Application de la Technique de prévision BOX-JENKINS

**2015**

Gaston

[Tapez le nom de la société]

17/07/2015



**Gastonfils LONZO LUBU**

Doctorant en Sciences Economiques/PTCI

Université de Kinshasa

[gastonfils@hotmail.fr](mailto:gastonfils@hotmail.fr)

+243.81.69.20.221

# APPLICATION DE LA METHODE DE PREVISION DE BOX-JENKINS

# Par

# **Lonzo Lubu Gastonfils**

Doctorant en Sciences Economiques

[gastonfils@outlook.com](mailto:gastonfils@outlook.com)

# Résumé

# L’objectif de ce papier est de permettre aux étudiants en sciences économiques, non seulement de la Faculté des Sciences Economiques et de Gestion (FASEG) de l’Université de Kinshasa mais également ceux qui évoluent dans la cette filière ou recourent souvent à cette technique de prévision, de trouver un précieux repère ou guide pour parvenir avec toute indépendance de traiter leurs données et d’appliquer la technique de prévision de Box-Jenkins sans commettre d’erreurs notamment dans la prévision fore casting comme in casting.

# **Introduction**

La maitrise de l’environnement économique est devenue avec le temps la priorité des acteurs des pouvoirs publics et des entreprises. Ils doivent anticiper les incertitudes liées à l’avenir afin de mieux les aborder qu’elles soient bonnes ou mauvaises. Pour contourner l’incertitude liée au futur à travers des bonnes décisions, sur le plan économique et de gestion, les décideurs recourent à des techniques de prévisions du comportement des variables économiques soit macro ou microéconomiques.

D’où l’impérieuse nécessité d’apprentissage de ces techniques scientifiques des théories et techniques de prévisions dans le cadre des enseignements de deuxième cycle universitaire en Sciences Economiques à l’Université de Kinshasa (UNIKIN).

Dans le cadre de l’examen relatif à ces enseignements, nous allons faire notre application sur la série gdpcauta PIB de l’Australie de la Source : The Conference Board Total Economy Database, January 2011, <http://www.conference-board.org/data/economydatabase/>. La série GDPCAUTA de l’Australie est annuelle et varie de 1950 à 2010.

Outre l’introduction et la conclusion, le canevas de notre étude se résume de la manière suivante :

1. L’évolution graphique des données de la série de 1950 à 2010.
2. Synthèse sur les statistiques descriptives de la série
3. l’identification de la série ;
4. Traitement et choix de la méthode de prévision.

La méthodologie de Box-Jenkins a été conçue pour faire de la prévision en se basant sur l’évolution passée de la variable elle-même. Pour son application il faut déterminer au préalable la nature de la chronique en présence car une série chronologique peut être générée à travers quatre types différents de processus qui sont : les processus AR, MA, ARMA et ARIMA. Le dernier processus peut lui-même être affecté d’une saisonnalité donnant lieu au processus appelé SARIMA. Le choix d’un processus ainsi que la détermination de son ordre constituent la phase la plus importante de cette méthodologie.

L’application de la méthodologie de Box-Jenkins passe par quatre étapes.

1. **Identification**
2. **Estimation**
3. **Validation**
4. **Prévision**

La première étape consiste à identifier la chronique en utilisant des tests informels basés sur l’examen visuel des corrélogrammes. Ceux-ci sont la représentation graphique des différents coefficients d’autocorrélation (simples et partiels). Ils montrent ainsi l’évolution des fonctions d’autocorélation simple et d’autocorrélation partielle. Ce qui permet de déceler si la chronique est AR, MA ou ARMA. En effet, en se basant sur les propriétés d’autocorrélation simple et d’autocorrélation partielle, (Cuthbertson et al.(1992)) donnent des critères indicatifs pouvant aider à détecter la nature du processus générateur d’une chronique. En mettant en parallèle les deux fonctions, ces auteurs synthétisent ces critères tels qu’indiqués au tableau ci-après :

|  |  |  |
| --- | --- | --- |
| **Fonction d’Autocorrelation Simple** | **Processus générateur** | **Fonction d’Autocorrelation Partielle** |
| Un pic pour k=1 | AR(1) | Décroissance exponentielle |
| Décroissance exponentielle | MA(1) | Un pic pour k=1 |
| \*Un pic k=1 suivi d’une décroissance exponentielle | ARMA(1,1) | \*Un pic k=1 suivi d’une  décroissance exponentielle |
| \*Deux pics k=1 et k=2  suivis par une  décroissance  exponentielle | ARMA(2,1) | \*Un pic k=1 suivi d’une  décroissance exponentielle |
| \*Un pic k=1 suivi d’une  décroissance exponentielle | ARMA(1,2) | \*Deux pics k=1 et k=2  suivis par une  décroissance  exponentielle |
| \*Deux pics k=1 et k=2  suivis par une  décroissance  exponentielle | ARMA(2,2) | \*Deux pics k=1 et k=2  suivis par une  décroissance  exponentielle |

Comme principe général on retient que les coefficients d’autocorrélation simple font voir un processus AR et ceux d’autocorrélation partielle font voir un processus MA. Quant à leurs ordres respectifs le processus AR a pour ordrele nombre des coefficients d’autocorrélation partielle non nulsc’est-à-dire significativement différents de zéro tandis que celui du processus MA est donné par le nombre des coefficients d’autocorrélation simple non nuls c’est-à-dire significativement différents de zéro.

Ces critères bien que simples ne peuvent pas être réduits à un niveau trop mécanique car les autocorrélations simples et les autocorrélations partielles n’indiquent pas toujours d’une manière aussi claire un modèle spécifique, elles peuvent également indiquer plus d’un modèle.

Au cours de cette première étape, on vérifie également la stationnarité de la chronique généralement par les tests de Dickey-Fuller. Une chronique non stationnaire peut être affectée soit d’une tendance déterministe (processus TS), c’est-à-dire être un processus stationnaire en tendance ou soit être un processus stationnaire en différence (processus DS).

Une série affectée d’une tendance déterministe est stationnarisée par les MCO tandis qu’une série générée par un processus DS est stationnarisée en prenant les différences.

* La deuxième étape est celle de l’estimation proprement dite laquelle se fait généralement par les MCO.
* La troisième étape se réfère à la vérification du diagnostique.
* Enfin la quatrième étape concerne la prédiction.

La résolution du problème se fait en quatre étapes que voici :

Première étape :

On fait une analyse exploratoire des données en identifiant la série à travers le corrélogramme et en vérifiant la stationnarité avec le test ADF.

Deuxième étape :

On fait l’estimation par la méthode appropriée

Troisième étape :

On fait la vérification du diagnostique

Quatrième étape :

On fait la prévision proprement dite

## I. IDENTIFICATION DE LA SERIE

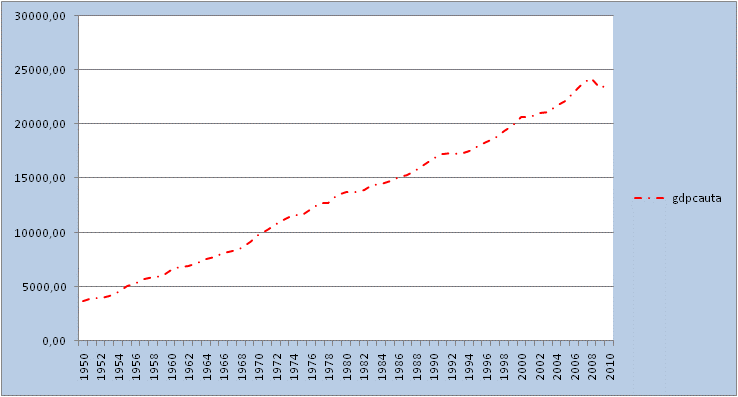
Pour identifier le type de chronique, nous allons procéder à des tests formels et informels afin de connaître la saisonnalité, la tendance et la stationnarité.

## A. LES TESTS INFORMELS

### A.1. L’ANALYSE GRAPHIQUE DE LA SERIE

L’analyse graphique nous permet de visualiser l’évolution temporelle de la série du PIB Australien (GDPCAUTA) de 1950 à 2010. Ainsi, à partir des données de la série, nous obtenons le graphique suivant :

**Figure 1.1. Évolution graphique de GDPCAUTA**



Source : Auteur, sur base des données The Conference Board Total Economy Database, Jan. 2011,

Source : Auteur, TCBTED, 2011

De ce graphique, nous observons que la série GDPCAUTA évolue de manière croissante dans le temps. Ce qui nous permet de tirer quelques conclusions intéressantes au vu de l’allure de cette courbe ci-dessus :

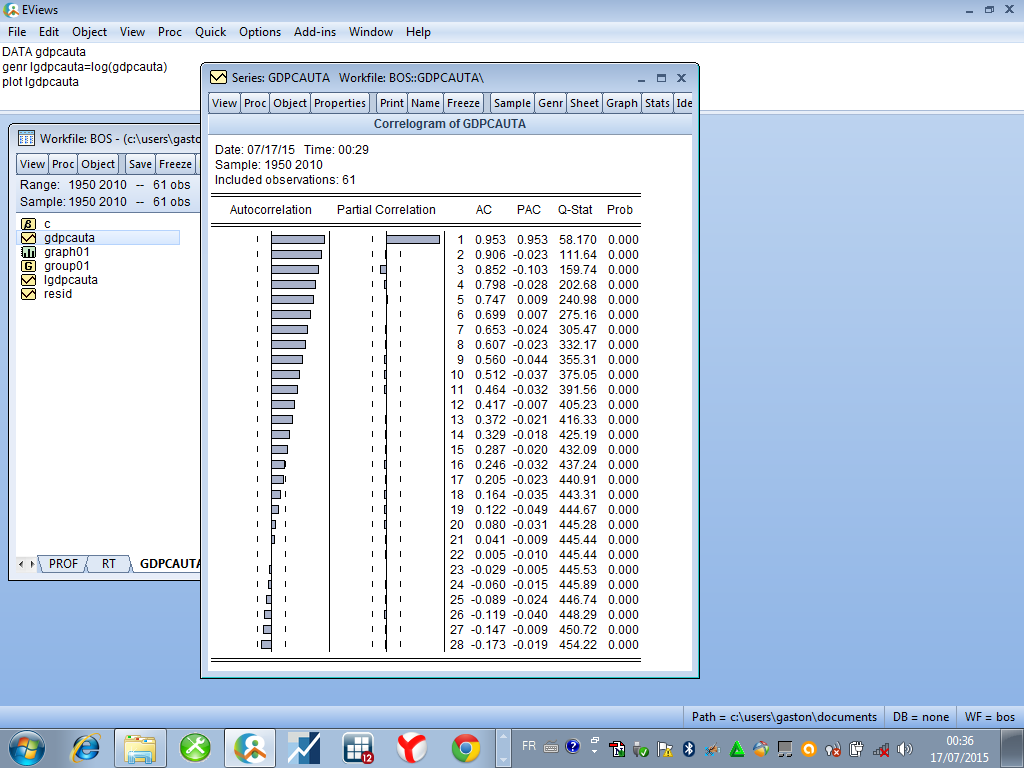
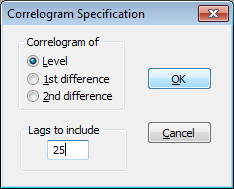
* La présomption de l’absence de saisonnalité du fait qu’il n’y a pas des points de changement de structure avec oscillation dans une série annuelle;
* La présomption d’une tendance de la série ;
* La série n’est pas stationnaire du fait de la présomption d’une faible volatilité autour de sa moyenne.

Les conclusions de l’analyse graphique restent les mêmes sur la série GDPCAUTA transformée en logarithme.

Un dernier test peut alors être effectué, il s’agit de l’analyse de la fonction d’autocorrélation qui vise avant tout à déterminer le caractère significatif des coefficients d’autocorrélation simples et partiels par l’observation du corrélogramme de la série GDPCAUTA:

Correlogram série GDPCAUTA

Commande : **IDENT GDPCAUTA**



Source : Auteur, à l’aide du logiciel Eviews 9.0

Les coefficients d’autocorrélation simples décroissent lentement, de manière linéaire, ce qui appuie notre affirmation sur la non stationnarité de la série.

Nous observons aussi que le premier coefficient d’autocorrélation partielle est significatif.

Seulement ces indicateurs sont nécessaires mais non-suffisants.

Par conséquent, nous utilisons le Test de la racine unitaire de Dickey- Fuller (DF) et Dickey-Fuller Augmenté (ADF) qui sont plus performant dans l’étude de la stationnarité des séries.

**1.1.2. Tests formels : Test de racine unitaire ( DF & ADF).**

Ce test a un double objectif :

- Il permet de vérifier la stationnarité d’une série ;

- Il donne une idée sur la structure de la série.

La formulation générale de celui-ci est :

avec avec

Nous avançons les hypothèses suivantes pour notre test :

Ho : , présence de la racine unitaire càd la série est non stationnaire ;

H1 : , absence de la racine unitaire càd la série est stationnaire ;

Ainsi, nous rencontrons les cas de figures ci-après :

* Si ,on accepte l’hypothèse nulle (A.Ho), la série est non stationnaire[[1]](#footnote-2)
* Si ,on rejette l’hypothèse nulle (R.Ho), la série est non stationnaire

Significativité du Trend

Ho : , le trend est non significatif ;

H1 : , le trend est significatif ;

Ainsi, nous rencontrons les cas de figures ci-après :

* Si , on accepte l’hypothèse nulle (A.Ho)
* Si , on rejette l’hypothèse nulle (R.Ho),

Significativité de l’intercept

Ho : , la constante est non significatif ; modèle sans dérive

H1 : , le constante est significatif ; modèle avec dérive

Ainsi, nous rencontrons les cas de figures ci-après :

* Si , on accepte l’hypothèse nulle (A.Ho)
* Si , on rejette l’hypothèse nulle (R.Ho),

**Tableau 1.5. Test de racine unitaire de DICKEY-FULLER AUGMENTÉ**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Null Hypothesis: GDPCAUTA has a unit root | | | |  |
| Exogenous: Constant, Linear Trend | | | |  |
| Lag Length: 0 (Automatic - based on SIC, maxlag=10) | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  | t-Statistic | Prob.\* |
|  |  |  |  |  |
|  |  |  |  |  |
| Augmented Dickey-Fuller test statistic | | | -3.097627 | 0.1163 |
| Test critical values: | 1% level |  | -4.118444 |  |
|  | 5% level |  | -3.486509 |  |
|  | 10% level |  | -3.171541 |  |
|  |  |  |  |  |
|  |  |  |  |  |
| \*MacKinnon (1996) one-sided p-values. | | | |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Augmented Dickey-Fuller Test Equation | | | |  |
| Dependent Variable: D(GDPCAUTA) | | | |  |
| Method: Least Squares | | |  |  |
| Date: 07/17/15 Time: 00:49 | | |  |  |
| Sample (adjusted): 1951 2010 | | |  |  |
| Included observations: 60 after adjustments | | | |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| GDPCAUTA(-1) | -0.258651 | 0.083500 | -3.097627 | 0.0030 |
| C | 991.3829 | 233.8343 | 4.239681 | 0.0001 |
| @TREND("1950") | 90.28674 | 28.81753 | 3.133049 | 0.0027 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.149342 | Mean dependent var | | 333.7876 |
| Adjusted R-squared | 0.119495 | S.D. dependent var | | 270.2865 |
| S.E. of regression | 253.6239 | Akaike info criterion | | 13.95829 |
| Sum squared resid | 3666531. | Schwarz criterion | | 14.06301 |
| Log likelihood | -415.7487 | Hannan-Quinn criter. | | 13.99925 |
| F-statistic | 5.003491 | Durbin-Watson stat | | 1.708452 |
| Prob(F-statistic) | 0.009954 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Source : Auteur, à l’aide du logiciel Eviews 9.0

Il ressort de cet output du test d’ADF, que la série de GDPCAUTA est non stationnaire (, et le trend est significatif, la constante est significatif.

