



UNIVERSITÉ DE NANTES



IAE NANTES
ÉCONOMIE & MANAGEMENT

Prédiction et Financiarisation des Matières Premières

Dossier Série Temporelle et modèle ARIMA

Kyllien ROMAND

2020 – 2021

SOMMAIRE

Introduction

Analyse économique du sujet

Analyse des trois variables

Présentation du modèle ARIMA

Choix du modèle ARIMA pour le SPGSCI

Analyse des variables WTI et S&P 500

Test de cointégration selon Engle-Granger

Conclusion et discussion

Bibliographie

Annexes

Tables des Matières

Introduction

Le processus de financiarisation fait référence à la situation dans laquelle le prix d'une matière première n'est pas déterminé que par l'offre et la demande, mais aussi par des acteurs financiers et le comportement des investisseurs sur les marchés dérivés.¹

La financiarisation des matières premières est un sujet de plus en plus présent dans le monde économique et financier. Avant les années 2000, ce marché était surtout investi par des spécialistes et les producteurs physiques afin de se couvrir par un rapport un risque, il ne faut pas oublier que les premiers traders étaient les producteurs de matières agricoles puis est venu le marché action. Depuis cette date, de nombreux nouveaux investisseurs et produits ont fait surface : fond spécialisé en matières premières (OPVCM), tracker de cours boursier d'une matière première (ETF) ; normalement ces acteurs ne devraient pas avoir d'impact sur le cours des matières, cependant il en ressort que de nombreuses spéculations ont lieu dorénavant surtout pendant la période 2008 mais aussi en 2011/2012, où les prix ont connu une flambée et a amené certaines régions du monde dans une crise alimentaire.

La période de 2007-2008 étant une période trouble et de "folie financière", période où la spéculation et la financiarisation étant très importante, la période 2011/2012 sera donc préféré à étudier car suite à la période décrite précédemment les marchés agricoles ont été l'objet de nouvelle régulation : tel que la mise en place d'un bulletin hebdomadaire de la Commodity Futures Trading Commission (CFTC) avec les changements de positions des acteurs financiers sur le marché des futures des matières agricoles.

Dans ce dossier, il sera effectué dans un premier temps une analyse économique du sujet à l'aide d'une revue de la littérature économique, pour ensuite présenter les variables utilisées, puis sera détaillé la méthode utilisée (modèle ARIMA) et enfin une présentation des résultats du modèle ARIMA et du Test de cointégration selon Engle-Granger.

¹ Créti et al. 2013

Analyse Economique du Dossier

Le premier marché financier (marché dérivé) de matières premières a été créé au XIX^{ème} siècle à Chicago, il permettait l'achat et la vente de ces dernières de gré à gré ; les marchés dérivés ont un rôle primordial dans l'économie puisqu'ils permettent d'avoir un lieu d'échange à terme pour les matières premières, d'avoir une meilleure information et visibilité des prix. Les marchés dérivés sont donc le lieu de formation des prix de référence (*prix spot*) mais aussi le lieu de couverture (*contrat futures*) pour les différents acteurs du marché des matières premières.

Les marchés dérivés sont constitués de quatre types de contrat dont le sous-jacent peut être une matière première :

- Les contrats *Forward* qui sont échangés sur les marchés de gré à gré, ces contrats sont échangés bilatéralement entre les acteurs, c'est à dire que le contrat s'échange entre un acheteur et un vendeur, les deux parties se sont accordées sur la quantité de sous-jacent à livrer, le prix à payer et la date de livraison. Dans ce cadre de contrat, la livraison intervient obligatoirement contrairement aux contrats *Futures*, cependant un problème systémique se pose avec ce type de contrat : si un des deux parties est amené à ne pas pouvoir honorer le contrat, l'autre partie est lésée. Les marchés OTC représentent la majorité des échanges de matières premières dans le monde.
- Les contrats *Futures* comme explicité précédemment n'amènent pas forcément livraison du sous-jacent (seulement 1% des contrats *Futures* amènent à une livraison du sous-jacent), ces contrats interviennent sur des marchés dits organisés, les contrats sont standardisés : les quantités du sous-jacent sont les mêmes pour chaque contrat, les prix et les contrats aussi. Tout ceci est déterminé par le marché, de plus afin d'éviter le défaut d'un des deux parties, un troisième acteur apparaît : le marché où sont échangés les contrats avec la Chambre de Compensation, qui permet de faire des appels de marge après chaque journée et donc de vérifier si les acteurs présents sur ce marché sont toujours viables, si un des deux parties ne peut valider l'appel de marge alors sa position est clôturée et la chambre de compensation reprend sa position afin que la partie qui honore sa part du contrat ne soit lésée.
- Les *Options* permettent, moyennant une prime d'option, à un opérateur d'exercer une option sur une position c'est à dire que l'option peut être prise ou non tout dépend si le marché évolue dans le sens de l'acteur.
- Les *SWAPS* sont plus complexes, créés par des ingénieurs financiers, ils permettent un transfert de risque entre deux acteurs financiers.

