + \Phi_1 x_{t-1} + a_t$ n'est pas significative.

Test de Student $c=0$

On teste l'hypothèse $H_0 : c$ n'est pas significativement différent de 0.

Règles de décision :

Si le Student calculé $>$ Student tabulé alors on rejette H_0

Si le Student calculé $<$ Student tabulé alors on accepte H_0

Student calculé = $6,424437 >$ Student tabulé = $1,972$ alors on rejette H_0 ,
« c » est donc significativement différent de 0.

La constante du modèle : $X_t = c + bt + \Phi_1 x_{t-1} + a_t$ est significative.

D'après la figure de stratégie des tests de Dickey-Fuller notre processus générateur est un processus autorégressif d'ordre 1 [AR(1)] sans tendance et avec constante. Il ne dépend donc pas du temps.

$$x_t = c + \Phi_1 x_{t-1} + a_t$$

6.2. Pr vision de la chronique par r gression

Mod le th orique

Mod le I : $X_{\text{census}} = aT + b + \epsilon$ et avec un sch ma de d composition de type multiplicatif.

T : le temps, correspondant au nombre d'observation de la s rie soit $n=228$.

b : la constante

ϵ : le terme d'erreur

On estime ce mod le par la m thode des Moindres Carr s Ordinaires.

Tableau N  13 : MCO-1

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| C | 1939.259 | 5.778632 | 335.5914 | 0.0000 |
| T | -0.040617 | 0.043755 | -0.928288 | 0.3542 |
| R-squared | 0.003798 | Mean dependent var | 1934.609 | |
| Adjusted R-squared | -0.000610 | S.D. dependent var | 43.47104 | |
| S.E. of regression | 43.48428 | Akaike info criterion | 10.39141 | |
| Sum squared resid | 427339.5 | Schwarz criterion | 10.42149 | |
| Log likelihood | -1182.621 | F-statistic | 0.861719 | |
| Durbin-Watson stat | 0.605909 | Prob(F-statistic) | 0.354248 | |

On remarque directement un coefficient de d termination R tr s faible, seulement 0.37% de la variance totale de collecte_sa est expliqu e par la r gression.

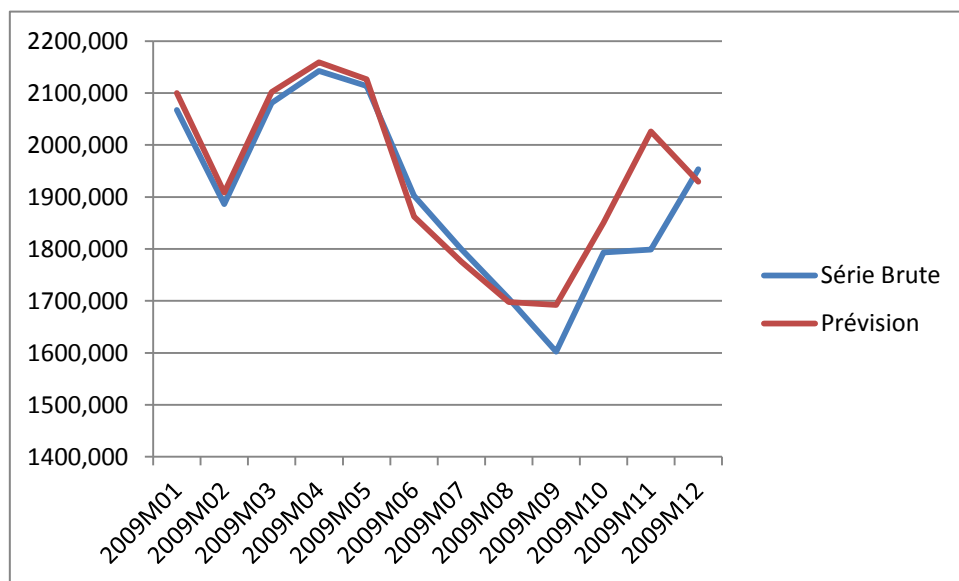
$X_{\text{census}} = \text{Tendance} = aT + b + \epsilon$ avec $a = -0,040617$ et $b = 1939,259$

Le sch ma  tant multiplicatif (Test de Buys-Ballot) on obtient la pr vision en multipliant les valeurs de X_{census} estim es par les coefficients saisonniers pr visionnels de la m thode Census.