Ce qui nous amène à mettre fin à l’étude de la stationnarité de notre série et à rejoindre la conclusion des tests préliminaires : notre série présente donc une non stationnarité de type Déterministe (TS : Trend stationnarity).

**STATIONNARISATION DE LA SERIE GDPCAUTA**

Etant donné que la série que la série GDPCAUTA présente donc une non stationnarité de type Déterministe (TS : Trend stationnarity).

Nous allons la stationnariser par l’extraction du trend grâce à la méthode d’écart à la tendance. En d’autres termes, la stationnarisation de GDPCAUTA se fera en deux étapes suivantes :

* Estimer par MCO : avec =le résidu
* Générer le résidu:

Lorsque nous sommes en présence d’une série non stationnaire de type TS, c’est la tendance déterministe qui perturbe la série. Ainsi pour stationnariser une variable TS il faut appliquer l’écart à la tendance.

Supposons une variable MT, elle est dite non stationnaire si l’on constate qu’avoir effectué le test de Dickey Fuller : . Le coefficient est significatif.

Les trois étapes à suivre sont les suivantes :

- **GENR T = @TREND(t-1)** : générer le temps, si l’année sous étude comme à 1970 taper 1969 [ex : GENR T = @TREND(1969)]

- **LS GDPCAUTA C T** : estimer le modèle qui prend en compte la tendance

- **GENR WT = RESID** : la commande qui nous permet d’élaguer la tendance

L’élimination de la tendance dans la variable residu peut être observée par l’estimation suivante :

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: GDPCAUTA | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 07/17/15 Time: 01:04 | | |  |  |
| Sample: 1950 2010 | | |  |  |
| Included observations: 61 | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 3028.558 | 100.0376 | 30.27420 | 0.0000 |
| T | 344.4427 | 2.875878 | 119.7696 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.995904 | Mean dependent var | | 13361.84 |
| Adjusted R-squared | 0.995834 | S.D. dependent var | | 6127.430 |
| S.E. of regression | 395.4724 | Akaike info criterion | | 14.83028 |
| Sum squared resid | 9227508. | Schwarz criterion | | 14.89949 |
| Log likelihood | -450.3234 | Hannan-Quinn criter. | | 14.85740 |
| F-statistic | 14344.75 | Durbin-Watson stat | | 0.467845 |
| Prob(F-statistic) | 0.000000 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Nous constatons que le coefficient de tendance est significatif. D’où nous allons extraire le temps qui est significatif dans la série GDPCAUTA.



Nous constatons que le coefficient de la tendance n’est plus significativement différent de 0. De plus, notre variable WT n’est pas encore stationnaire comme nous l’indiquent les résultats suivants :

Test de Stationnarité sur la série extrait du trend

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Null Hypothesis: WT has a unit root | | | |  |
| Exogenous: Constant, Linear Trend | | | |  |
| Lag Length: 0 (Automatic - based on SIC, maxlag=10) | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  | t-Statistic | Prob.\* |
|  |  |  |  |  |
|  |  |  |  |  |
| Augmented Dickey-Fuller test statistic | | | -3.097627 | 0.1163 |
| Test critical values: | 1% level |  | -4.118444 |  |
|  | 5% level |  | -3.486509 |  |
|  | 10% level |  | -3.171541 |  |
|  |  |  |  |  |
|  |  |  |  |  |
| \*MacKinnon (1996) one-sided p-values. | | | |  |
|  |  |  |  |  |
| Augmented Dickey-Fuller Test Equation | | | |  |
| Dependent Variable: D(WT) | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 07/17/15 Time: 01:09 | | |  |  |
| Sample (adjusted): 1951 2010 | | |  |  |
| Included observations: 60 after adjustments | | | |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| WT(-1) | -0.258651 | 0.083500 | -3.097627 | 0.0030 |
| C | -47.30847 | 66.31273 | -0.713415 | 0.4785 |
| @TREND("1950") | 1.196350 | 1.890672 | 0.632765 | 0.5294 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.149342 | Mean dependent var | | -10.65507 |
| Adjusted R-squared | 0.119495 | S.D. dependent var | | 270.2865 |
| S.E. of regression | 253.6239 | Akaike info criterion | | 13.95829 |
| Sum squared resid | 3666531. | Schwarz criterion | | 14.06301 |
| Log likelihood | -415.7487 | Hannan-Quinn criter. | | 13.99925 |
| F-statistic | 5.003491 | Durbin-Watson stat | | 1.708452 |
| Prob(F-statistic) | 0.009954 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Ainsi, nous allons stationnariser cette série au moyen de la méthode de différence première

**GENR DWT=WT-WT(-1) ou GENR DWT=D(WT).**

Ce qui conduit donc à la série DWT

Test d’ADF

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Null Hypothesis: D(WT) has a unit root | | | |  |
| Exogenous: None | | |  |  |
| Lag Length: 0 (Automatic - based on SIC, maxlag=10) | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  | t-Statistic | Prob.\* |
|  |  |  |  |  |
|  |  |  |  |  |
| Augmented Dickey-Fuller test statistic | | | -7.138024 | 0.0000 |
| Test critical values: | 1% level |  | -2.604746 |  |
|  | 5% level |  | -1.946447 |  |
|  | 10% level |  | -1.613238 |  |
|  |  |  |  |  |
|  |  |  |  |  |
| \*MacKinnon (1996) one-sided p-values. | | | |  |

Source: Auteur, calcul àl’aide du logiciel Eviews

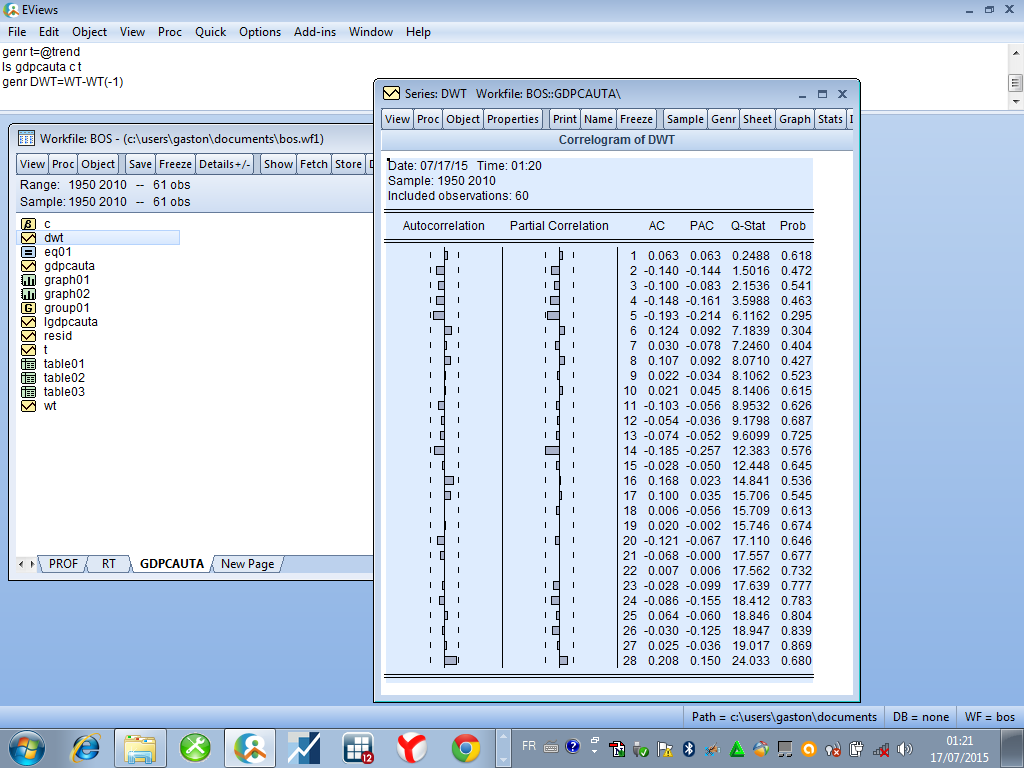
La représentation graphique de la série stationnarisée est donnée sur la figure ci-dessous.



Source : Auteur, Eviews 9.0

Identification de la série stationnaire

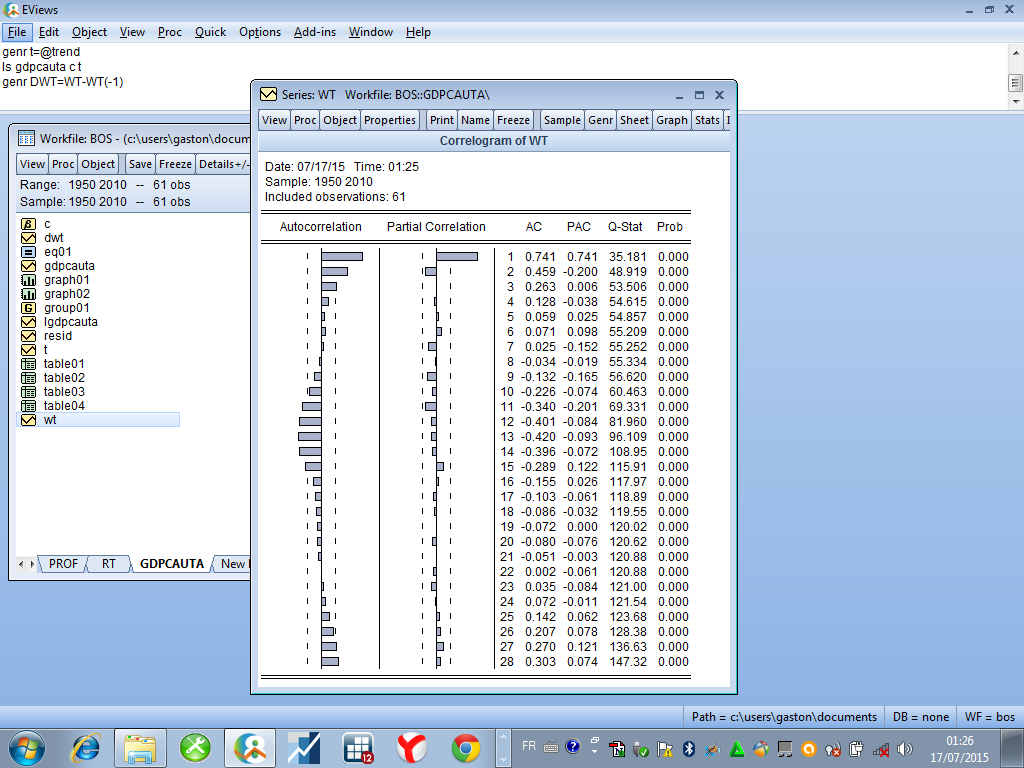
Commande : Ident DWT



Source : Auteur, Eviews 9.0

**IDENTIFICATION DU MODELE OPTIMAL PROVISOIRE**

Cette phase consiste à déterminer les modèles adéquats dans la famille ARIMA, SARIMA. Elle est fondée sur l’étude des corrélogrammes simples et partiels de la série étudiée (éventuellement stationnarisée). Ainsi, observons le corrélogramme de WT



Source : Auteur, Eviews 9.0

Nous constatons que les corrélogrammes partiels ont leur premier terme différent de 0. Nous pouvons donc anticiper un processus de type ARIMA(1,0,0). Ceci ne pourra être confirmer qu’en identifiant le résidu de l’estimation de l’équation :

  ou

**3) ESTIMATION.**

Après avoir identifié le modèle, il convient de passer à la phase d’estimation des paramètres de ce modèle. Elle consiste à trouver une chronique permettant de transformer la variable observée et d’obtenir un ou plusieurs estimateurs. En estimant par les MCO notre modèle, nous avons :

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: WT | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 07/17/15 Time: 01:30 | | |  |  |
| Sample (adjusted): 1951 2010 | | |  |  |
| Included observations: 60 after adjustments | | | |  |
| Convergence achieved after 4 iterations | | | |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -41.80762 | 126.5526 | -0.330358 | 0.7423 |
| AR(1) | 0.741198 | 0.083067 | 8.922937 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.578546 | Mean dependent var | | -11.29195 |
| Adjusted R-squared | 0.571280 | S.D. dependent var | | 385.3422 |
| S.E. of regression | 252.3095 | Akaike info criterion | | 13.93196 |
| Sum squared resid | 3692286. | Schwarz criterion | | 14.00177 |
| Log likelihood | -415.9587 | Hannan-Quinn criter. | | 13.95926 |
| F-statistic | 79.61881 | Durbin-Watson stat | | 1.696273 |
| Prob(F-statistic) | 0.000000 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Inverted AR Roots | .74 | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Le modèle AR(1) estimé est

  ou   avec

**t-stat (8,9229)**

le coefficient est statistiquement significatif. La constante étant non significatif, il ne sera pas pris en compte dans le modèle final.