Tous les différents acteurs des marchés financiers ont pour but d'être rentable, et donc de créer un bénéfice en achetant ou vendant des produits financiers. La volatilité des marchés financiers étant très importante, les variations de prix le sont tout autant, c'est pourquoi depuis des années de nombreux mathématiciens et statisticiens ont étudiés les évolutions des prix des marchés financiers, afin d'établir des modèles de prédictions pour les prix et donc de savoir si un prix va être sur une tendance haussière ou baissière, et de permettre de

prendre position soit en *long* soit en *short* afin de s'assurer un bénéfice. De plus depuis l'arrivée de l'algorithmie et du machine learning, la volatilité permet aux algorithmes de créer un bénéfice sur de très courte période (algorithme de haute de fréquence) en leur permettant d'acheter bas pour vendre haut sur des plus bas ou plus hauts journaliers. La mise en place d'un modèle ARIMA permet de faire des prédictions de prix et déterminer la tendance future de la variable étudiée.

Graphique 1 : Evolution du VIX (indice de volatilité sur le S&P 500) sur la journée du 26/02/2021

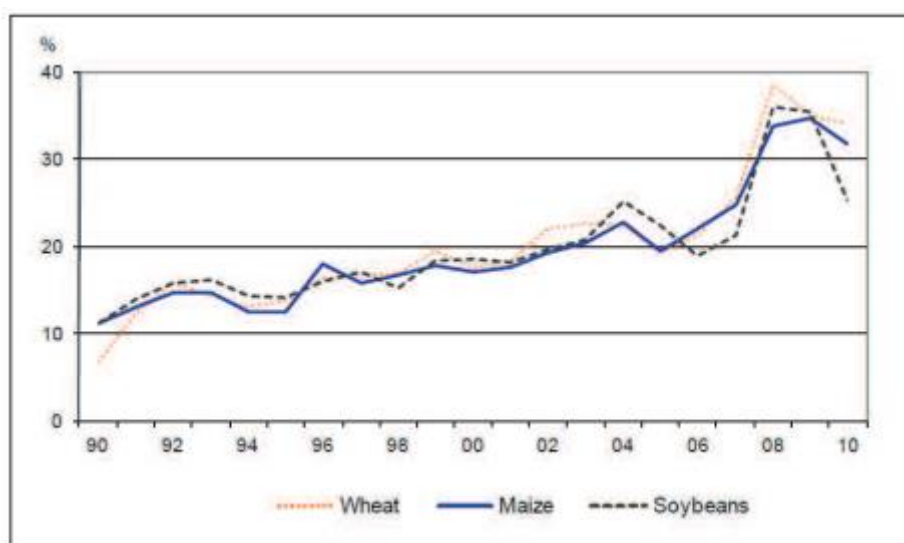


Source : Yahoo Finance

Comme le montre le graphique 1, si un bon modèle de prédiction aurait pu permettre de prédire correctement les variations de l'indice VIX, un trader ou un algorithme aurait pu être *long* sur chaque début de flèche verte (tendance haussière) et clôturer sa position au bout de celle-ci, puis être *short* lors des flèches rouges (tendance baissière) et clôturer à la fin de cette tendance afin d'engranger des profits. Certains algorithmes de haute fréquence permettent de prendre positions et de clôturer en quelques ms, ces premiers "jouent" sur des changements de valeurs (*ticks*) sur de très courte durée et engrangent de plus petits bénéfices à chaque *trade*, mais en effectuent beaucoup plus sur toute une journée. L'arrivée de ces nouveaux acteurs est aussi un facilitateur de volatilité mais ne sera pas étudié dans ce dossier.

Concernant la financiarisation des matières premières, depuis les années 1980, les marchés de matières premières connaissent un vrai essor suite à la dérèglementation aux Etats-Unis de ces derniers, mais aussi en Grande-Bretagne avec l'abolition des contrôles de changes. En 2000 un nouveau pas est franchi aux Etats-Unis, avec la totale dérèglementation des marchés de matières premières et l'arrivée de nouveaux acteurs : les spéculateurs. Ce nouveau pas franchi amène une hausse de la volatilité des prix, cependant la véritable hausse des denrées alimentaires commencent en 2007 puisque les banques d'investissement se sont pourvus d'infrastructure de stockage, de transport et de distribution pour les matières premières, ceux qui leur permettent de ne plus être spectateurs de ce marché mais aussi acteurs, et donc d'influencer (voire manipuler) les prix de ces dernières. Cette dérèglementation des années 2000 est remise en cause par la FED depuis 2014.