Tableau N  14 : R capitulatif-1

| Date | Tendance | Coef Pr v Census | Pr vision | S rie | MSE |
|---------|------------|------------------|------------|----------|------------|
| 2009M01 | 1929,83586 | 1,088 | 2099,66141 | 2067,105 | 1059,91992 |
| 2009M02 | 1929,79524 | 0,989 | 1908,56749 | 1885,962 | 511,00824 |
| 2009M03 | 1929,75462 | 1,089 | 2101,50278 | 2081,064 | 417,743865 |
| 2009M04 | 1929,71401 | 1,119 | 2159,34997 | 2142,176 | 294,9453 |
| 2009M05 | 1929,67339 | 1,102 | 2126,50007 | 2113,554 | 167,600821 |
| 2009M06 | 1929,63277 | 0,965 | 1862,09562 | 1902,408 | 1625,08766 |
| 2009M07 | 1929,59215 | 0,92 | 1775,22478 | 1799,622 | 595,224262 |
| 2009M08 | 1929,55154 | 0,88 | 1698,00535 | 1704,503 | 42,2194223 |
| 2009M09 | 1929,51092 | 0,877 | 1692,18108 | 1602,158 | 8104,15436 |
| 2009M10 | 1929,4703 | 0,959 | 1850,36202 | 1793,324 | 3253,33579 |
| 2009M11 | 1929,42969 | 1,05 | 2025,90117 | 1798,9 | 51529,5313 |
| 2009M12 | 1929,38907 | 1 | 1929,38907 | 1953,643 | 588,253169 |
| Somme : | | | | | 68189,0241 |

Graphique N°14 : Prévion-1



Modèle II : $X_{\text{Moyennes Mobiles}} = aT + b + \varepsilon_t$ avec un schéma de décomposition de type multiplicatif

T : le temps, correspondant au nombre d'observation de la série soit $n=228$.

b : la constante

ε_t : le terme d'erreur

On estime ce modèle par la méthode des Moindres Carrés Ordinaires.

Tableau N°15 : MCO-2

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| C | 1928.066 | 8.729796 | 220.8603 | 0.0000 |
| T | -0.013203 | 0.066100 | -0.199749 | 0.8419 |
| R-squared | 0.000177 | Mean dependent var | 1926.554 | |
| Adjusted R-squared | -0.004247 | S.D. dependent var | 65.55276 | |
| S.E. of regression | 65.69183 | Akaike info criterion | 11.21656 | |
| Sum squared resid | 975284.1 | Schwarz criterion | 11.24664 | |
| Log likelihood | -1276.688 | F-statistic | 0.039900 | |
| Durbin-Watson stat | 0.487943 | Prob(F-statistic) | 0.841856 | |

Même conclusion pour le modèle à partir de la série désaisonnalisée par moyennes mobiles, le coefficient de détermination étant encore plus faible que précédemment.