1. **VALIDATION DU MODELE**

Le modèle ne peut être mis à défaut. Ces tests dans Eviews sont dans **Views/ Residual Tests**. Nous allons diagnostiquer notre modèle à partir des tests suivants :

**Test de portemanteau/Test de bruit blanc(Q-Q plot simple)**

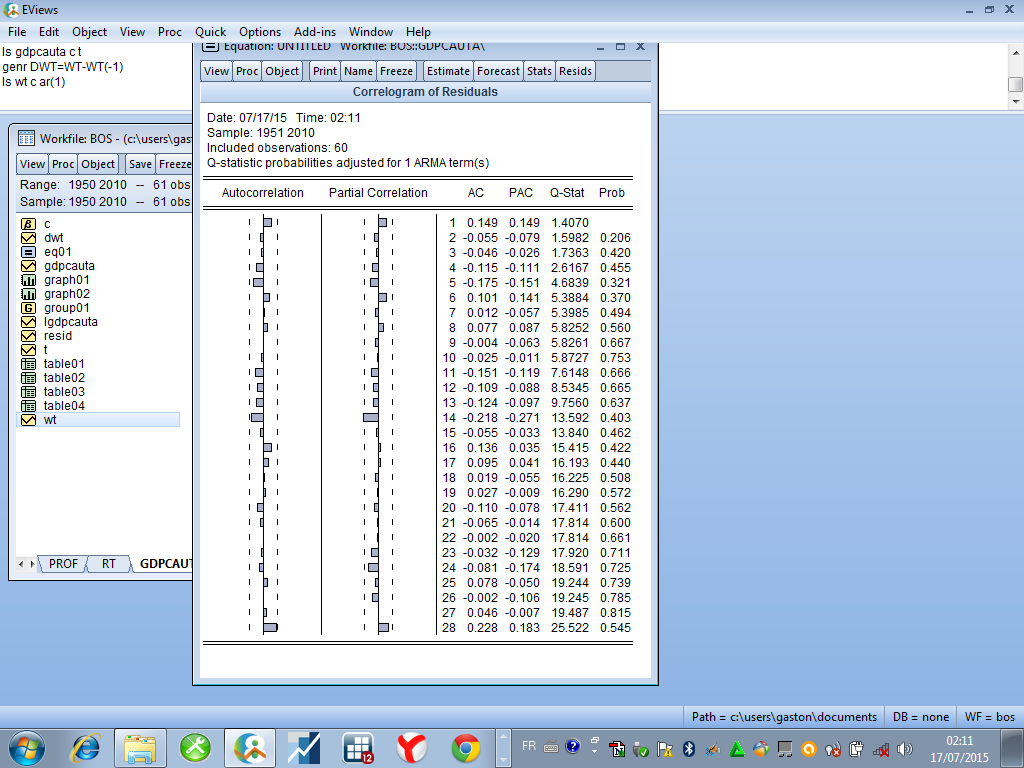
Une des premières validations du modèle consiste à vérifier que le résidu issu de l’estimation est un bruit blanc. Cela équivaut à faire le Test de bruit blanc normal.

Les hypothèses de ce test sont :

, la série des résidus est un bruit blanc

Au moins un , la série des résidus n’est pas un bruit blanc

Il est basé sur l’observation des corrélogrammes et sur le test de BOX-PIERCE & LJUNG-BOX. Ainsi en visualisant les corrélogrammes du résidu, nous avons :



Source : Auteur, Eviews 9.0

Les hypothèses du test sont :

H0 : les résidus sont des bruits blancs (prob-kième lag>5%)

H1 : les résidus ne sont pas des bruits blancs (prob-k ième lag<5%)

Les Prob (Q-stat) n’étant pas inférieures à 5%, nous acceptons l’hypothèse nulle. Les residus du modèle estimé sont des bruits blancs. Par conséquent, le modèle ARIMA(1,0,0) estimé est celui qui a généré notre variable(GDPCAUTA). Il est donc opportun de choisir un modèle AR dont le décalage est égal à 1.

**Condition de stationnarité**

Soit notre modèle théorique : [1]. Tester les conditions de stationnarité sur le modèle [1] estimé revient à vérifier que les racines caractéristiques du polynôme de retard associé à ce modèle – soit le polynôme : – sont toutes supérieures à 1 en valeur absolue (càd : ) et que le paramètre « » estimé soit, en valeur absolue, inférieur à l’unité (compris dans le cercle unité du plan complexe. Càd ). Ainsi, écrivons notre polynôme retard d’ordre 1 (degré du lag optimal : ) et calculons la racine caractéristique associée « » comme suit : avec :

Le modèle estimé respecte les conditions de stationnarité. Le modèle est convergent, et la série GDPCAUTA est stationnaire.

**Test de Student des paramètres (significativité statistique)**

Ce test consiste à vérifier que les paramètres du modèle qui ont été estimés sont statistiquement différents de 0, en supposant que les estimateurs sont normalement distribués.

Les hypothèses du test sont :

, le coefficient est non significatif

, le coefficient est significatif

Ainsi, au risque de 5%, tous le paramètre du modèle est statistiquement différent de Zéro car ils sont en valeur absolue supérieurs à 2.

Le paramètre «» estimé est statistiquement significatif (prob-t student<5%|tc|>|tt|).

**TEST DE LINÉARITÉ : View/Stability Test/Ramsey RESET Test…→Fitted : 2→ok.**

Ce test consiste à vérifier la bonté de la relation présentée par le modèle. On va recourir à la statistique de Fisher. Les hypothèses du test sont :

, mauvaise spécification

Au moins un , Bonne spécification

coefficient d’autocorrélation

Généralement, lorsque l’hypothèse nulle du test de student est rejetée, nous pouvons accepter l’hypothèse alternative pour le test de linéarité c’est-à-dire le modèle est globalement bon.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Ramsey RESET Test | | |  |  |
| Equation: UNTITLED | | |  |  |
| Specification: WT AR(1) | | |  |  |
| Omitted Variables: Powers of fitted values from 2 to 3 | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
|  | Value | df | Probability |  |
| F-statistic | 0.737476 | (2, 57) | 0.4828 |  |
| Likelihood ratio | 1.532833 | 2 | 0.4647 |  |
|  |  |  |  |  |
|  |  |  |  |  |
| F-test summary: | | |  |  |
|  | Sum of Sq. | df | Mean Squares |  |
| Test SSR | 93310.10 | 2 | 46655.05 |  |
| Restricted SSR | 3699310. | 59 | 62700.18 |  |
| Unrestricted SSR | 3606000. | 57 | 63263.16 |  |
| Unrestricted SSR | 3606000. | 57 | 63263.16 |  |
|  |  |  |  |  |
|  |  |  |  |  |
| LR test summary: | | |  |  |
|  | Value | df |  |  |
| Restricted LogL | -416.0157 | 59 |  |  |
| Unrestricted LogL | -415.2493 | 57 |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Avec la Prob/F-stat 0,4828 >0, on accepte H0. Le modèle respecte l’hypothèse de linéarité.

### Test d’autocorrélation des erreurs : VIEW/RESIDUAL TESTS/ SERIAL CORRELATION LM TEST

Le test de Breusch-Godfrey s’utilise pour des processus autorégressifs d’ordre 2 .

Les hypothèses du test sont :

H0 : ρ= 0 : absence d’autocorrélation

H1 : ρ ≠ 0 : présence d’autocorrélation

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Breusch-Godfrey Serial Correlation LM Test: | | | |  |
|  |  |  |  |  |
| F-statistic | 1.625714 | Probability | | 0.201779 |
| Obs\*R-squared | 1.417629 | Probability | | 0.492227 |
|  |  |  |  |  |

La statistique chi2 (obs\*R-squared) a une probabilité qui est supérieure au seuil de significativité de 5% nous rejetons l’hypothèse de recherche. Donc il y a absence d’autocorrélation des erreurs.

### Test d’hétéroscédasticité : view/residual test /arch lm test

Nous avons choisit le test d’Hétéroscédasticité conditionnelle autorégressive, en sigle ARCH. Les hypothèses du test sont :

H0 : ρ1= ρ2 =… = ρk= 0 : homoscédasticité

H1 : ρ1≠ ρ2≠…..≠ ρk≠ 0 : hétéroscédasticité

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Heteroskedasticity Test: ARCH | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| F-statistic | 8.840199 | Prob. F(1,105) | | 0.0037 |
| Obs\*R-squared | 8.309027 | Prob. Chi-Square(1) | | 0.0039 |
|  |  |  |  |  |

La probabilité associée à la statistique chi2 (obs\*R-squared) est inférieure au seuil de significativité de 5%, nous rejetons l’hypothèse nulle donc le modèle estimé est hétéroscéstastique.

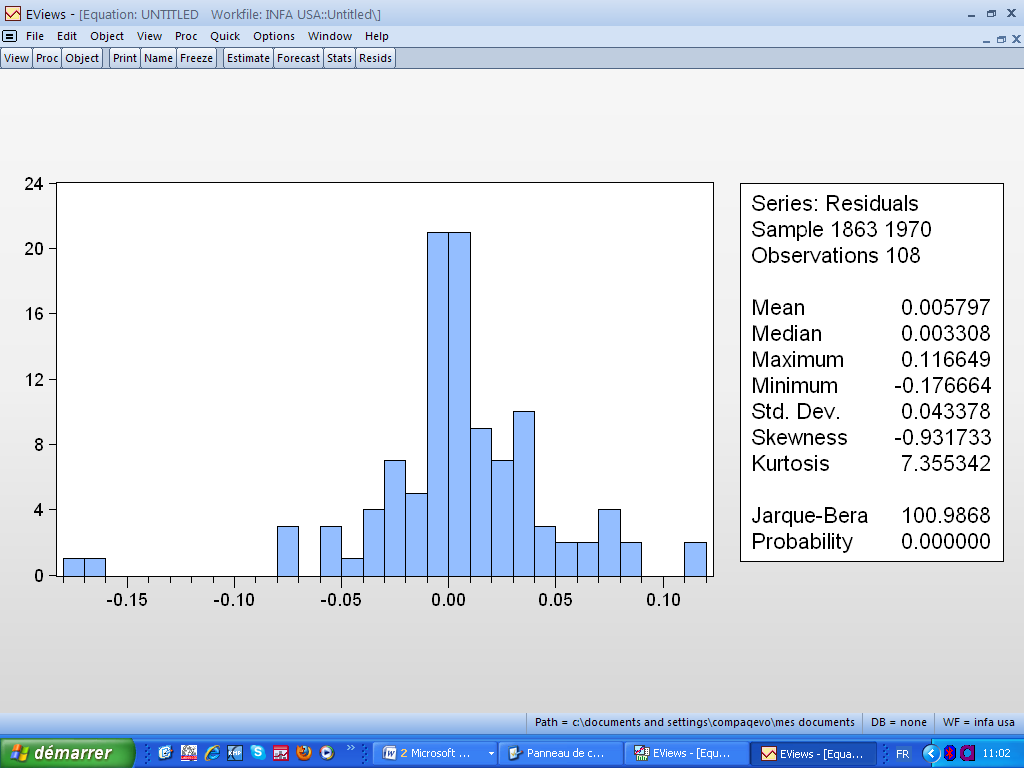
### Test de normalité des erreurs : VIEW/ RESIDUAL TESTS/ HISTOGRAM – NORMALITY TEST

Le test que nous utilisons est celui de Jarque et Bera. Il stipule que les erreurs suivent une distribution normale de moyenne nulle et de variance constante.

Ce test se formule avec les hypothèses suivantes :

H0: les résidus sont normalement distribués

H1: les résidus ne sont par normalement distribués



Les résidus de ce modèle ne sont pas normalement distribués. Ils suivent un processus bruit blanc non gaussien.

**PERFORMANCE DE LA PREVISION.**

Nous la lisons en comparant les valeurs prévues calculées aux données réelles dans l’échantillon (l’on parle mieux de la « prévision in casting »).

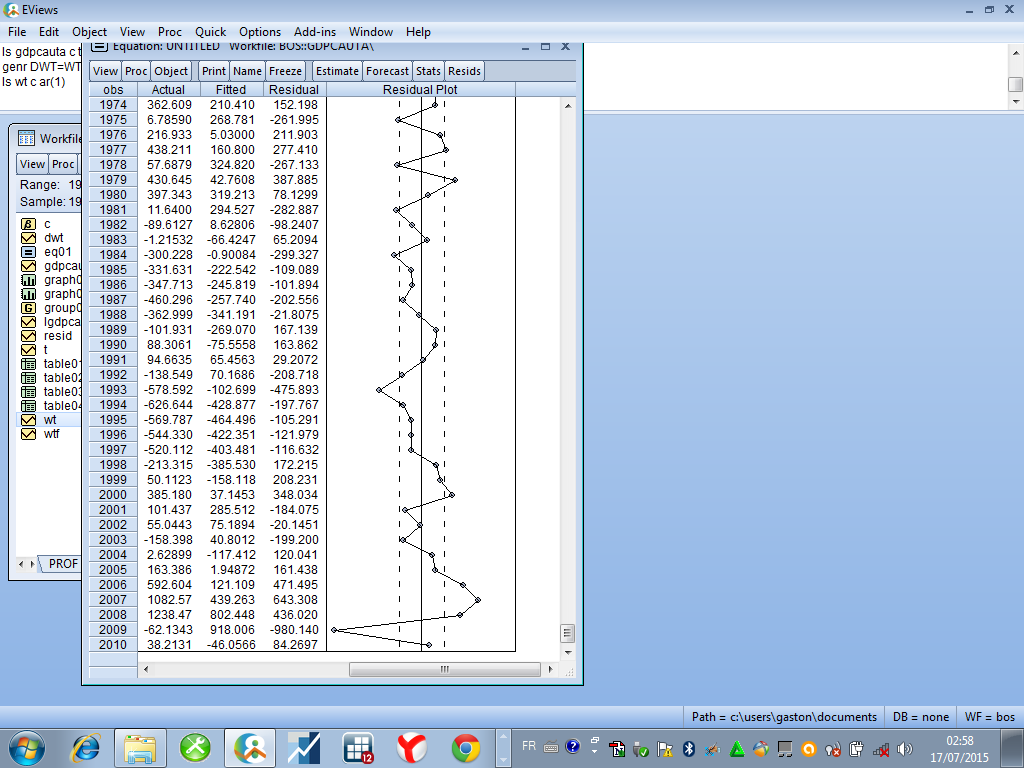


L’ajustement de la prévision in casting est bien fait car la série prédite est plus proche de la série actuelle de GDPCAUTA.

La visualisation des graphiques de la prévision dans l’intervalle de temps de la période sous étude permet de vérifier la performance de la prévision. Partant de l’équation suivante :

, nous pouvons calculés les valeurs prévues et les comparer aux valeurs actuelles en vue de dégager les erreurs de prévision dans l’intervalle de temps considéré

Prévision dans l’echantillon (Prévision in casting) *View/Actual,Fitted,Resid Table*



**PREVISION FORECASTING**

Calculs préliminaires

avec

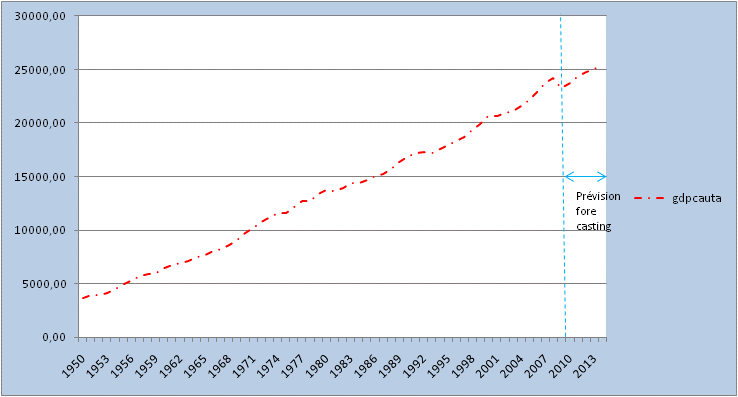
Prévision (Ponctuelle pour h=4 années)

Pour h=1 (année 2011)

🡺  **s**

**Cet encadré est une feuille EXCEL active, vous pouvez double cliquer pour obtenir les formules**





Série AA

Q1.