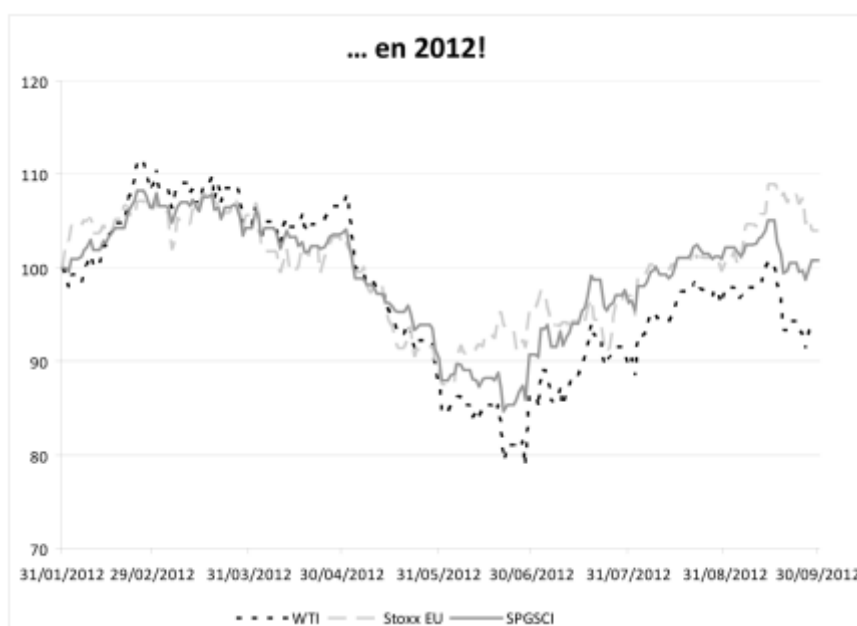
Graphique 2 : Evolution de la volatilité implicite des Matières Premières entre 1990 et 2010



Source : Rapport sur la volatilité des prix agricoles de 10 organisations internationales pour le G20

Comme le montre le graphique 2, une hausse de la volatilité sur les marchés agricoles à lieu depuis les années 1990. Ces marchés normalement peu volatils ont connu de forte hausse de variations de prix depuis l'arrivée de la dérèglementation et des nouveaux acteurs financiers, cette hausse de la volatilité est un problème puisqu'elle amène à donner une mauvaise information concernant le prix du sous-jacent, et peut donc amener à des périodes de tensions alors que la production (l'offre) se porte correctement. Cette hausse de la volatilité peut faire penser qu'entre les marchés des actions et les marchés de matières premières, le même niveau de volatilité est apparu comme le montre le graphique 3, où des co-mouvements de ces 3 actifs qui ne sont normalement peu corrélés apparaissent. Ce graphique de l'année 2012, montre l'hypothèse d'une cointégration entre ces actifs et la possibilité d'établir une relation entre eux.

Graphique 3 : Évolution de l'Euro Stoxx 600, du pétrole brut WTI et de l'indice des matières premières de Standard & Poor et Goldman Sachs (SPGSCI) – 31 janvier au 30 septembre 2012



Source : calculs du secrétariat de la CNUCED basés sur la base de données de Bloomberg

La financiarisation des matières premières est un long débat qui est ressorti suite aux différentes crises alimentaires apparues suite à la crise financière de 2007-2008, certains défendent l'idée que la hausse des prix survenue est dû à l'augmentation de population et au changement de mode de consommation des populations des pays émergents, d'autres affirment que la dérèglementation des années 2000 et la folie financière apparue jusqu'en 2008 en sont la cause. La cointégration selon Engle-Granger permettra de déterminer une relation de court terme et de long terme entre les variables de ce dossier, si une régression de bonne qualité est possible entre un indice de marché boursier et un indice de prix des différentes matières premières alors une question se pose concernant la création de prix des matières premières qui ne serait donc plus dû aux variations des variables fondamentales : les stocks, la production, l'offre et la demande, mais aussi à l'activité sur les marchés dérivés des matières premières des différents acteurs financiers présents sur ces marchés.

Analyse des 3 Variables

Les données proviennent du site internet : <https://fr.investing.com/> . Elles commencent le 03/01/2012 et finissent le 28/09/2012, cependant des jours fériés apparaissent qui ne sont pas pris en compte par le logiciel Gretl, c'est pourquoi la date de fin est modifiée dans ce logiciel.

Le Standard & Poors and Goldman Sachs Commodity Index (SPGSCI)

Cet indice boursier est développé par la banque d'investissement Goldman Sachs et passé sous le contrôle de Standard & Poors, c'est une référence des prix des matières premières dans le milieu financier. Il est calculé à l'aide d'un panier de différentes *commodities*, dont la composition évolue chaque année en fonction du volume de production de chaque matière première.