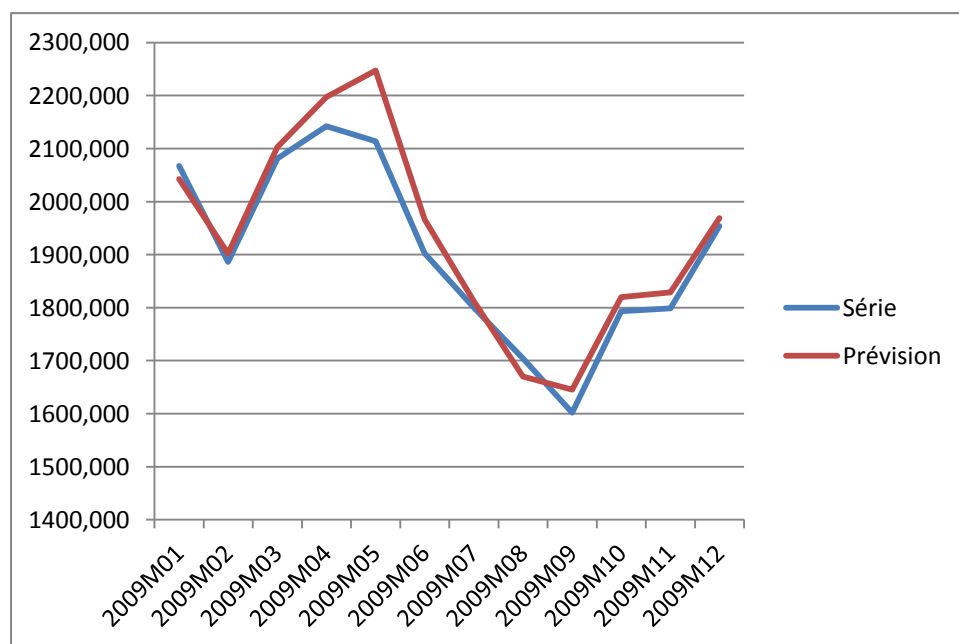
$X_{\text{Moyennes Mobiles}} = aT + b + \varepsilon_t$ avec $a = 0,013203$ et $b = 1928,066$

Le schéma étant multiplicatif (Test de Buys-Ballot) on obtient la prévision en multipliant les valeurs de $X_{\text{Moyennes Mobiles}}$ estimées par les coefficients saisonniers prévisionnels de la méthode des Moyennes Mobiles.

Tableau N°16 : Récapitulatif-2

| Date | Tendance | CoefPrév MAM | Prévision | Série | MSE |
|---------|------------|--------------|-------------|----------|------------|
| 2009M01 | 1925,0029 | 1,060983 | 2042,395356 | 2067,105 | 610,566502 |
| 2009M02 | 1924,9897 | 0,987899 | 1901,695401 | 1885,962 | 247,539895 |
| 2009M03 | 1924,9765 | 1,092315 | 2102,680703 | 2081,064 | 467,281866 |
| 2009M04 | 1924,9633 | 1,141369 | 2197,093431 | 2142,176 | 3015,92423 |
| 2009M05 | 1924,95009 | 1,167304 | 2247,001942 | 2113,554 | 17808,3533 |
| 2009M06 | 1924,93689 | 1,021617 | 1966,54825 | 1902,408 | 4113,97164 |
| 2009M07 | 1924,92369 | 0,941314 | 1811,957615 | 1799,622 | 152,167387 |
| 2009M08 | 1924,91048 | 0,867662 | 1670,17168 | 1704,503 | 1178,63957 |
| 2009M09 | 1924,89728 | 0,854917 | 1645,627408 | 1602,158 | 1889,58943 |
| 2009M10 | 1924,88408 | 0,94548 | 1819,939397 | 1793,324 | 708,379364 |
| 2009M11 | 1924,87087 | 0,95026 | 1829,127797 | 1798,9 | 913,719695 |
| 2009M12 | 1924,85767 | 1,02289 | 1968,917663 | 1953,643 | 233,315332 |
| Somme : | | | | | 31339,4482 |

Graphique N°15 : Prévision-2





Modèle III : $XCVSRégT = aT + b + \varepsilon$ et avec un schéma de décomposition de type additif.

Tableau N°17 : MCO-3

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| C | 1919.759 | 7.639512 | 251.2934 | 0.0000 |
| T | 0.113268 | 0.068649 | 1.649973 | 0.1006 |
| R-squared | 0.014126 | Mean dependent var | 1930.689 | |
| Adjusted R-squared | 0.008937 | S.D. dependent var | 52.95866 | |
| S.E. of regression | 52.72147 | Akaike info criterion | 10.77828 | |
| Sum squared resid | 528115.2 | Schwarz criterion | 10.81222 | |
| Log likelihood | -1032.715 | F-statistic | 2.722410 | |
| Durbin-Watson stat | 0.654325 | Prob(F-statistic) | 0.100601 | |

Le R-squared nous indique que 1.4% de la variance total de notre Y_t est expliqué par la régression, ce qui est très faible.