Question 2.



Question 3



Série AA

Q1.





Question 2.



Question 3



Théorie et Pratique de Prévision

Introduction

Une première information sur des erreurs de prévision nous est fournie par le diagramme des réalisations et des prévisions. Chaque point du diagramme représente le couple réalisation ‑ prévision. Traçons une droite à 45 degrés passant par l'origine. Cette droite est, par construction, le lieu des points où Yt = Pt. C'est la droite des prévisions parfaites. Chaque point situé sur cette droite représente un cas de ***prévision parfaite***, alors que chaque point situé en dehors représente une prévision erronée.

La dispersion des points autour de la droite à 45 degrés caractérise la prévision. Si cette dispersion est faible, les prévisions seront dans leur ensemble précises, elles seront imprécises dans le cas contraire. Une mesure de la dispersion autour de cette droite nous est donnée par l'**erreur quadratique moyenne** E(Yi - Pi)2. Il est à noter que ce critère ne fait aucune discrimination entre le signe des erreurs. On considère que le coût de l'erreur est identique quel que soit son signe. Considérons le point des coordonnées [E(Yi), E(Pi). La prévision sera considérée comme non biaisée si ce point se trouve sur la droite à 45. Dans le cas contraire E(Yi) différera de E(Pi). La différence E(Yi) - E(Pi) = E(ui) nous donne le biais de la prévision. ***D'après le signe du biais, les prévisions seront surestimées (signe négatif) ou sous-estimées (signe positif).***

Il existe en fait un grand nombre d'indicateurs statistiques permettant d'évaluer l'exactitude d'une prévision. Ainsi, pour déterminer la justesse d'une prévision, on pourrait ajouter les erreurs de plusieurs périodes afin d'obtenir une indication de l**'erreur moyenne** (le rapport entre la somme des erreurs prévisionnelles et le nombre de périodes de l'horizon temporel).

L’erreur moyenne est définie par :

 (1.1)

Cet indicateur n'est pas vraiment utile, la somme des erreurs s'annulant du fait de la compensation des erreurs positives par les erreurs négatives. Cet indicateur signale la présence ou l’apparition d’un biais systématique : prévision en moyenne trop forte ou trop faible. On peut donc apprécier le centrage statistique du modèle : un modèle correct avec variations aléatoires donnera une valeur nulle de e.

Puisque des termes de signes contraires, même importants, peuvent se compenser au moins partiellement pour donner une valeur de e qui semble acceptable, on définit l’écart moyen absolu (que nous noterons MAD, pour Mean Absolute Deviation, en anglais) donné parla formule :

MAD =  (1.2)

qui évite ces compensations et contrôle l’écart entre observation et prévision.

**Remarque : MAD et écart type**

La MAD est simple à calculer. Elle est souvent utilisée à la place de l’écart type. Il faut connaître la correspondance facile à retenir 3σ= 4MAD (ce qui correspond à un filtre à 99,7 % pour une loi normale, c’est-à-dire où le risque d’accepter une valeur à rejeter est inférieure à 0,3 %). On pourra utiliser cette grandeur pour évaluer les stocks de sécurité permettant de couvrir l’incertitude de la prévision.

**MAD lissée**

On préfère parfois calculer la MAD lissée (lissage exponentiel) :

MADi = b│yi– pi│– (1 – b) MADi-1  (1.3)

On choisit un coefficient b petit (par exemple 0,1), ce qui assure un lissage à long terme de la MAD

**Qualité du modèle de prévision**

L’observation simultanée de e et MAD permet d’avoir une bonne idée de la qualité du modèle de prévision. Afin de maîtriser un système de prévision de nombreux articles, il faut mettre en place des fourchettes pour ces indicateurs. Le suivi de ces indicateurs et de leur comportement nous alertera d’une quelconque modification et nous permettra de réagir.

Un autre indicateur utilisé pour prévenir d’un processus de prévision qui devient hors contrôle est le signal d’alerte Ai suivant :

Ai = (1.4)

Cette valeur peut naturellement être positive ou négative, mais doit rester dans des limites raisonnables et non biaisées (systématiquement négative ou positive). D’une manière analogue à un contrôle statistique de la qualité où l’on souhaite une valeur dans une fourchette de plus ou moins trois écarts types, si le signal d’alerte Ai dépasse quatre en valeur absolue (car 3σ= MAD), on soupçonnera un changement dans la demande. Il nous restera à en rechercher les causes et à modifier le modèle.

Il existe également d’autres indicateurs tels que la **moyenne des carrés des erreurs** appelée aussi l**'erreur carrée moyenne** (**Mean Square Error, MSE**) ou la racine carrée de cette dernière (**Root Mean Square Error, RMSE**) ou encore **la racine carrée de l'erreur carrée moyenne en pourcentage** (**Root Mean square Percentage Error, RMSPE**). On peut également chercher ***l'erreur absolue moyenne en pourcentage*** (**Mean Absolute Percent Error, MAPE**).

Désignons les prévisions par Pt, les réalisations par Yt et la moyenne des erreurs par0. En négligeant l'indice t, l'erreur carrée moyenne (MSE) pour n périodes sera définie par :

MSE =  1.5

ou

MSE =  1.6

MAPE =  1.7

RMSPE =  1.8

RMSE = = 1.9

Autant l'erreur calculée est minimale, autant la prévision sera exacte.

Un autre indicateur statistique fréquemment utilisé est le **coefficient d'inégalité de Theil** basé sur une comparaison des écarts prévisionnels avec les écarts observés. Il est défini par :

U= 1.10

où le numérateur est le RMSE tandis que le dénominateur est la somme des racines carrées des moyennes des carrées des prévisions et des réalisations. L'avantage de cet indicateur est que sa valeur est comprise entre 0 et 1. Si toutes les prévisions sont correctes, la valeur de U sera égale à 0. Dans ce cas, les prévisions se confondent avec les réalisations et Pt - Yt = 0. Si U = 1, la prévision est très mauvaise.

**Tableau 4. La *Prévision* de la consommation *intérieure de* ciment en *utilisant* la méthode**

**des moyennes mobiles simples**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **(1)** | **(2)** | **(3)** | **(4)** | **(5)** |
| **Années** | **Périodes** | **Tonnes/**  **1000** | **MM3** | **MM5** |
| 1990 | 1 | 417,6 | - | - |
| 1991 | 2 | 256,1 | - | - |
| 1992 | 3 | 198,0 | - | - |
| 1993 | 4 | 142,0 | 290,6 | - |
| 1994 | 5 | 149,0 | 198,7 | - |
| 1995 | 6 | 197,0 | 163,0 | 232,5 |
| 1996 | 7 | 231,0 | 162,7 | 188,4 |
| 1997 | 8 | 128,4 | 192,3 | 183,4 |
| 1998 | 9 | 138,0 | 185,5 | 169,5 |
| 1999 | 10 | 150,1 | 165,8 | 168,7 |
| 2000 | 11 | 135,5 | 138,8 | 168,9 |
| 2001 | 12 | 137,7 | 141,2 | 156,6 |
| 2002 | 13 | 176,8 | 141,1 | 137,9 |
| 2003 | 14 | - | 150,0 | 147,6 |

La colonne 4 du tableau 4 se rapporte aux moyennes mobiles d’ordre trois calculées à partir des valeurs de la période courante et de celles de deux périodes antérieures. Par exemple, au temps **t** 3, la valeur 290,6 est la moyenne de la série entre les périodes 1, 2 et 3, c’est-à-dire, (417,6 + *256,1* + 198,0)/3. Cette valeur moyenne a été utilisée en vue de prévoir la consommation de la période 4 (l’année 1993). La dernière figure de cette colonne (150,0) est en fait la moyenne des valeurs des périodes il, 12 et 13. C’est cette valeur qui prédit la production de la période 13 (l’année 1992). La colonne 5 reprend les moyennes mobiles sur cinq ans. Ainsi, la valeur 232,5 est la moyenne mobile des valeurs observées aux temps 1, 2, 3, 4 et *5.* Cette moyenne sert de prévision pour la période 6. Pour prévoir la consommation de l’année 2003, on a retenu la moyenne des valeurs des périodes 9 à 13.

Au vu de la figure 1, il peut être noté que l’effet de lissage sur les données de la série augmente avec le nombre d’observations incluses dans le calcul des moyennes.

Une autre manière plus concise et plus claire de calculer la moyenne mobile est donnée par

**Pt + 1**= –  (2.7)

Dans cette dernière expression, chaque nouvelle prévision est tout simplement un ajustement de la prévision précédente (Pt). Cet ajustement incorpore les observations les plus récentes Yt, mais laisse tomber la valeur la moins récente (Yt - N). Ainsi, l’effet de lissage s’accroît avec l’augmentation du nombre N des séries à traiter.

**2.1.2.1.2 La moyenne mobile pondérée**

Leprincipe de base consiste à calculer la tendance sous forme d’une moyenne pondérée d’un certain nombre d’observations successives précédant la période courante **t.** Etant donné que les informations contenues dans la série chronologique sont d’allure complexe et imprécise, il peut être nécessaire de les pondérer en donnant, par exemple, plus de poids aux informations récentes, représentatives de la situation actuelle, qu’aux informations anciennes.

En pratique, cette pondération consiste à multiplier chaque valeur de la série par un coefficient représentant le pourcentage d’information pour lequel cette valeur contribue à la définition de la moyenne. La moyenne pour la période t+1 sera ainsi donnée par :

 (2.8)

ou encore :

 (2.9)

dans laquelle b0, b1, ..., bi, ... seront les coefficients de pondération respectifs aux informations Yt, Yt-1, ... Yi, ... avec la contrainte que la somme de termes de ces coefficients (puisqu’ils représentent des pourcentages) soit égale à 1.

**Tableau 5 : Prévision de la consommation du ciment en** RDC **par la technique de la   
 moyenne mobile pondérée**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | **Coefficients de pondération** | | |
|  |  |  | **b0 = 0,6** | **b0 = 0,8** | **b0 = 0,85** |
|  |  |  | **b1 = 0,3** | **b1 = 0,15** | **b1 = 0,10** |
|  |  |  | **b2 = 0,1** | **b2 = 0,05** | **b2 = 0,05** |
| Années | Périodes | Tonnes /1000 |  |  |  |
| 1990 | 1 | 417,6 | - | - | - |
| 1991 | 2 | 256,1 | - | - | - |
| 1992 | 3 | 198,0 | - | - | - |
| 1993 | 4 | 142,0 | 237,4 | 217,7 | 214,8 |
| 1994 | 5 | 149,0 | 170,2 | 156,1 | 153,3 |
| 1995 | 6 | 197,0 | 151,8 | 150,4 | 150,8 |
| 1996 | 7 | 231,0 | 177,1 | 187,1 | 189,5 |
| 1997 | 8 | 128,4 | 212,6 | 221,8 | 223,5 |
| 1998 | 9 | 138,0 | 166,0 | 147,2 | 142,1 |
| 1999 | 10 | 150,1 | 144,4 | 141,2 | 141,7 |
| 2000 | 11 | 135,5 | 144,3 | 147,2 | 147,8 |
| 2001 | 12 | 137,7 | 140,1 | 137,8 | 137,1 |
| 2002 | 13 | 176,8 | 138,3 | 138,0 | 138,1 |
| 2003 | 14 | - | 160,9 | 168,9 | 170,8 |

Le tableau 5 ci-haut nous donne une idée des résultats obtenus par la méthode de la moyenne pondérée. Trois pondérations ont été prises en compte, la troisième est plus forte que les deux premières. Dans la deuxième, la moyenne est formée de 80 % de l’information contenue dans la dernière valeur Yt, 15 % de l’information de la valeur immédiatement antérieure (c’est-à-dire, de la période t-1) et seulement de 5 % de l’information contenue dans la valeur de la période t-2.

**Avantages et inconvénients :**

Le problème qui se pose en pratique consiste à déterminer la structure des coefficients ***bi*** puisque de cette structure dépend le contenu informationnel de la moyenne. Etant donné qu’il y a une infinité de combinaisons possibles, il en résulte une infinité de moyennes possibles, toutes issues de la même série chronologique ; d’où, difficulté de stockage et d’interprétation des coefficients de pondération

La méthode possède néanmoins des avantages évidents sur celle de la moyenne mobile simple. D’abord, la moyenne calculée ici est valable pour la dernière période, c’est-à-dire, celle correspondant à la dernière information de la série. Il n’y a donc plus de phénomène de retard comme dans le cas de la moyenne mobile simple. Ensuite, en fonction de la structure fixée pour ces coefficients, il est possible de rendre la moyenne plus ou moins sensible aux changements de tendance.

**2.1.2.1.3 Les Moyennes mobiles linéaires** ou **Moyennes mobiles doubles**

La méthode des moyennes mobiles linéaires a été développée afin d’éviter la sous-estimation des valeurs actuelles, sous-estimation qui survient lorsque la moyenne mobile simple est appliquée aux données présentant une tendance linéaire.

La méthode des moyennes mobiles linéaires (ou doubles) calcule au départ un jeu de moyennes mobiles simples, exactement comme nous l’avons fait précédemment, et calcule ensuite une seconde moyenne mobile basée sur les valeurs de la première moyenne mobile simple.

Si S’t est la moyenne mobile simple, S”t, la moyenne mobile double, Pt + m la prévision pour m périodes à l’avance, et b, l’ajustement qu’on fait subir à la prévision, la prévision fournie par la moyenne mobile linéaire peut s’obtenir mathématiquement de la façon suivante :

 (2.10)

 (2.11)

at = 2S’t – S’’t (2.12)

 (2.13)

Pt + m = at + btm (2.14)

La relation (2.13) est un terme correctif pour rendre la méthode plus précise.

Le tableau 6 contient les résultats de l’application des moyennes mobiles doubles (avec N = 4 et m = 1) à un exemple de prévision de la consommation trimestrielle de ciment en RDC. La colonne *5* donne les résultats qu’on obtient lorsqu’on ajoute à la colonne 3 la différence entre les colonnes 3 et 4 ; la colonne 6 calcule un terme correctif qui rend la méthode encore plus précise.