Tableau 1 : Composition de l'indice avec le poids de chaque matière première dans l'indice

Sub-Index	2019 RPDW*	2020 RPDW	Included Commodities
Energy	62.63%	61.71%	Crude Oil (and supporting contracts) and Natural Gas
Petroleum	59.52%	58.47%	Crude Oil (and supporting contracts)
Non-Energy	37.37%	38.29%	All commodities not included in Energy Sub-Index
Agriculture	15.41%	15.89%	Wheat (Chi. & Kan.), Corn, Soybeans, Coffee, Sugar, Cocoa and Cotton
Grains	11.42%	12.12%	Wheat (Chi. & Kan.), Corn and Soybeans
Livestock	6.65%	7.25%	Lean Hogs, Live Cattle and Feeder Cattle
Industrial Metals	11.16%	10.65%	Aluminum, Copper, Lead, Nickel and Zinc
Precious Metals	4.14%	4.50%	Gold and Silver

Source : S&P GSCI Methodology

Comme le montre le Tableau 1, cet indice est principalement composé du prix du pétrole brut c'est pourquoi des co-mouvements seront remarqués avec le WTI cependant aucun lien ne devrait être fait avec le S&P 500.

Graphique 4 : Evolution du prix de l'indice S&P GSCI

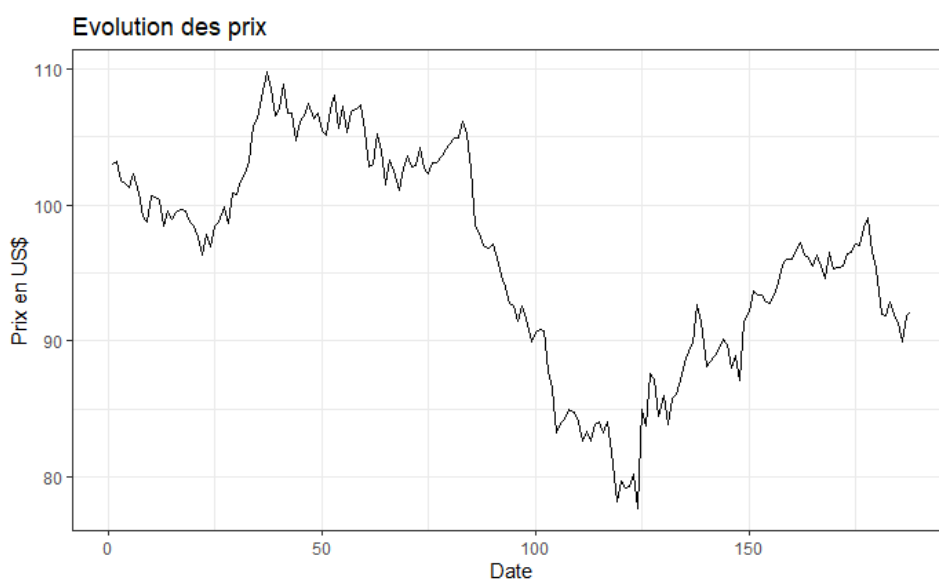


Source : Dossier

Futures Pétrole Brut West Texas Intermediate (WTI)

Cet actif correspond au contrat *futures* à 1 mois du pétrole brut WTI, qui est un pétrole extrait en Amérique du Nord, il est la référence mondiale concernant le pétrole avec le BRENT car ce sont tous les deux des bruts légers, et ont donc un coût faible en raffinage contrairement au pétrole de schiste par exemple. Ils permettent d'avoir les informations concernant le marché du pétrole.

Graphique 5 : Evolution du prix du WTI

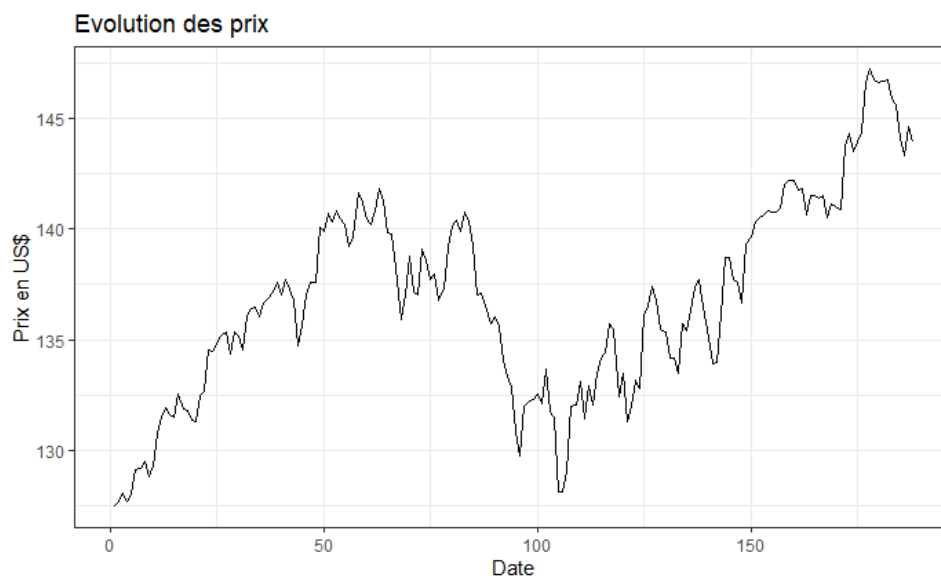


Source : Dossier

Indice S&P 500

L'indice boursier Standard and Poor's 500 correspond à l'évolution des 500 plus grandes entreprises américaines par capitalisation boursière, il permet d'évaluer les performances de l'économie américaine et du monde en général puisque l'économie américaine est le moteur de l'économie mondiale (de moins en moins vrai avec l'émergence de la Chine). Cependant de nombreux biais apparaissent dans cet indice, depuis quelques années les fortes capitalisations boursières (GAFA : Google, Amazon, Facebook, Apple ; Microsoft) ont un poids de plus en plus important dans cet indice, et fausse donc l'idée de se référer à cet indice afin d'avoir une idée sur l'économie américaine.