$XCVSRégT = aT + b + \varepsilon$ avec $a = 0,113268$ et $b = 1919,759$.

Le schéma étant additif (Test de Buys-Ballot) on obtient la prévision en ajoutant aux valeurs de $XCVSRégT$ estimées les coefficients saisonniers prévisionnels par régression sur le temps.

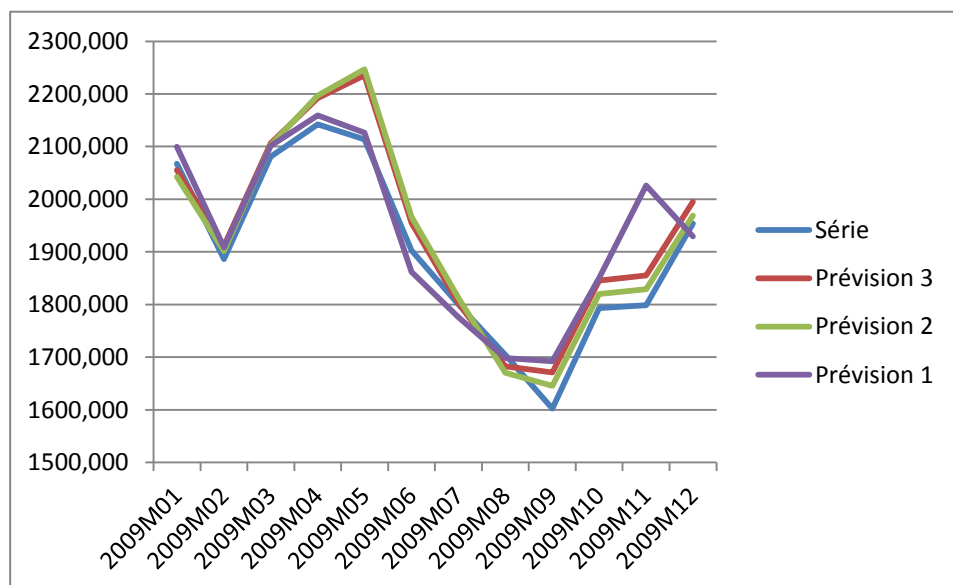
Tableau N°18 : Récapitulatif-3

| Date | Tendance | Coef Reg | Prévision | Série | MSE |
|---------|------------|------------|------------|----------|------------|
| 2009M01 | 1941,61972 | 113,420565 | 2055,04029 | 2067,105 | 145,557264 |
| 2009M02 | 1941,73299 | 30,7332885 | 1910,9997 | 1885,962 | 626,886597 |
| 2009M03 | 1941,84626 | 164,739734 | 2106,58599 | 2081,064 | 651,372152 |
| 2009M04 | 1941,95953 | 250,389068 | 2192,3486 | 2142,176 | 2517,28939 |
| 2009M05 | 1942,0728 | 293,193028 | 2235,26582 | 2113,554 | 14813,768 |
| 2009M06 | 1942,18606 | 12,9821745 | 1955,16824 | 1902,408 | 2783,64277 |
| 2009M07 | 1942,29933 | 139,850991 | 1802,44834 | 1799,622 | 7,98820345 |
| 2009M08 | 1942,4126 | 259,974469 | 1682,43813 | 1704,503 | 486,858444 |
| 2009M09 | 1942,52587 | 271,598135 | 1670,92773 | 1602,158 | 4729,27625 |
| 2009M10 | 1942,63914 | 97,2933625 | 1845,34577 | 1793,324 | 2706,26492 |
| 2009M11 | 1942,7524 | 87,7289655 | 1855,02344 | 1798,9 | 3149,84035 |
| 2009M12 | 1942,86567 | 52,4571815 | 1995,32285 | 1953,643 | 1737,21019 |
| Somme : | | | | | 34355,9545 |

Graphique N°16 : Préviation-3



Graphique N°17 : Comparatif des 3 modèles



Les 3 prévisions par régression confirment ce que nous a appris le test de Dickey-Fuller, à savoir que notre modèle à partir de la série désaisonnalisée ne contient pas de tendance, n'est pas fonction du temps.