**Tableau 6 : Prévision de la consommation de ciment en RDC par la technique des   
 moyennes mobiles linéaires (doubles)**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** | **(8)** | **(9)** |
| Années | Trimestres | Périodes | Observé | S′t = MM4 | S″t = MM4 | at | bt | bt+1 = at + btm |
|  | 1 | 1 | 29444 | - | - | - | - | - |
|  | 2 | 2 | 40244 | - | - | - | - | - |
|  | 3 | 3 | 28420 | - | - | - | - | - |
| 2000 | 4 | 4 | 34013 | 33030,3 | - | - | - | - |
|  | 1 | 5 | 17267 | 29986,0 | - | - | - | - |
|  | 2 | 6 | 35127 | 28706,8 | - | - | - | - |
|  | 3 | 7 | 37662 | 31017,3 | 30685,1 | 31349,4 | 221,5 | - |
| 2001 | 4 | 8 | 37871 | 3191,8 | 30422,9 | 33540,6 | 1039,2 | 31570,9 |
|  | 1 | 9 | 40208 | 37717,0 | 32355,7 | 43078,3 | 3574,2 | 34579,9 |
|  | 2 | 10 | 42304 | 39511,3 | 35056,8 | 43965,7 | 2969,6 | 46652,5 |
|  | 3 | 11 | 50536 | 42729,8 | 37984,9 | 47474,6 | 3163,2 | 46935,3 |
| 2002 | 4 | 12 | 38484 | 42883,0 | 40710,3 | 45055,8 | 1448,5 | 50637,8 |
|  | 1 | 13 | 50622 | 45486,5 | 42652,6 | 48320,4 | 1889,3 | 46504,3 |
|  | 2 | 14 | 59735 | 49844,3 | 45235,9 | 54452,6 | 3072,3 | 50209,6 |
|  | 3 | 15 | 73145 | 55496,5 | 48427,6 | 62565,4 | 4712,6 | 57524,9 |
| 2003 | 4 | 16 | 64166 | 61917,0 | 53186,1 | 70647,9 | 5820,6 | 67278,1 |
|  | 1 | 17 | 68216 | 66315,5 | 58393,3 | 74237,7 | 5281,5 | 76468,6 |
| 2004 | 2 | 18 | - | - | - | - | - | 79519,1 |

Finalement, à la colonne 7, on présente la prévision elle-même. La prévision est simplement la somme de la colonne *5* et de la colonne 6 multipliée par le nombre de périodes à projeter. Pour le prévisionniste effectuant sa prévision une période à l’avance, la prévision est simplement la somme des colonnes *5* et 6. Pour deux périodes à l’avance, la prévision serait donnée par la colonne *5* plus deux fois la colonne 6, et ainsi de suite.

En utilisant cet ensemble d’équations pour préparer une prévision de la période 18, sachant que l’on est à la période 17, on obtient :

**Pt + 1 = P18 = a17 + b17 (1)**

= [2(66315,5) – 58393,3] + (5281,5) = 74237,7 + 5281,5 = 79519,1

Pour que le prévisionniste puisse appliquer cette méthode de prévision, il faut qu’il ait à sa disposition 2N-1 relevés de données, ou pratiquement deux fois le nombre exigé par le lissage simple. Il est clair que cette nécessité d’un stockage de données substantiel rend la méthode des moyennes mobiles doubles moins attrayantes que le lissage exponentiel double décrit à la section suivante.

**2.1.2.2 Le lissage exponentiel**

Nous avons vu que les moyennes mobiles calculées précédemment avaient deux inconvénients majeurs. Primo, pour les obtenir, il faut disposer de N observations sur le passé. Or, dans le cas où le nombre des séries chronologiques devenait important, ceci poserait un problème de stockage de l’information. Secundo, un même poids est donné à chacune des N observations antérieures (c’est-à-dire, l’estimation est linéaire), et aucun poids n’est donné aux observations antérieures à la période (t - N).

Avec la moyenne mobile, un second filtre linéaire est couramment utilisé pour la décomposition des chroniques : il s’agit du **lissage exponentiel** qui est une technique introduite par Holt en 1957, mais surtout par Brown en 1962 et utilisée dans le cas d’une chronique affectée d’une tendance aléatoire.

Le lissage regroupe l’ensemble des techniques empiriques qui ont pour caractéristiques communes d’accorder un poids plus important aux valeurs récentes de la série chronologique. On donne ainsi à chaque observation un poids d’autant plus faible qu’elle est ancienne et ces poids décroissent exponentiellement dans le passé. Ces méthodes portent aussi le nom de **filtrage**, car il s’agit d’une opération mathématique transformant un entrant xt en une nouvelle chronique sortante yt.

Nous étudierons plusieurs types de lissage exponentiel à savoir :

* le **lissage exponentiel simple** qui s’applique lorsque la série chronologique est sans tendance ni saisonnalité ;
* le **lissage exponentiel double ou linéaire ;**
* le **lissage exponentiel quadratique.**

***2.1.2.2.1 Le lissage exponentiel simple (L.E.S.) ou Lissage de Brown***

**Le lissage exponentiel simple** *(single exponential smoothing), développé*par **R.G. BROWN**,est un cas particulier de la moyenne mobile pondérée. Cette technique, qui ne peut pas être employée en présence d’une tendance ou d’une saisonnalité, est basée sur l’idée que les observations récentes contiennent plus d’information sur le futur que les anciennes. Par conséquent, les observations récentes devront être affectées d’un poids plus important que les anciennes.

La technique du lissage exponentiel peut être aisément dérivée de la relation (2.7) des moyennes mobiles simples que nous reprenons ci-après :

**Pt + 1**= –  (2.7)

Dans le cas où Yt – N n’est pas disponible, on peut supposer que sa valeur peut être approximativement égale à la valeur prédite de la période précédente, Pt. En fait, si les données sont stationnaires, cette approximation est plutôt bonne. Nous pouvons donc écrire (2.7) comme :

**Pt + 1**= -  (2.15)

**Pt + 1**=  (2.16)

Il est clair qu’une telle prévision est basée sur la pondération de l’information plus récente par un poids plus élevé et celle de la prévision la plus récente par un poids de   
valeur . Puisque N est un nombre positif,  devra se situer entre 0 (si N = ∞)et 1 (si N = 1). En posant α = , l’équation (2.16) devient :

**Pt + 1**=  (2.17)

avec 0 ≤ α ≤ 1 où

Yt = valeur de la série chronologique à la période t ;

Pt + 1 = prévision pour la période t+1;

Pt = valeur prédite à la période t ;

α = coefficient de pondération appelé aussi “**constante de lissage**”. Il est généralement compris entre 0 et 1. Une valeur recommandée par Brown, l’auteur de la méthode, est α = 0,3.

Si α = 0, on a Pt + 1 = Pt, de sorte que les observations n’interviennent pas dans la prévision.

Si α = 1, on a Pt + 1 = Yt, ce qui signifie que la prévision est égale à la dernière valeur observée, et on retrouve la formule de la méthode de prévision naïve n°1.

C’est la formule générale utilisée dans le calcul des prévisions par la méthode du lissage exponentiel. Cette formule réduit sensiblement le problème de stockage, étant donné qu’il n’est plus nécessaire de conserver toutes les données historiques (comme dans le cas de la moyenne) ou une partie d’entre elles (comme dans le cas de la moyenne mobile). Mieux, seules l’observation récente, la prévision récente et une valeur de **α** ont besoin d’être stockées.

Avant d’appliquer la formule (2.17), il est nécessaire de disposer au préalable de la valeur de α et d’une prévision initiale (P1). Bien que le choix de la constante **α** constitue une question cruciale pour l’opération de lissage, il n’existe aucun critère absolu pour déterminer la meilleure constante α, cette dernière dépendant de l’objectif poursuivi par le prévisionniste. La prévision, par exemple, peut nécessiter une sensibilité assez élevée aux dernières observations.

Généralement, la valeur de α peut être obtenue par tâtonnements. On essaie un certain nombre de valeurs de α telles que 0,1, 0,3, *0,5,* 0,7 et 0,9 et on conserve celle qui minimise **l’écart- type** **de l’erreur quadratique moyenne** (RMSE, Root Mean Square Error) ou toute autre mesure d’erreur prévisionnelle, **l’erreur quadratique moyenne** (MSE, Mean Square Error) notamment. Ce dernier critère est le plus largement utilisé.

Le tableau 7 illustre l’application du lissage exponentiel simple à la série se rapportant à la production du pétrole brut en RDC pour la période allant de janvier à décembre 2003. Trois différentes valeurs de α (α = 0,1, α = *0,5* et α = 0,9) ont été utilisées dans les calculs. Dans la colonne *(5),* la prévision pour le mois de février est supposée identique à la production réalisée en janvier (669123).

Pour prévoir la demande du mois de janvier 2004, nous utilisons la formule suivante:

**Pjanvier 2004** = αY12 + **(1** –α **)P12**

**= 0,1(801468) + 0,9(740449,9) = 80146,8 + 666404,91 = 746651,7**

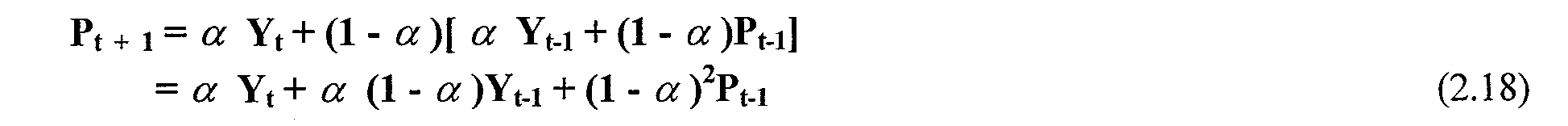
Quant au choix de la prévision initiale P1, soulignons que ce choix a d’autant plus d’importance que α est proche de zéro et que la série chronologique est courte. On peut par exemple prendre P1 = Y1 ou la moyenne de N dernières observations. Toutes les autres possibilités de choix sont détaillées au paragraphe 2.1.2.3.6 consacré aux valeurs d’origine dans les méthodes de lissage exponentiel.

Les prévisions pour α = *0,5* et α = 0,9 sont obtenues de la même manière.

Dans ce tableau, l’erreur quadratique moyenne a été calculée pour les trois constantes de lissage α = 0,1, α ***0,5*** et α = 0,9 et on a obtenu respectivement *7442,465, 3514,743* et 3 *552,646.* Ainsi, α = *0,5* serait la meilleure constante de lissage.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Tableau 7 : Prévision de la production du pétrole brut par la**  **technique de lissage exponentiel simple** | | | | | | |
| Mois | Années | Périodes | Production observée | Valeurs lissées exponentiellement | | |
|  |  |  |  | α = 0,1 | α = 0,5 | α = 0,9 |
| **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** |
| Janvier | 2003 | 1 | 669,123 | - | - | - |
| Février | 2003 | 2 | 638,963 | 669,123 | 669,123 | 669,123 |
| Mars | 2003 | 3 | 803,996 | 666,107 | 654043 | 641,979 |
| Avril | 2003 | 4 | 765,073 | 679,895 | 729,019 | 787,794 |
| Mai | 2003 | 5 | 839,131 | 688,413 | 747,046 | 767,345 |
| Juin | 2003 | 6 | 798,341 | 703,485 | 793,088 | 831,952 |
| Juillet | 2003 | 7 | 808,675 | 712,970 | 795,714 | 801,702 |
| Août | 2003 | 8 | 775,990 | 722,541 | 802,194 | 807,977 |
| Septembre | 2003 | 9 | 736,753 | 727,886 | 789,092 | 779,188 |
| Octobre | 2003 | 10 | 774,116 | 728,772 | 762,922 | 740,996 |
| Novembre | 2003 | 11 | 804,734 | 733,307 | 768,519 | 770,804 |
| Décembre | 2003 | 12 | 801,468 | 740,449 | 786,626 | 801,341 |
| Janvier | 2004 | 13 | - | 746,551 | 794,047 | 801,455 |
|  | | | MSE | 7442,465 | 3514,743 | 3552,646 |

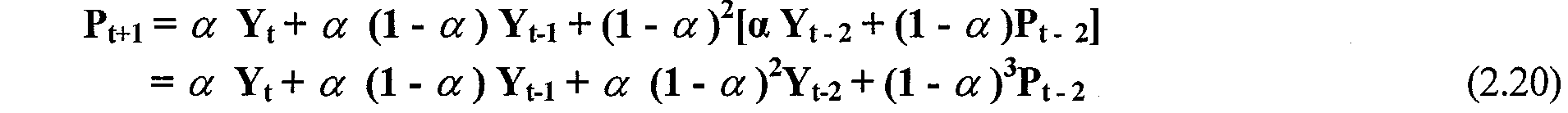
Nous pouvons aussi exprimer la prévision en fonction des données passées. Pour ce faire, on substitue dans la relation (2.17) le terme Pt par sa valeur et obtenir :



D’une manière analogue, puisque à la période t-1, on a :

**Pt + 1**=  (2.19)

On peut porter cette dernière dans (2.18) et l’on obtient :



Remplaçons à présent P2 par sa valeur, et, en continuant ce processus indéfiniment, nous arrivons au résultat suivant :

**Pt + 1**=   (2.21)

Nous voyons que la fonction Pt + **1** est une combinaison linéaire de toutes les observations passées. Les poids donnés aux observations antérieures décroissent exponentiellement avec l’ancienneté de l’information, d’où le nom de *lissage exponentiel.*

De plus, l’influence des termes décroît rapidement et en particulier, le dernier terme a une influence négligeable si **α** n’est pas trop voisin de zéro et si la série dont on dispose est suffisamment longue (N suffisamment grand). *α* grand signifie que la prépondérance est accordée au premier terme de l’équation (2.21) et que le passé s’efface très vite. Au contraire, α petit signifie que la prépondérance s’étale sur un grand nombre d’observations.

De tout ce qui précède, nous pouvons déduire les constatations suivantes :

1. Le développement de la formule (2.21) peut encore s’écrire :

Pt + 1 = b0 Yt + b1 Y t – 1 + b2Yt – 2 + … + biYt – i + … (2.22)

avec :

b0 = α , b1 = , b2 = , … bi  = …

Comme nous le voyons, c’est bien là une moyenne mobile pondérée dont la particularité est que les coefficients décroissent exponentiellement en fonction de l’âge ou de l’ancienneté de l’information. L’appellation de « **lissage exponentiel** » se justifie donc.

Une fois fixé, le coefficient de pondération ou de lissage a l’avantage de déterminer automatiquement toute la structure des coefficients bi, (cfr tableau 8) : bi = .