Graphique 6 : Evolution du Prix de l'indice S&P 500



Source : Dossier

Présentation du Modèle ARIMA

Lors d'une modélisation de série temporelle, trois modèles sont possibles en fonction de la série à étudier :

- Les Modèles d'Ajustements
- Les Modèles Autorégressifs ou Autoprojectifs
- Les Modèles Explicatifs

Le modèle ARIMA est un modèle autorégressif, il correspond au Moyenne Mobile Autorégressive Intégré en anglais. C'est un modèle qui se constitue de deux parties :

- Le Modèle AR qui est le modèle autorégressif
- Le Modèle MA qui est le modèle des moyennes mobiles

La mise en place de ces modèles se décide à l'aide l'analyse du corrélogramme, qui sera expliqué plus tard dans le dossier, si le modèle choisit mélange modèle AR et modèle MA alors ce modèle est mixte. Une variante du modèle ARIMA permet d'intégrer la saisonnalité de la série chronologique : modèle SARIMA.

Tous ces modèles permettent de faire de la prédiction d'une série temporelle avec des niveaux de qualité qui dépendront du modèle choisit.

Stationnarité

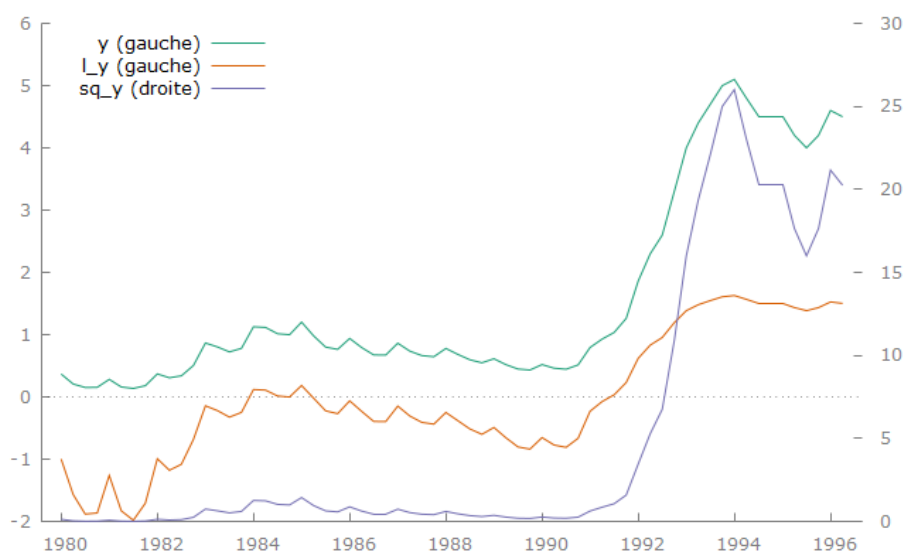
Avant d'utiliser un modèle ARIMA, il nécessite de vérifier si la structure de la série est stable, si elle est stationnaire. Pour qu'une série soit stationnaire, il faut que la série soit stable au niveau de la variance, c'est à dire que les covariances sont stables par translation du temps (h) : **$\text{Cov}[X_i, X_{i+h}] = \sigma(h)$ pour tout i** , ou **$\text{Var}[X_i] = \sigma(0)$** ; mais aussi stable par la moyenne c'est à dire que les espérances sont les constantes : **$E[X_i] = \mu$ pour tout i** . Il faut donc tout d'abord vérifier la stationnarité au niveau de la variance puis au niveau de la moyenne.

Pour vérifier la stationnarité, on le vérifie graphiquement, il ne faut pas que les écarts entre un point et la tendance ne soit trop grande pour ça il existe deux méthodes :

- Passer la série chronologique en logarithme : $y_t = \log(Y_t)$
- Passer la série chronologique à la racine carré : $y_t = \sqrt{Y_t}$

Si graphiquement la série se modifie lorsqu'on utilise une des deux méthodes, la série brute n'était pas stationnaire au niveau de la variance, si aucune grande modification apparaît alors la série brute est stationnaire au niveau de la variance. Pour illustrer ceci, une série d'exemple du logiciel Gretl est utilisé : *Swiss Unemployment*. Le graphique 7 met en évidence les modifications apportées par les différentes transformations, ici la série brute n'était pas stationnaire par la variance puisque les trois courbes ne sont pas "proches", il sera préféré une transformation logarithmique pour cette série.