En effet les coefficients associés à cette composante sont respectivement pour les 3 modèles : $-0,040617$ / $0,013203$ / $0,113268$, qui sont des valeurs très faible.

A contrario la constante occupe une place importante dans l'explication des 3 modèles, ce qui vient confirmer le test de Dickey et Fuller qui stipule que nous sommes en présence d'un modèle sans tendance MAIS avec Constante.

La minimisation du critère des MSE nous amène à préférer le modèle II.

Nous allons réitérer le test de Dickey-Fuller sans tendance sur la série désaisonnalisée par Census puisque nous sommes en présence d'un processus autorégressif d'ordre 1 [AR(1)] sans tendance et avec constante.

Tableau N°19 : Estimation

Sample (adjusted): 1990M02 2008M12
Included observations: 227 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| COLLECTE_SA(-1) | -0.307320 | 0.047614 | -6.454363 | 0.0000 |
| C | 594.1491 | 92.14377 | 6.448066 | 0.0000 |
| R-squared | 0.156225 | Mean dependent var | -0.430428 | |
| Adjusted R-squared | 0.152475 | S.D. dependent var | 33.84599 | |
| S.E. of regression | 31.15900 | Akaike info criterion | 9.724855 | |
| Sum squared resid | 218448.7 | Schwarz criterion | 9.755031 | |
| Log likelihood | -1101.771 | F-statistic | 41.65880 | |
| Durbin-Watson stat | 2.031000 | Prob(F-statistic) | 0.000000 | |

Le modèle AR(1) s'écrit donc : $x_t = c + \phi_1 x_{t-1} + a_t$

Avec comme valeurs :

$$c = 594.1491$$

$$\phi_1 = -0,307320$$

Nous générerons le modèle par la suite.



7. Algorithme de Box et Jenkins sur série désaisonnalisée

35

Transformation de la chronique afin d'éliminer une éventuelle non stationnarité.

Nous allons utiliser la transformation de Box-Cox afin de rendre linéaire notre extra-saisonnier et afin de rendre le schéma de notre chronique additif.

La transformation de Box-Cox est une transformation logarithmique pour un paramètre inconnu $\lambda = 0$. On obtient alors notre nouvelle série désaisonnalisée transformée en \ln par Box-Cox : $yt = \ln xt$.

7.1. Identification du processus ARMA de notre série

Pour identifier le processus, nous allons étudier la FAP et la FAC des résidus de notre chronique transformée.