**Tableau 8 : La structure de pondération α (1 - α)i suivant diverses valeurs de α**.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **I** | **α = 0,1** | **α = 0,2** | **α = 0,3** | **α = 0,4** | **α = 0,5** |
| 0 | 0,1 | 0,2 | 0,3 | 0,4 | 0,5 |
| 1 | 0,09 | 0,16 | 0,21 | 0,24 | 0,25 |
| 2 | 0,061 | 0,128 | 0,147 | 0,144 | 0,125 |
| 3 | 0,073 | 0,1024 | 0,1029 | 0,0864 | 0,062 |
| 4 | 0,066 | 0,0819 | 0,0720 | 0,0518 | 0,031 |
| *5* | 0,059 | 0,0655 | 0,0504 | 0,0311 | 0,016 |
| *6* | 0,053 | 0,0524 | 0,0352 | 0,0186 | 0,008 |
| *7* | 0,048 | 0,0419 | 0,0247 | 0,0111 | 0,004 |

Liés par une fonction exponentielle, les coefficients de pondération décroissent très rapidement au fur et à mesure que l’âge de l’information est grand. Ainsi, par exemple, pour une valeur de α = 0,4 (tableau 8) :

- la dernière valeur connue de la série Yt est pondérée avec b0 = 0,4 (40 % de sa valeur contribue à former la moyenne cherchée à la période t).

- l’avant dernière valeur connue Yt - 1 est pondérée avec b1 = 0,24 (24 % de son volume contribue à former la moyenne à la période t).

- les informations de cette série vieilles de 8 périodes ou plus (Yt - 8, Yt - 9, etc.) sont pondérées avec des coefficients b8, b9, ..., inférieurs à 0,01. Leur contenu informationnel pour le calcul de la moyenne de la période t peut être considéré comme négligeable.

**2**. La condition que la somme de ces coefficients bi est égale à 1 est ici satisfaite. En effet,

****

est une série géométrique. En remplaçant  par a et (1- ) par q, on obtient :

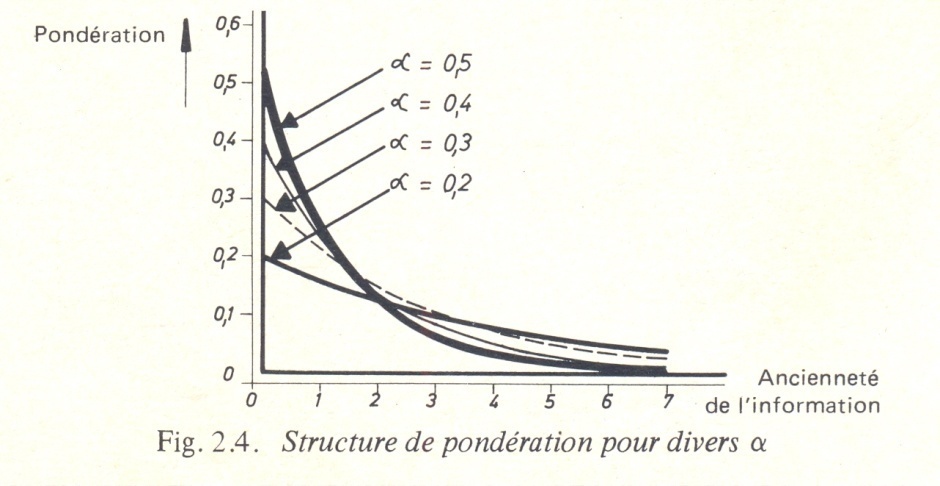
**,**

série géométrique dont la somme est :

****

Ce qui vérifie la condition **.**

1. Comme on vient de le voir, le fait de fixer le coefficient de lissage définit toute une structure de pondération des informations vieilles telle que le montre la figure ci-après :



**4.** Enfin, d’après la formule (2.17), il n’y a que deux additions et deux multiplications. Donc, le calcul est simple. Etant donné que toute l’information ancienne se trouve concentrée dans le terme Pt, on n’a pas besoin de conserver en mémoire toute l’historique passée pour effectuer le calcul de la moyenne ; une seule valeur suffit Pt. Cette économie de stockage de l’information est d’une importance capitale pour l’analyse d’un grand nombre de séries.

A partir de la formule (2.17), nous pouvons écrire :

**Pt + 1**=  0 ≤ α ≤ 1 (2.17)

= Pt + α (Yt – Pt)

ou simplement :

**Pt +1 = Pt + α еt** (2.23)

Nous voyons que la nouvelle prévision apparaît comme l’ancienne prévision corrigée car la prévision pour la période t+1 est calculée en ajoutant à la prévision de la période t un certain pourcentage **α** de la différence entre la valeur actuelle à la période t et la prévision de l’instant t.

C’est cette différence que nous appelons « **erreur prévisionnelle** » constatée à la période t.

Du point de vue statistique, **α** est la vitesse à laquelle une nouvelle prévision s’ajuste à l’erreur. L’ampleur de la correction est proportionnelle à la dernière erreur de prévision. C’est ce qu’on appelle « ***la présentation de mise à jour par l’erreur du lissage exponentiel simple*** ».

D’après cette formule, ***la prévision fournie par le lissage exponentiel est simplement égale à l’ancienne prévision plus un ajustement à cause de l’erreur constatée dans la dernière prévision***.

Sous cette forme, il est clair que si **α** est proche de 1, on ajoute à l’ancienne prévision Pt la totalité de l’erreur prévisionnelle ***et = Yt*** – ***Pt***. Dans ce cas, on considère que cette erreur provient du fait que la moyenne a augmenté de ***et***. Inversement, si **α** = 0, il n’y aura pas d’ajustement. On considère alors que l’erreur de prévision ***et*** est purement aléatoire et n’est pas la conséquence d’un changement possible de la moyenne.

Le rôle de la constante de lissage est prépondérant dans le processus de lissage car c’est de sa valeur que dépendent la stabilité et le taux de réponse de la série lissée. Ces deux caractéristiques sont complémentaires : la série lissée sera d’autant plus stable qu’elle est moins sensible aux fluctuations de la série observée, et le taux de réponse sera d’autant plus élevé que le modèle est plus sensible à la dernière observation. Il est donc clair que si l’on désire un taux de réponse élevé de la part de la série lissée, le poids accordé à la dernière observation doit être relativement fort, ce qui implique le choix d’une valeur élevée pour **α**. Si au contraire, on désire que la série lissée soit très stable, on choisira une valeur relativement faible pour la constante de lissage.

D’autre part, l’équation (2.21) nous indique que les différentes observations reçoivent un poids décroissant en fonction de leur ancienneté. Il est aisé de comprendre que plus le coefficient **α** sera faible, plus grand sera le nombre d’observations qui interviendront réellement dans l’estimation de Pt + 1 et par conséquent, plus accusé sera le lissage.

**2.1.2.2.2 Le lissage exponentiel simple avec taux de réponse adaptatif**

Le lissage exponentiel simple adaptatif a unavantage sur le lissage exponentiel simple: c’est qu’il n’exige aucune valeur spécifique de **α**. Cette caractéristique est un atout particulièrement intéressant lorsque la prévision porte sur des centaines de milliers d’articles. En plus, cette méthode peut modifier la valeur de **α** lorsque cette dernière n’est plus appropriée à cause de changements intervenant dans la série.

La technique **ARRSES** (Adaptive Response Rate Single Exponential Smoothing) est adaptative dans ce sens que la valeur de **α** changera automatiquement en cas d’un changement dans le modèle de base, lequel changement requiert une différente valeur de **α**. La relation de base pour la prévision par la méthode ARRSES est identique à celle de l’équation (2.17) sauf que **α** est remplacé par **α** t. Nous l’écrirons comme :

 (2.24)

où

[[2]](#footnote-3) (2.25)

 (2.26)

 (2.27)

***et = Yt***  ***Pt (2.28)***

= 0,2.

 désigne la valeur absolue. *et* est l’erreur lissée et Mt l’erreur lissée absolue

Le tableau 9 ci-dessous illustre la prévision de la production du pétrole brut en RDC par la méthode de lissage exponentiel simple avec taux adaptatif (ARRSES) avec une valeur initiale ***α*** = 0,5.

La prévision pour la période 11, par exemple, est donnée par :

P11 = 

= 0,005 (774,1) + 0,995(767,9) = 768,0

e11 = 804,7 – 768 = 36,769 par la formule (2.28)

E11 = 0,2(36,8) + 0,8(6,2) = 8,24 par la formule (2 .26)

M11 = 0,2(36,8) + 0,8(30,25) = 31,55 par la formule (2 .27)

= = 0,26 par la formule (2 .25)

De même, la prévision pour la période 13 est :

Pjanvier 2004 = 

= 0,26(801,5) + 0,74(769,3) = 208,35 + 569,282 = 777,67

**Tableau 9 : Prévision de la production du pétrole brut par la technique de lissage   
 exponentiel simple avec taux de réponse adaptatif [[3]](#footnote-4)**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Mois** | **Années** | **Périodes** | **Production observée** | **Prévision Pt** | **Erreur et** | **Erreur lissée Et**  **avec**  **β = 0,2** | **Erreur lissée absolue Mt** | **Valeur de *α* t** |
|  |  |
| Janvier | 2003 | 1 | 669,1 |  |  |  |  |  |
| Février | 2003 | 2 | 639,0 | 669,1 | -30,1 | -6,03 | 6,03 | 0,5 |
| Mars | 2003 | 3 | 804,0 | 654,0 | 150,0 | 25,17 | 34,82 | 0,5 |
| Avril | 2003 | 4 | 765,1 | 729,0 | 36,1 | 27,35 | 35,07 | 0,7 |
| Mai | 2003 | 5 | 839,1 | 755,1 | 84,0 | 38,69 | 44,86 | 0,8 |
| Juin | 2003 | 6 | 798,3 | 820,6 | -22,3 | 26,49 | 40,35 | 0,9 |
| Juillet | 2003 | 7 | 808,7 | 801,4 | 7,3 | 22,65 | 33,73 | 0,7 |
| Août | 2003 | 8 | 776,0 | 806,2 | -30,2 | 12,08 | 33,03 | 0,7 |
| Septembre | 2003 | 9 | 736,8 | 785,9 | -49,2 | -0,17 | 36,26 | 0,4 |
| Octobre | 2003 | 10 | 774,1 | 767,9 | 6,2 | 1,10 | 30,25 | 0,005 |
| Novembre | 2003 | 11 | 804,7 | 768,0 | 36,8 | 8,24 | 31,55 | 0,04 |
| Décembre | 2003 | 12 | 801,5 | 769,3 | 32,2 | 13,02 | 31,67 | 0,26 |
| Janvier | 2004 | 13 | - | 777,7 |  |  |  | 0,41 |

**2.1.2.2.3 Le lissage exponentiel linéaire (double)**

Nous avons dit précédemment que la méthode des moyennes mobiles comportait deuxlimitations principales. D’abord, pour calculer une prévision, les N dernières données doivent être disponibles et stockées. Ensuite, les moyennes mobiles accordent un poids égal à chacune de ces observations et un poids nul aux observations précédentes. Les deux limitations s’appliquent également à la méthode des moyennes mobiles doubles. En revanche, le lissage exponentiel double peut faire le même travail que les moyennes mobiles doubles sans souffrir de ces deux limitations.

En fait, dans le lissage exponentiel double, on a besoin de ne stocker que trois données et une valeur de α. Cette approche donne aussi des poids décroissants aux observations passées. C’est pour cette raison qu’on la préfère aux moyennes mobiles doubles.

**2.1.2.2.3.1 Le lissage exponentiel linéaire de Brown**

Le concept de base implicitement contenu dans le lissage exponentiel linéaire (double) de Brown est tout à fait identique à celui des moyennes mobiles doubles. L’application du lissage exponentiel simple à une série chronologique comportant une loi de tendance donne des résultats systématiquement inférieurs à la tendance. Une seconde application du procédé à ces valeurs lissées produit de nouveau des valeurs inférieures à la tendance modifiée. Comme dans les moyennes mobiles doubles, nous pouvons ajouter à la valeur résultant du lissage exponentiel simple la différence entre elle-même et le lissage double, puis ajuster pour tenir compte de la tendance.

Les notations mathématiques utilisées dans le lissage exponentiel double de Brown sont :

**S’t**=  (2.29)

**S’’t**=  (2.30)

at = 2S’t – S’’t (2.31)

**bt**=  =  (2.32)

Pt + m = at + btm (2.33)

où

S’t est la valeur lissée exponentielle simple à la période t

S”t est la valeur doublement lissée de la série

α est la constante du lissage exponentiel

m est le décalage de la prévision dans le futur, exprimé en nombre de périodes (l’horizon).

La valeur de la constante α est choisie de façon à minimiser l’EQM (MSE). Toute tendance linéaire dans les données sera lissée et extrapolée.

Avec les données du tableau 6 et une valeur de α = 0,7, nous pouvons utiliser les relations (2.29) à (2.33) pour prévoir la consommation intérieure de ciment en février 2004 pour m = 1.

Tous les résultats sont consignés dans le tableau 10.

P18 = a17 + b17

= (2S’17 – S”17) + 

= [2(67001,1) – 66744,7] + (0,7/0,3)(67001,1 – 66744,7] = 67257,3 + 598,1 = 67855,4

Il est utile de rappeler que les formules (2.29) et (2.30) nécessitent d’initialiser les valeurs de S’t-1 et S’’t-1 à la période t = 1. C’est ce que nous avons fait en posant S’1 = S”1 = Y1.