Graphique 7 : Série Brute du Chômage Suisse, Transformer en Log et à la Racine Carré



Source : Gretl

La série brute du Chômage Suisse n'est pas stationnaire au niveau de la moyenne, afin de la rendre stable au niveau de l'espérance la méthode utilisée est la différenciation d'ordre 1 ou 2, cependant si après cette méthode la série n'est toujours pas stationnaire il faudra appliquer d'autres méthodes : tel qu'une différenciation saisonnière afin d'intégrer une composante

saisonniers mais cela ne correspondra plus à l'utilisation d'un modèle ARIMA mais SARIMA.
Formule de la différenciation d'ordre d :

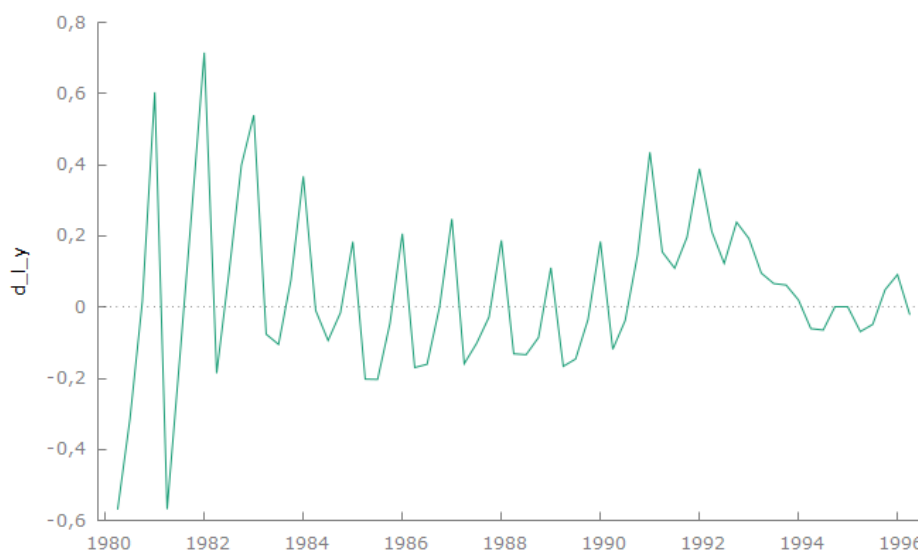
$$Z_t = \nabla^d Y_t$$

Avec :

- Y_t : La série brut stationnaire à la variance et non à la moyenne
- Z_t : La série stationnaire à la variance et à la moyenne
- ∇ : Signe de la différenciation
- d : Ordre de la différenciation

Le graphique 8 permet de montrer que la série différenciée varie autour de la moyenne qui correspond à la ligne en pointillé tracé sur le graphique ($y=0$), la série semble être stationnaire.

Graphique 8 : Série du Chômage Suisse Transformer Logarithmiquement Différencié à l'ordre 1



Source : Gretl

Afin de vérifier la stabilité de la série différenciée, il est utile par la suite d'étudier son corrélogramme.

Toutefois outre le fait de vérifier graphiquement si une série est stationnaire, il est utile d'appliquer un test statistique sur la série stationnaire graphiquement afin de s'assurer de ceci : le test augmenté de Dickey-Fuller ou test ADF qui permet de savoir si les différentes propriétés statistiques (espérance, variance et auto-corrélation) varient ou pas dans le temps.

Corrélogramme

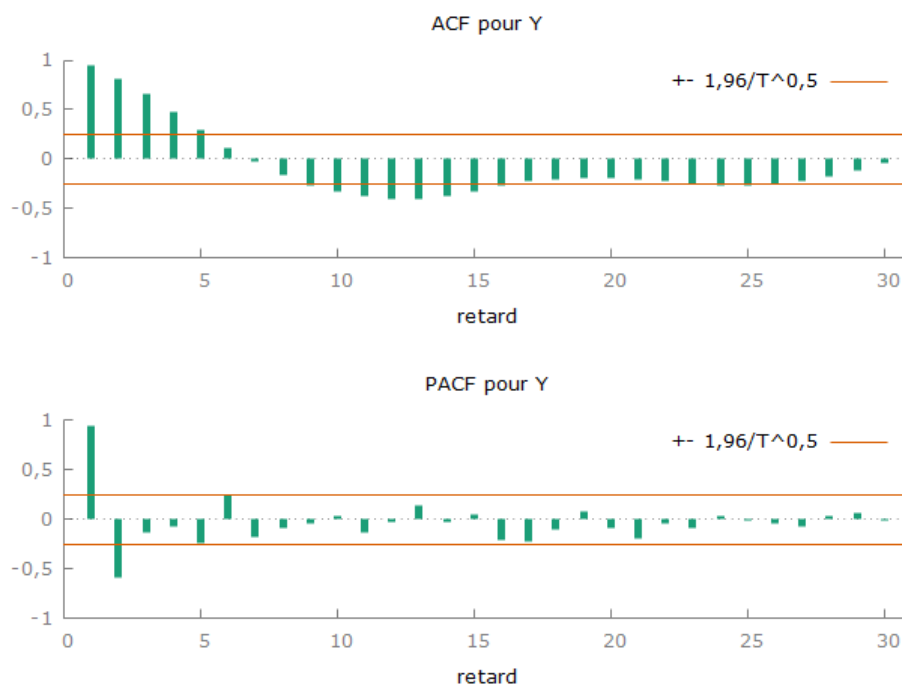
Le corrélogramme est une représentation graphique qui met en évidence une corrélation entre deux séries, dans le cas d'une étude de série temporelle les corrélations s'effectuent en fonction de la même série mais avec un retard. Avec ce graphique il sera permis de choisir entre les deux processus : $AR(p)$ et $MA(q)$. Pour cela, lors de la mise en place du corrélogramme il apparaît deux graphiques distincts, la fonction d'autocorrélation (FAC ou ACF en anglais) et la fonction d'autocorrélation partielle (PFAC ou PACF en anglais) :