| Autocorrelation | Partial Correlation | AC | PAC | Q-Stat | Prob |
|-----------------|---------------------|-----------|--------|--------|-------|
| | | 1 0.689 | 0.689 | 109.54 | 0.000 |
| | | 2 0.480 | 0.012 | 163.06 | 0.000 |
| | | 3 0.363 | 0.053 | 193.76 | 0.000 |
| | | 4 0.253 | -0.035 | 208.78 | 0.000 |
| | | 5 0.186 | 0.018 | 216.88 | 0.000 |
| | | 6 0.092 | -0.088 | 218.85 | 0.000 |
| | | 7 0.060 | 0.045 | 219.71 | 0.000 |
| | | 8 0.005 | -0.073 | 219.72 | 0.000 |
| | | 9 -0.013 | 0.030 | 219.76 | 0.000 |
| | | 10 0.070 | 0.161 | 220.93 | 0.000 |
| | | 11 0.057 | -0.071 | 221.72 | 0.000 |
| | | 12 0.062 | 0.033 | 222.66 | 0.000 |
| | | 13 0.062 | -0.008 | 223.59 | 0.000 |
| | | 14 0.036 | -0.039 | 223.90 | 0.000 |
| | | 15 0.039 | 0.016 | 224.27 | 0.000 |
| | | 16 0.029 | 0.010 | 224.47 | 0.000 |
| | | 17 0.008 | -0.049 | 224.49 | 0.000 |
| | | 18 -0.028 | -0.024 | 224.69 | 0.000 |
| | | 19 -0.033 | 0.037 | 224.95 | 0.000 |
| | | 20 -0.019 | -0.013 | 225.04 | 0.000 |
| | | 21 -0.023 | 0.007 | 225.18 | 0.000 |
| | | 22 -0.063 | -0.094 | 226.18 | 0.000 |
| | | 23 -0.058 | 0.028 | 227.05 | 0.000 |
| | | 24 -0.122 | -0.142 | 230.89 | 0.000 |
| | | 25 -0.112 | 0.070 | 234.15 | 0.000 |
| | | 26 -0.106 | -0.048 | 237.04 | 0.000 |
| | | 27 -0.101 | 0.019 | 239.69 | 0.000 |
| | | 28 -0.101 | -0.034 | 242.38 | 0.000 |
| | | 29 -0.115 | -0.018 | 245.87 | 0.000 |
| | | 30 -0.107 | -0.030 | 248.89 | 0.000 |
| | | 31 -0.079 | 0.034 | 250.56 | 0.000 |
| | | 32 -0.061 | 0.001 | 251.54 | 0.000 |
| | | 33 -0.066 | -0.073 | 252.73 | 0.000 |
| | | 34 -0.092 | 0.000 | 255.04 | 0.000 |
| | | 35 -0.111 | -0.066 | 258.39 | 0.000 |
| | | 36 -0.151 | -0.075 | 264.62 | 0.000 |

Graphique N°18 : Corrélogramme

Ici, à la vue du corrélogramme, nous pouvons distinguer uniquement un AR(1). La phase d'identification est donc très simplifiée et les étapes d'estimation et de tests d'adéquation peuvent être rassemblées en une seule étape.



Tableau N°20 : Estimation et tests d'adéquation du processus AR(1) :

Modèle retenu : les résidus (RES1) suivent un processus AR(1)

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| AR(1) | 0.691265 | 0.047668 | 14.50177 | 0.0000 |
| R-squared | 0.481986 | Mean dependent var | -0.285063 | |
| Adjusted R-squared | 0.481986 | S.D. dependent var | 42.85456 | |
| S.E. of regression | 30.84378 | Akaike info criterion | 9.700143 | |
| Sum squared resid | 215002.6 | Schwarz criterion | 9.715231 | |
| Log likelihood | -1099.966 | Durbin-Watson stat | 2.026507 | |
| Inverted AR Roots | .69 | | | |

Tests de validité du modèle :

Test de significativité du coefficient d'AR (1), test de Student (n-2) :

$H_0 : b=0$ contre $H_1 : b \neq 0$

$t\text{-Statistic} = 14,50177 > t_{lu} = 1,96$

On rejette H_0 , le paramètre de AR(1) est significativement différent de 0.

Test du coefficient de détermination R^2 :