**Tableau 10 : Prévision de la consommation de ciment en RDC par la technique de lissage**

**exponentiel linéaire de Brown**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Années** | **Trimestres** | **Périodes** | **Observé** | **S’t**  **(α = 0,7)** | **S”t** | **at=2S’t -S”t** | **bt** | **Pt + m = at + btm** |
| 2000 | 1 | 1 | 29444 | 29444,0 | 29444,0 | 29444,0 | 0,0 |  |
| 2000 | 2 | 2 | 40244 | 37004,0 | 34736,0 | 39272,0 | 5292,0 |  |
| 2000 | 3 | 3 | 28420 | 31967,2 | 32797,8 | 31136,6 | -1938,2 | 44564,0 |
| 2000 | 4 | 4 | 34013 | 32335,1 | 32473,9 | 32196,3 | -323,9 | 29198,4 |
| 2001 | 1 | 5 | 17267 | 22290,8 | 25345,7 | 19235,9 | -7128,2 | 31872,4 |
| 2001 | 2 | 6 | 35127 | 29769,0 | 28442,0 | 31096,0 | 3096,3 | 12107,7 |
| 2001 | 3 | 7 | 37662 | 36901,5 | 34363,7 | 39439,3 | 5921,6 | 34192,3 |
| 2001 | 4 | 8 | 37871 | 37808,3 | 36774,9 | 38841,7 | 2411,3 | 45361,0 |
| 2002 | 1 | 9 | 40208 | 39506,9 | 38687,3 | 40326,5 | 1912,4 | 41252,9 |
| 2002 | 2 | 10 | 42304 | 41675,2 | 40778,8 | 42571,6 | 2091,5 | 42238,9 |
| 2002 | 3 | 11 | 50536 | 48066,4 | 45880,1 | 50252,7 | 5101,3 | 44663,1 |
| 2002 | 4 | 12 | 38484 | 42099,6 | 43233,8 | 40965,4 | -2646,4 | 55354,0 |
| 2003 | 1 | 13 | 50622 | 46980,6 | 45856,5 | 48104,7 | 2622,8 | 38319,1 |
| 2003 | 2 | 14 | 59735 | 57001,1 | 53657,7 | 60344,5 | 7801,2 | 50727,4 |
| 2003 | 3 | 15 | 73145 | 69122,0 | 64482,7 | 73761,3 | 10825,0 | 68145,7 |
| 2003 | 4 | 16 | 64166 | 66859,7 | 66146,6 | 67572,8 | 1663,9 | 84586,3 |
| 2004 | 1 | 17 | 68216 | 67001,0 | 66744,7 | 67257,3 | 598,1 | 69236,7 |
| 2004 | 2 | 18 | -- |  |  |  |  | 67855,4 |

**2.1.2.2.3.2 Le lissage exponentiel linéaire de Holt**

Connu également sous le nom de **« méthode de Holt-Winters, le lissage exponentiel linéaire de Holt est appliqué au cas d’une série avec tendance (ou des évolutions de tendance), mais sans saisonnalité.**

Cette méthode de lissage linéaire est semblable, en principe, à celle de Brown, sauf qu’elle ne fait pas intervenir la formule du lissage double. Elle lisse directement les valeurs du trend.

La prévision fournie par le lissage exponentiel linéaire de Holt est obtenue en utilisant deux constantes de lissage ***α*** et ***β***(avec 0 < ***α***, ***β <***1) et les trois équations ci-après :

St =  (2.34)

Tt =  (2.35)

Pt + m = St + Ttm (2.36)

avec :   
St = niveau de la tendance défini comme une moyenne pondérée de la dernière   
 observation Yt et de la prévision réalisée en temps t  l pour la période t ;   
Tt = pente de la tendance en période t. Elle est une moyenne pondérée de la nouvelle pente   
 (St St - 1)et de l’ancienne pente Tt -1

***α =***  constante de lissage correspondant au niveau de la série;

***β*** *=* constante de lissage associée à la tendance;

m = horizon supposé égal à 1.

L’équation (2.34) ajuste directement St pour tenir compte de la tendance de la période précédente, Tt - 1, en l’ajoutant à la dernière valeur lissée, St - 1. Cet ajustement élimine le décalage et porte St à une base approximative de la valeur des données actuelles.

La relation *(2.35)* actualise le trend qui est exprimé ici comme la différence entre deux valeurs successives du lissage exponentiel. Ce qui est normal puisque, lorsque les données présentent un trend, les nouvelles valeurs pourraient être plus grandes ou plus faibles que les précédentes. Cette procédure est identique à celle décrite dans le cas du lissage exponentiel de Brown où l’on a soustrait les valeurs du lissage exponentiel double des valeurs du lissage simple.

Pour effectuer la prévision, le trend est multiplié par le nombre de périodes antérieures pour lesquelles on désire cette prévision. Le résultat est ensuite ajouté à St (le niveau des données réelles qui sont lissées afin d’éliminer les fluctuations aléatoires).

**Tableau 11 : Prévision de la consommation de ciment en RDC par la technique de lissage   
 exponentiel linéaire de Holt**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Années** | **Trimestres** | **Périodes** | **Observé** | **St (*α* = 0,7)** | **Tt (*β* = 0,1)** | **Pt+rn = St + Ttm** |
| 2000 | 1 | 1 | 29444 | 29444,0 | 0,0 | - |
| 2000 | 2 | 2 | 40244 | 37004,0 | 756,0 | - |
| 2000 | 3 | 3 | 28420 | 31222,0 | 102,2 | 37760,0 |
| 2000 | 4 | 4 | 34013 | 33206,4 | 290,4 | 31324,2 |
| 2001 | 1 | 5 | 17267 | 22135,9 | -845,7 | 33496,8 |
| 2001 | 2 | 6 | 35127 | 30976,0 | 122,9 | 21290,3 |
| 2001 | 3 | 7 | 37662 | 35693,1 | 582,3 | 31098,9 |
| 2001 | 4 | 8 | 37871 | 37392,3 | 694,0 | 36275,4 |
| 2002 | 1 | 9 | 40208 | 39571,5 | 842,5 | 38086,3 |
| 2002 | 2 | 10 | 42304 | 41737,0 | 974,8 | 40414,0 |
| 2002 | 3 | 11 | 50536 | 48188,8 | 1522,5 | 42711,8 |
| 2002 | 4 | 12 | 38484 | 41852,2 | 736,6 | 49711,3 |
| 2003 | 1 | 13 | 50622 | 48212,0 | 1298,9 | 42588,8 |
| 2003 | 2 | 14 | 59735 | 56667,8 | 2014,6 | 49511,0 |
| 2003 | 3 | 15 | 73145 | 68806,2 | 3027,0 | 58682,4 |
| 2003 | 4 | 16 | 64166 | 66466,2 | 2490,3 | 71833,2 |
| 2004 | 1 | 17 | 68216 | 68438,1 | 2438,5 | 68956,5 |
| 2004 | 2 | 18 | - | - | - | 70876,6 |

Le tableau 11 illustre l’application du modèle de Holt à la série du tableau 10 pour ***α*** = 0,7 et ***β*** = 0,1. De plus, pour le choix des prévisions initiales, on a posé S1 = Y1 et T1 = 0.

Nous faisons observer que le choix optimal des paramètres ***α*** et ***β*** n’est pas aussi facile que celui de ***α*** dans les méthodes de lissage exponentiel simple (L.E.S.) ou double (L.E.L.), du fait de l’existence de deux paramètres ***α*** et ***β.***

En utilisant les équations de (2.34) à (2.36) pour préparer une prévision de la période 18, sachant que l’on est à la période 17, on obtient:

P18 = S17 + T17(1)

= [ ***α***(Y17)+ (1 ***α***) (S17 - 1 + T17 - 1)] + [***β***(S17  S17 - 1) + (1 ***β***)T17 - 1]

**=** [ ***α***(Y17) + (1 ***α***) (S16 + T16)] + [,8 (S17S16) + (1  ***β***)T16]

= [0,7 (68216) + 0,3 (66466,2 + 2490,3)] + [0,1(68438,1  66466,2) + 0,9(2490,3)]

= [47751,2 + 0,3 (68956,5)]+ [0,1(1971,9) + 2241,27]

= [47751,2 + **20686,95]** + [0,1(1971,9) +2241,27]

= [68438,1] + [2438,5]= 70876,6

**2.1.2.2.4 Le lissage exponentiel quadratique ou lissage exponentiel triple de Brown**

Nous inspirant des techniques de moyennes mobiles doubles et du lissage exponentiel double, nous pouvons .développer des modèles de lissage plus sophistiqués afin de prévoir des lois de nature quadratique ou de nature plus complexe. Le *lissage exponentiel quadratique* est l’un parmi ces modèles appelé aussi méthode de *lissage exponentiel triple.*

Il est utile de noter qu’on peut également passer du lissage quadratique au lissage cubique, du lissage cubique au lissage d’ordre supérieur, et ainsi de suite.

Le lissage exponentiel quadratique est défini par les trois équations de base suivantes:

**S’t**=  (2.37)

**S’’t**=  (2.38)

**S’’’t**=  (2.39)

at = 3S’t – 3S’’t + S’’t (2.40)

**bt**=  (2.41)

ct =  (2.42)

Pt + m = at + btm + ctm2 (2.43)

Les équations (2.37) et (2.38) se rapportent respectivement au lissage exponentiel simple et au lissage exponentiel double. L’équation (2.39) nous donne la formule du lissage exponentiel triple, tandis que les équations (2.40) à (2.42) déterminent respectivement la valeur des données actuelles, le trend linéaire et le trend quadratique. Pour prévoir les données des périodes futures, on utilise la formule (2.43) qui prend en compte le trend linéaire et le trend quadratique.

Comme on peut aisément le vérifier, les calculs nécessaires à ce type de prévision deviennent très lourds et sont plus complexes que ceux utilisés dans les lissages simple et linéaire.

Avec les données du tableau 11 et une valeur de ***α*** = 0,7, nous pouvons, par exemple, prévoir la consommation du ciment au cours du deuxième trimestre de l’année 2004 (période 18). Tous les éléments de calculs sont ‘consignés dans le tableau 12 ci-dessous où :

P18 = a17 + b17 + 0,5 c17

Avec :

at = 3S’17 – 3S’’17 + S’’17 = 3(67350,5) – 3(66588,0) + 65599,5 = 67666,9

**bt**= 

=

ct = 

Ainsi P18 = a17 + b17 + 0,5 c17 = 67886,9 + 108,7 – 1230,8 = 67380,2

**2.1.2.2.5 Le lissage exponentiel de Holt et Winters**

Cette technique, connue également sous le nom de **méthode de Holt - Winters avec saisonnalité** et mise au point en 1960, est une généralisation de la méthode de Holt conçue pour une fonction de prévision localement linéaire. Il s’agit d’un lissage exponentiel double (LED) de Holt à deux paramètres pour la partie non saisonnière et d’un lissage exponentiel saisonnier à un paramètre de Winters. Cette méthode de lissage exponentiel comporte donc trois paramètres à estimer et il existe deux manières de combiner la tendance linéaire et la composante saisonnière : par addition (modèle additif) et par multiplication (modèle multiplicatif).

La méthode donne des résultats identiques à ceux du lissage exponentiel double développé plus haut, mais a l'avantage supplémentaire d'incorporer un coefficient saisonnier It, avec une période de saisonnalité égale à **L**, c'est-à-dire que It = Tt-L pour tout t (dans la pratique, les estimations de It ne seront toutefois pas périodiques).

Le lissage exponentiel de Winters utilise trois équations reprises ci-après. Chacune d'elles a pour fonction de lisser l'une des trois composantes de la série : l'aléa, la tendance et la composante saisonnière. Elle ressemble en cela au lissage exponentiel double (de Holt) qui ajuste la tendance et lisse l'aléa: mais elle ajoute un paramètre supplémentaire  pour traiter la saisonnalité.

Dans le cas où la composante saisonnière est introduite de manière additive, les formules à utiliser sont :

 (lissage de la moyenne) (2.44)

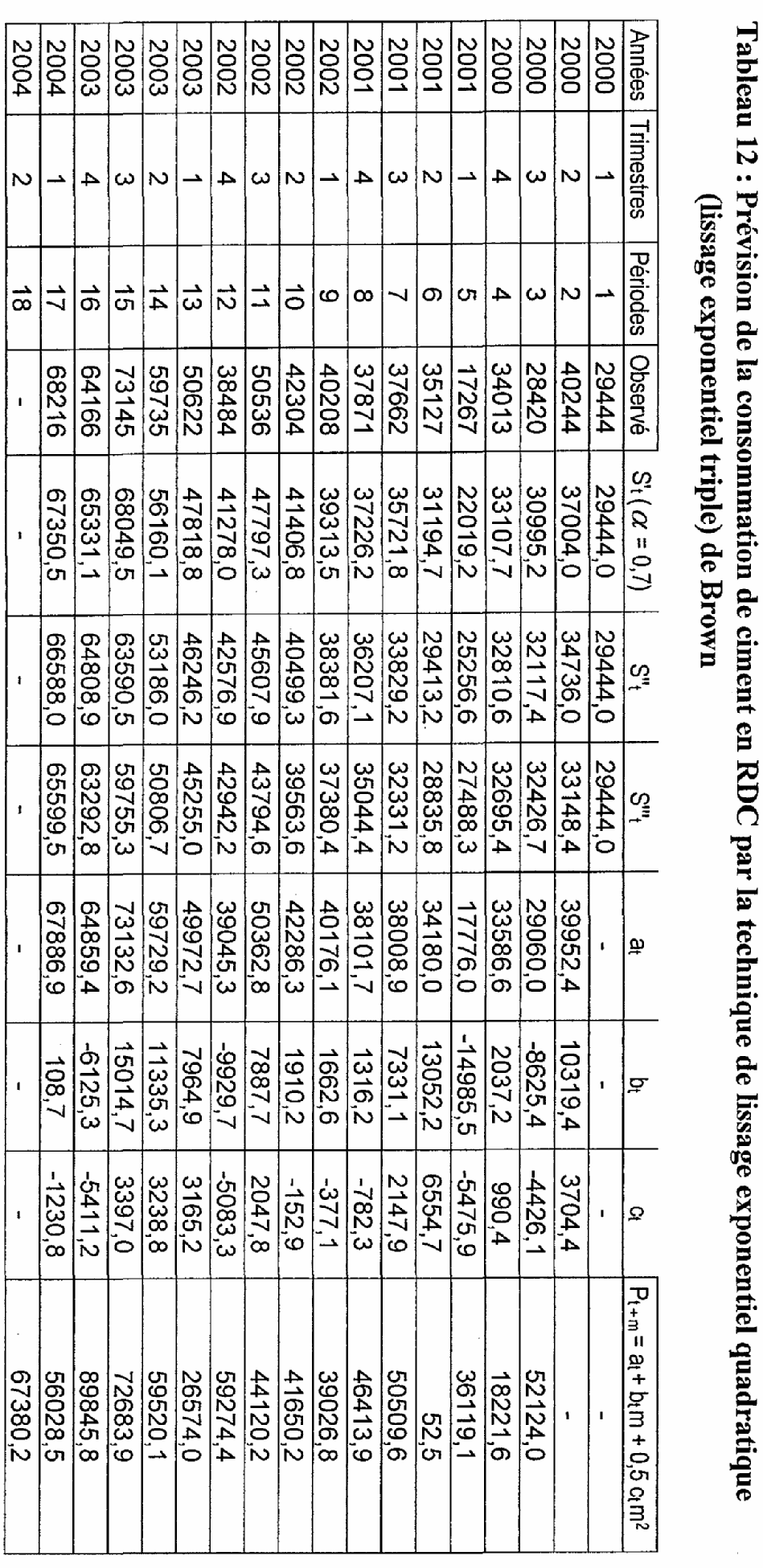
Tt =  (St  St ‑ 1) + (1  )Tt ‑ 1 (lissage de la tendance) (2.45)

It =  (Yt  St) + (1  )It ‑ L (lissage de la saisonnalité) (2.46)

Prévision à un horizon de m périodes :

Pt + m = St + Ttm + It ‑ L+ m (2.47)

Le principe de la conservation des aires implique : .