- L'ACF correspond à la corrélation entre les valeurs de la série séparée de k intervalles avec un décalage k (ou retard).
- La PACF correspond à la corrélation, au décalage k , entre les valeurs de séries séparées par k intervalles, compte tenu des valeurs des intervalles intermédiaires.

Sur le graphique 9, on remarque sur l'ACF une décroissance puis une sorte de stagnation dans l'intervalle de significativité, alors que sur la PACF on remarque une rupture au retard 2, le modèle à choisir pour cette série serait un processus $AR(2)$, si la rupture était au retard 1 le processus serait $AR(1)$.

Si à l'inverse la rupture s'effectue sur l'ACF avec un retard 1 et la décroissance sur la PACF, alors le processus utilisé est $MA(1)$. Ou encore s'il y a rupture sur l'ACF avec un retard 2 et rupture avec un retard 4, les processus utilisés sont $AR(4)$ et $MA(2)$.

Graphique 9 : Corrélogramme d'une Série Y



Source : Gretl

Processus Moyenne Mobile, MA(q)

Le processus MA(q) indique que l'erreur de régression est une combinaison linéaire de termes d'erreur dont les valeurs se sont produites simultanément et à différents moments dans le passé. Ce modèle est toujours stationnaire.

$$Z_t = \mu + a_t - \sum_{i=1}^q \theta_i a_{t-i}$$

Avec :

- μ : constante
- θ : paramètre estimer par le logiciel
- a_t : bruit blanc
- q : nombre de d'autocorrélation significatives dans l'ACF, ce qui correspond par la suite au nombre de paramètre à estimer pour ce processus

Processus AutoRegressif AR(p)

Le processus AR(q) indique que la variable d'intérêt évolutive est régressée sur ses propres valeurs décalées. Ce modèle n'est pas toujours stationnaire.

$$Z_t = \delta + a_t + \sum_{j=1}^p \Phi_j Z_{t-j}$$

Avec :

- δ : constante
- a_t : bruit blanc
- p : nombre de retard à la rupture, ce qui correspond par la suite au nombre de paramètre à estimer pour ce processus
- Φ : paramètre à estimer

Construction du modèle Box-Jenkins

Le modèle de Box-Jenkins est une technique de prévision mis en place en 1970 par Box et Jenkins, qui permet de prédire une série chronologique à partir du modèle ARIMA. Elle est constituée par trois étapes : l'identification, l'estimation et la vérification. Au préalable de ces trois étapes, la série doit être stationnaire afin que l'hypothèse : la série est généré aléatoirement soit validé. On peut donc parler de quatre étapes afin d'affecter ce modèle

itératif à toutes les séries chronologiques, pour vérifier la stationnarité elle s'effectuera lors de l'étape d'identification.

- L'identification : Cette étape consiste tout d'abord à rendre stationnaire la série à l'aide des étapes vu précédemment, puis à identifier à l'aide du corrélogramme le modèle ARIMA ou SARIMA à proposer.
- L'estimation : Lors de cette étape on applique le modèle identifier préalablement, et on vérifie la significativité des paramètres, la stationnarité des AR (ou SAR pour un modèle SARIMA) et l'invisibilité des paramètres des MA (ou SMA pour SARIMA). Si ces hypothèses ne sont pas vérifiées, une nouvelle identification des paramètres est obligatoire.
- La vérification : Cette dernière étape consiste à vérifier si le modèle estimé reproduit le modèle qui a engendré la série. Pour cela, il faut s'intéresser aux résidus du modèles estimé, et vérifier s'ils se comportent comme un bruit blanc, deux étapes sont nécessaires :
 - Test de Bartlett : Ce test unitaire permet de vérifier si les résidus suivent une loi normale. Pour analyser les résidus, il faut créer le corrélogramme de ces derniers et établir si les corrélations sont toutes comprises dans l'intervalle de significativité.
 - Test de Portmanteau : Ce test global permet de vérifier sur l'ensemble des résidus, et non un à un contrairement au test de Bartlett, si les résidus se comportent comme des bruits blancs, c'est à dire de vérifier l'hypothèse d'indépendance des résidus. Pour cela, on s'intéresse à un grand groupe de résidus k , et on analyse la statistique Q :