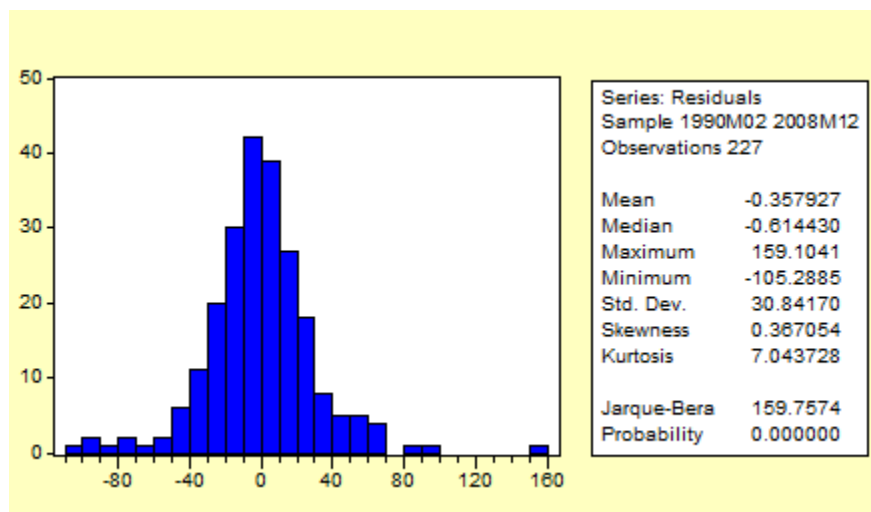
$H_0 : R^2=0$ contre $H_1 : R^2 \neq 0$

$R^2=0,481986$

48% de la variance est expliquée.

Le modèle AR(1) est un bon modèle.

Graphique N°19 : Tests sur les résidus du processus AR(1) (test de Bruit Blanc Normal)



Test de normalité des résidus (test de Jarques-Bera) :

H_0 : normalité H_1 : non normalité

$JB = 159,7574 > X^2(2) = 5,99$, alors on rejette H_0 et les résidus ne suivent pas une Loi Normale.

Test de nullité de la moyenne des résidus :

$$H_0 : \bar{e} = 0 \quad H_1 : \bar{e} \neq 0$$

$$\bar{e} = -0,357927$$

$$\sigma = 30,84170$$

$$\text{Racine de } n = 15,06651917$$

$|\bar{e} / (\sigma / \text{Racine de } n)| = -0,174851387 < 1,96$ alors on accepte l'hypothèse de nullité de la moyenne des résidus.

Test d'homoscédasticité des résidus, test ARCH :

$$H_0 : \Phi_1 = \Phi_2 = \dots = \Phi_p = 0 \text{ homoscédasticité}$$

$$H_1 : \text{au moins } 1 \neq 0 \text{ hétéroscédasticité}$$

Nous allons comparer la statistique de Lagrange nR^2 à un $X^2(p)$ ou p est le nombre de retards considéré. Commençons le test avec $p=1$.

Tableau N°21 : Arch. Test

ARCH Test:

| | | | |
|---------------|----------|-------------|----------|
| F-statistic | 1.121530 | Probability | 0.290730 |
| Obs*R-squared | 1.125906 | Probability | 0.288650 |

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 11/29/10 Time: 08:17

Sample (adjusted): 1990M03 2008M12

Included observations: 226 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| C | 874.0386 | 167.6248 | 5.214254 | 0.0000 |
| RESID^2(-1) | 0.070555 | 0.066623 | 1.059023 | 0.2907 |
| R-squared | 0.004982 | Mean dependent var | 941.1568 | |
| Adjusted R-squared | 0.000540 | S.D. dependent var | 2333.524 | |
| S.E. of regression | 2332.894 | Akaike info criterion | 18.35642 | |
| Sum squared resid | 1.22E+09 | Schwarz criterion | 18.38669 | |
| Log likelihood | -2072.275 | F-statistic | 1.121530 | |
| Durbin-Watson stat | 2.016637 | Prob(F-statistic) | 0.290730 | |

$$nR^2 = 227 * 0,004982 = 1,130914$$

$$X^2(1) = 3,84$$

$nR^2 = 1,130914 < X^2(1) = 3,84$, alors on accepte H_0 , les résidus au carrés de notre estimation ne sont pas hétéroscédastiques.

Test de non autocorrélation des résidus

Tableau N° 22

Il existe deux méthodes pour déceler de l'auto-corrélation de nos résidus : la statistique de Durbin-Watson et la comparaison de la Q-stat avec un $X^2(p)$ sur le corrélogramme.

Nous ne mettons en œuvre que le premier.

Test de Durbin-Watson

D'après notre estimation.

DW= 2,026507

DW (n=227, k=1) : $d_1 = 1,73$ et $d_2 = 1,78$

Le DW appartient à l'intervalle $[d_2 ; 4-d_2]$, alors on accepte H_0 , ce qui signifie que nos résidus ne sont pas auto-corrélés.