Les formules suivantes sont utilisées dans l'hypothèse d'une décomposition multiplicative.

 (2.44a)

Tt =  (St – St ‑ 1) + (1 )Tt – 1 (lissage de la tendance) (2.45a)

It =   + (1  ) It ‑ L (lissage de la saisonnalité) (2.46a)

Prévision à un horizon de m périodes :

Pt + m = (St + Ttm) It ‑ L + m (2.47a)

avec

St = moyenne lissée de la série en t ;

Tt = tendance estimée en t ;

Yt = valeur observée en t ;

L = périodicité des données ou longueur de la saisonnalité (L = 12 en mensuel, L = 4 en

trimestriel) dans une année) ;

It = facteur d'ajustement saisonnier en t.

Dans le modèle multiplicatif, les coefficients saisonniers vérifient la propriété : .

L'équation (2.46) est comparable à celle d'un indice saisonnier : c'est la différence entre la dernière valeur de la série, Yt, et la valeur prévue par lissage simple,St (Pour le modèle multiplicatif, c'est le quotient de Yt par St). Rappelons que St est une valeur lissée (moyenne) de la série d'où le saisonnier est absent. La valeur de la donnée Yt, elle, inclut la composante saisonnière et un facteur aléatoire de la série. Pour la lisser, l'équation (2.46) pondère la donnée corrigée des variations saisonnières qu'on vient de calculer (Yt - St) (ou Yt/St pour un modèle multiplicatif) par , et le plus récent coefficient saisonnier correspondant à la même saison (It ‑ L) par (1  ).

L'équation (2.45) relative à Tt a seulement pour but de lisser la tendance, en pondérant par  l'accroissement tendanciel (St  St ‑ 1) et l'ancienne valeur de la tendance (Tt ‑ 1) par (1 ).

Dans l'équation de lissage St (relation (2.44)), le coefficient saisonnier It - L est enlevé à la variable observée Yt (It - L divise Yt pour un modèle multiplicatif), afin de désaisonnaliser cette dernière (c'est‑à-dire, afin d'éliminer les fluctuations saisonnières de Yt).

L'équation (2.45) est identique à la relation (2.34) de la méthode de lissage de Holt. Le terme It ‑ L+ m étant le dernier coefficient saisonnier disponible, il est utilisé pour réajuster la prévision (par l'addition à Ttm). La prévision dans la méthode de Holt - Winters fait intervenir ainsi les effets saisonniers.

Les tableaux 14 et 15 permettent d'illustrer la méthode de Holt - Winters dans les hypothèses additive et multiplicative respectivement, avec les valeurs suivantes des coefficients ***α*** = 0,8,  = 0,1 et  = 0,3, choisies conformément au critère MSE.

Comme pour les autres méthodes de lissage, dans la pratique, il y a un problème de démarrage de la technique : les valeurs de départ peuvent être estimées par la méthode des moindres carrés ordinaires ou plus simplement initialisées pour la première année en prenant les moyennes arithmétiques des k premières périodes .

Nous avons pris les valeurs initiales suivantes :

St vaut, par exemple, la moyenne de la première année disponible;

Tt = 0 ;

It = Yt  St, t =1, 2, ..., L.

Des variantes sont évidemment possibles (Voir le paragraphe suivant).

Les valeurs initiales des trois composantes ont été déterminées sur base de la première année des données. S4 =287,2.

**Tableau 14 : Prévision des ventes trimestrielles de la bière Skol par la méthode de Holt -**

**Winters de 1988 à 1993** **(en milliers de hl) : Modèle additif**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Périodes | Années | Trimes-  tres | Vente  ob­servée Yt | Valeur lissée  St (***α*** = 0,8) | Valeur lissée  Tt ( = 0,1) | Indice Saisonnier  It ( = 0,3) | Prévision  avec m =1 |
| 1 | 1988 | I | 260,8 | ‑ | ‑ | 26,4 | ‑ |
| 2 | II | 290,6 | ‑ | ‑ | 3,4 | ‑ |
| 3 | III | 270,6 | ‑ | ‑ | 16,6 | ‑ |
| 4 | IV | 326,7 | 287,2 | 0,0 | 39,5 | - |
| 5 | 1989 | I | 243,0 | 272,9 | 1,4 | 27,4 | 260,8 |
| 6 | II | 270,7 | 268,1 | 1,8 | 3,2 | 274,9 |
| 7 | III | 206,4 | 231,7 | 5,2 | 19,2 | 249,8 |
| 8 | IV | 249,2 | 213,0 | 6,6 | 38,5 | 265,9 |
| 9 | 1990 | I | 198,3 | 221,9 | 5,0 | 26,3 | 179,0 |
| 10 | II | 223,6 | 219,7 | 4,7 | 3,4 | 220,0 |
| 11 | III | 207,1 | 224,0 | 3,8 | 18,5 | 195,8 |
| 12 | IV | 250,0 | 213,2 | 4,5 | 38,0 | 258,7 |
| 13 | 1991 | I | 255,4 | 267,1 | 1,3 | 21,9 | 182,4 |
| 14 | II | 284,6 | 278,6 | 2,3 | 4,2 | 271,8 |
| 15 | III | 99,9 | 150,9 | 10,7 | 28,3 | 262,5 |
| 16 | IV | 120,7 | 94,2 | 15,3 | 34,5 | 178,2 |
| 17 | 1992 | I | 143,6 | 148,2 | 8,4 | 16,7 | 57,0 |
| 18 | II | 160,0 | 152,6 | 7,1 | 5,1 | 144,0 |
| 19 | III | 150,7 | 172,3 | 4,4 | 26,3 | 117,3 |
| 20 | IV | 182,0 | 151,5 | 6,0 | 33,3 | 202,4 |
| 21 | 1993 | I | 134,6 | 150,1 | 5,6 | 16,4 | 128,8 |
| 22 | II | 150,0 | 144,8 | 5,5 | 5,1 | 149,7 |
| 23 | III | 141,1 | 161,7 | 3,3 | 24,6 | 113,0 |
| 24 | IV | 170,7 | 141,6 | 5,0 | 32,1 | 191,8 |
| 25 | 1994 | I | ‑ | - | ‑ | ‑ | 120,2 |

Les coefficients saisonniers initiaux It sont obtenus par différence des données de la première année et de cette moyenne. Par exemple, dans l'hypothèse additive, ­pour les premier, deuxième, troisième et quatrième trimestres, on a respectivement :

260,8  287,2 = 26,4 ;

290,6  287,2 = 3,4 ;

270,6  287,2 = 16,6 et

326,7  287,2 = 39,5.

La pente initiale Tt est supposée être égale à 0.

Les calculs à faire pour obtenir la prévision des ventes du premier trimestre de l'année 1994 (période 25) sont les suivants :

***Pour le modèle additif,***

 (lissage de la moyenne) (2.44)



= 0,8(170,7  33,3) + 0,2[(161,7 + (3,3)] = 141,6

Tt =  (St  St ‑ 1) + (1  )Tt ‑ 1 (lissage de la tendance) (2.45)

T24 = 0,1(S24  S23) + 0,9 T23

= 0,1(141,6  161,7) + 0,9(3,8) = 5,0

It =  (Yt  St) + (1  )It ‑ L (lissage de la saisonnalité) (2.46)

I24 = 0,3(Y24  S24) + (0,7)I20

= 0,3(170,7  141,6) + 0,7(33,3) = 32,1

et

Pt + m = St + Ttm + It ‑ L+ m (2.47)

P25 = S24 + T24(1)] + I22

= [141,6 + (5,0) + 5,1 = 120,2

***Pour le modèle multiplicatif,***



S24 = 0,8(170/1,1) + 0,2(S23 + T23)

= 0,8(170,7/1,1) + 0,2[(157,2 + (–3,8)] = 152,4

T24 =  (S24 – S24 ‑ 1) + (1  )T24 – 1 (lissage de la tendance) (2.45a)

T24 = 0,1(S24  S23) + 0,9 T23

= 0,1(152,4  157,2) + 0,9(3,8) = 3,9

I24 =   + (1  ) I24 ‑ L (lissage de la saisonnalité) (2.46)

I24 =   + (1  ) I24 ‑ L

= 0,3(Y24/S24) + (0,7)I20 = 0,3(170,7/152,4) + 0,7(1,1) = 1,1

Pt + m = (St + Ttm) It ‑ L + m (2.47a)

P25 = [S24 + T24(1)]I22 = 152,4 + {(3,9)(1)}] (1,0) = 139,0

Pour déterminer lequel de deux modèles il est préférable d’utiliser pour faire les prévisions, il est recommandé de comparer les indicateurs de performance habituellement utilisés notamment le MSE et le MAPE. Le meilleur modèle est celui qui minimisera le MSE ou le MAPE.

**Tableau 15 : Prévision des ventes trimestrielles de la bière Skol par la méthode**

**de Holt Winters de 1988 à 1993 (en milliers de hl) : Schéma multiplicatif**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Périodes | Années | Trimestres | Vente observée Yt | Valeur lissée  St ***α*** =0,8 | Valeur lissée Tt  = 0,1 | Indice Saisonnier  It = 0,3 | Prévision  avec m =1 |
| 1 | 1988 | I | 260,8 | ‑ | ‑ | 0,9 | ‑ |
| 2 | II | 290,6 | ‑ | ‑ | 1,0 | ‑ |
| 3 | III | 270,6 | ‑ | ‑ | 0,9 | ‑ |
| 4 | IV | 326,7 | 287,2 | 0,0 | 1,1 | ‑ |
| 5 | 1989 | I | 243,0 | 271,5 | 1,6 | 0,9 | 260,8 |
| 6 | II | 270,7 | 268,0 | 1,8 | 1,0 | 273,1 |
| 7 | III | 206,4 | 228,5 | 5,5 | 0,9 | 250,9 |
| 8 | IV | 249,2 | 219,8 | 5,8 | 1,1 | 253,6 |
| 9 | 1990 | I | 198,3 | 218,2 | 5,4 | 0,9 | 193,5 |
| 10 | II | 223,6 | 219,4 | 4,8 | 1,0 | 215,2 |
| 11 | III | 207,1 | 221,0 | 4,1 | 0,9 | 199,8 |
| 12 | IV | 250,0 | 219,4 | 3,9 | 1,1 | 246,4 |
| 13 | 1991 | I | 255,4 | 268,7 | 1,4 | 0,9 | 195,1 |
| 14 | II | 284,6 | 278,6 | 2,3 | 1,0 | 273,9 |
| 15 | III | 99,9 | 141,9 | 11,6 | 0,9 | 262,0 |
| 16 | IV | 120,7 | 110,9 | 13,5 | 1,1 | 148,2 |
| 17 | 1992 | I | 143,6 | 144,5 | 8,8 | 0,9 | 89,5 |
| 18 | II | 160,0 | 153,1 | 7,1 | 1,0 | 137,8 |
| 19 | III | 150,7 | 168,7 | 4,8 | 0,9 | 126,2 |
| 20 | IV | 182,0 | 162,5 | 5,0 | 1,1 | 184,0 |
| 21 | 1993 | I | 134,6 | 145,9 | 6,1 | 0,9 | 148,3 |
| 22 | II | 150,0 | 145,1 | 5,6 | 1,0 | 143,2 |
| 23 | III | 141,1 | 157,2 | 3,8 | 0,9 | 121,7 |
| 24 | IV | 170,7 | 152,4 | 3,9 | 1,1 | 172,1 |
| 25 | 1994 | I | ‑ | ‑ | ‑ | ‑ | 139,0 |

**2.1.2.2.6 Les Valeurs d'origine dans les Méthodes de Lissage Exponentiel : Besoin d'une initialisation**

La raison pour laquelle on a besoin de valeurs d'origine dans les méthodes de lissage exponentiel apparaît à l'examen de la formule du lissage exponentiel simple :

Pt+1 = ***α*** Yt + (1  ***α***)Pt (2.11)

où

Yt = dernière valeur observée ;

Pt = dernière prévision ;

Pt + 1 = prévision pour la période suivante, et

***α*** = constante de lissage.

Si t = 1, l'équation (2.11) devient :

P2 = ***α*** Y1 + (1  ***α***)P1

On a besoin, pour connaître la valeur de P2, de calculer celle de P1. Ce devrait être :

P1 = ***α*** Y0 + (1 ***α***)P0

On voit bien ici que Y0 n'existe pas, et que P0 ne peut être calculé. C'est là le problème : la valeur de P1, dans l'équation (2.11), doit être connue pour calculer P2, mais on ne peut la dériver des données qu'on a; il faut donc trouver une autre approche pour estimer la valeur initiale de P1 dans (2.11).

De même, on a besoin de valeurs initiales pour n'importe quel type de lissage exponentiel: leur nombre et leur type dépendent de la méthode de lissage exponentiel utilisée. Nous décrivons ci-après les méthodes les plus utilisées pour initialiser les valeurs d'un lissage exponentiel. On supposera qu'on dispose de données passées en nombre suffisant pour employer une des méthodes suivantes :

**1. Séparation des données en deux parties**

La première sert à estimer les valeurs initiales, et la seconde à estimer les valeurs optimales des paramètres. On considère empiriquement que les données de 3 à 8 (L) suffisent pour l'estimation des valeurs d'origine (L étant la longueur du cycle saisonnier). Si l'on dispose d'assez de données (observations sur 6 à 16 (L), on en emploiera la moitié pour l'estimation initiale, et l'autre moitié pour optimiser les paramètres.

**2. Prévision à rebours**

C'est une technique largement utilisée dans la méthode de Box‑Jenkins, mais elle peut aussi s'appliquer pour le lissage exponentiel. Elle consiste à inverser la série de données, et à partir des dernières (les plus récentes) pour terminer par les premières (les plus anciennes). Ceci faisant, on obtient des prévisions et/ou une estimation des paramètres pour le début de la série; on peut les utiliser comme valeurs de départ lorsque l'on effectue la prévision dans le sens normal, du début à la fin.

1. Avec VCM=Valeur Critique de MacKinnon ; =significativité au seuil de 5% ; =significativité au seuil de 1%

   [↑](#footnote-ref-2)
2. On a préféré utiliser ***αt+1*** plutôt que ***αt*** car le lissage exponentiel adaptatif est très sensible aux changements. En se servant de ***αt***, on introduit un petit décalage d’une période, lequel décalage permet au système de s’adapter un tout petit peu et de prévoir selon la méthode consacrée. [↑](#footnote-ref-3)
3. [↑](#footnote-ref-4)