En conclusion, les résidus de notre processus AR(1) ne passent pas tous les tests du bruit blanc normal.

Les résidus sont non auto-corrélés, de moyenne des résidus nulle, non hétéroscédastiques mais ne suivent pas une loi normale.

Enfin, le modèle estimé est un bon modèle d'après le coefficient de détermination.

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| AR(1) | 0.691265 | 0.047668 | 14.50177 | 0.0000 |
| R-squared | 0.481986 | Mean dependent var | -0.285063 | |
| Adjusted R-squared | 0.481986 | S.D. dependent var | 42.85456 | |
| S.E. of regression | 30.84378 | Akaike info criterion | 9.700143 | |
| Sum squared resid | 215002.6 | Schwarz criterion | 9.715231 | |
| Log likelihood | -1099.966 | Durbin-Watson stat | 2.026507 | |
| Inverted AR Roots | .69 | | | |

Tableau N°23 : Prévisions

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| AR(1) | 0.999640 | 0.001157 | 863.8464 | 0.0000 |
| R-squared | 0.386014 | Mean dependent var | 1924.997 | |
| Adjusted R-squared | 0.386014 | S.D. dependent var | 42.85176 | |
| S.E. of regression | 33.57746 | Akaike info criterion | 9.869983 | |
| Sum squared resid | 254802.8 | Schwarz criterion | 9.885071 | |
| Log likelihood | -1119.243 | Durbin-Watson stat | 2.336220 | |
| Inverted AR Roots | 1.00 | | | |

| | PrevAR(1) | Coef | Résidu | Série | prévision |
|---------|------------|-------|------------|----------|------------|
| 2009M01 | 1834,01135 | 100,2 | 42,6866143 | 2067,105 | 1976,89797 |
| 2009M02 | 1833,35199 | 100 | 130,560232 | 1885,962 | 2063,91222 |
| 2009M03 | 1832,69286 | 99,9 | 70,8443579 | 2081,064 | 1861,7485 |
| 2009M04 | 1832,03397 | 99,9 | 143,780179 | 2142,176 | 1788,15379 |
| 2009M05 | 1831,37531 | 96,6 | 53,4661873 | 2113,554 | 1981,4415 |
| 2009M06 | 1830,7169 | 95,5 | 11,9068371 | 1902,408 | 1914,31006 |
| 2009M07 | 1830,05872 | 100,1 | 1,31639287 | 1799,622 | 1928,84232 |
| 2009M08 | 1829,40077 | 100,2 | 20,9499485 | 1704,503 | 1908,65082 |
| 2009M09 | 1828,74306 | 100,3 | 3,48830171 | 1602,158 | 1925,55476 |
| 2009M10 | 1828,08559 | 100,2 | 26,5482201 | 1793,324 | 1901,73737 |
| 2009M11 | 1827,42836 | 100,1 | -22,745671 | 1798,9 | 1904,78269 |
| 2009M12 | 1826,77136 | 100,1 | 6,61540496 | 1953,643 | 1933,48676 |



7.2. Remarques

La production laitière a eu tendance à baisser ces deux dernières décennies.

Cette diminution est due à différents facteurs :

- La crise à laquelle est confronté l'ensemble du secteur agricole.
- La politique de contrôle mis en place par l'Europe au travers des quotas laitiers.

En ce qui concerne la saisonnalité de la collecte de lait de vache celle-ci est en partie la conséquence de la répartition des vêlages (mise bas d'une vache) sur l'année et aux saisons structurant le climat français.

Pour prévoir cette collecte nous pouvons conclure qu'il nous faut préférer un modèle avec constante et sans tendance, avec une désaisonnalisation par Moyennes Mobiles correspondant à la nature rigide de sa saisonnalité.

En effet nous avons pu remarquer grâce à cette étude que notre série était composée essentiellement d'une constante et de variations saisonnières sensiblement identiques.