$$Q = T \sum_{k=1}^k \left(\frac{1}{T-k} \right) \hat{\rho}_k(\hat{a})$$

Pour vérifier l'hypothèse d'indépendance, il faut que Q soit supérieur à 0.05 afin d'accepter H_0 au seuil de risque 5%.

Si un des deux tests cités précédemment n'est pas valide, alors il faut recommencer l'étape d'identification.

Une fois toutes les étapes validées on peut une prévision à n'importe quel horizon, même si plus l'horizon est grand plus l'erreur est grande, car l'horizon $l=2$ utilise la prévision à horizon $l=1$.

Modèle Multiplicatif

Le modèle multiplicatif est l'utilisation d'un processus ARIMA et d'un modèle SARIMA, il s'écrit : $ARIMA(p,d,q) \times (P,D,Q)_S$ avec S la composante saisonnière (fréquence de la série). Le modèle multiplicatif est une décomposition en deux sous-modèles, ARIMA et SARIMA, et

ensuite les multiplie entre eux. Il est donc composé d'une partie classique et d'une composante saisonnière.

$$[Y_t = (1 - B)^d (1 - Bs)^D X_t - \mu ; \Phi(B)\Theta(Bs) Y_t = \theta(B) \Theta(Bs) Z_t, Z_t \sim N(0, \sigma^2)]$$

Pour un modèle ARIMA(1,0,0)(1,0,0) S=4 :

- Modèle Classique AR(1) :
$$e_t = \frac{a_t}{1 - \phi_1 B}$$
- Modèle Saisonnier SAR(1) :
$$Z_t = e_t + \Phi Z_{t-4}$$
- Modèle Multiplicatif :
$$Z_t = a_t + \phi + Z_{t-1} + \Phi Z_{t-4} - \Phi_1 \phi_1 Z_{t-5}$$

Choix du modèle ARIMA pour le SPGSCI

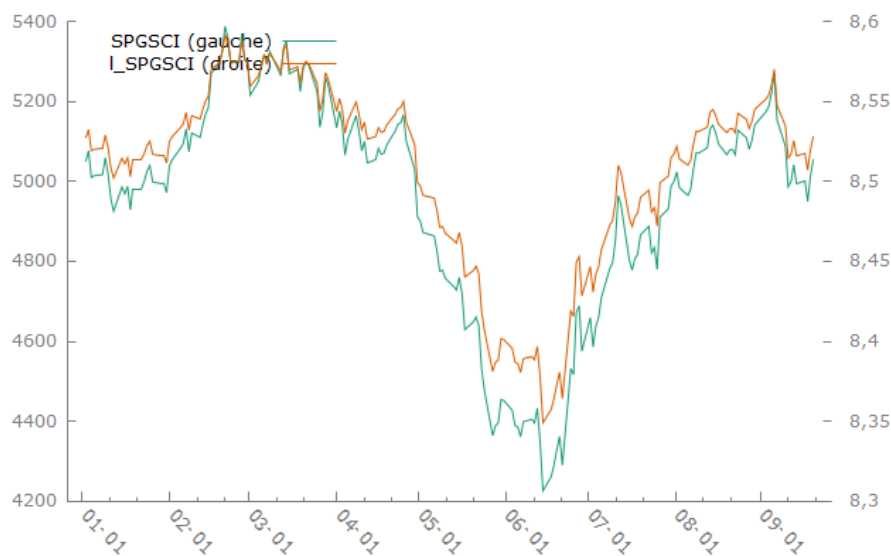
Stationnarité

Comme vu précédemment, il est nécessaire de vérifier la stationnarité de la série chronologique pour cela il sera essentiel de la vérifier graphiquement puis à l'aide du Test de Dickey-Fuller augmenté afin de s'en assurer.

Pour une série de cours boursier, il est obligatoirement nécessaire de la transformer logarithmiquement puis de la différencier afin d'obtenir les variations journalières. Il reste à le vérifier à l'aide du logiciel Gretl.

Le graphique 10 permet de remarquer que les fluctuations de la série transformé logarithmiquement correspondent aux mêmes mouvements que la série brute, la série semble être stationnaire au niveau de la variance. Cependant il reste à déterminer graphiquement si elle est stable au niveau de la moyenne.

Graphique 10 : Série Brute et Série Transformer Logarithmiquement



Source : Dossier

Dans le graphique suivant, une différenciation d'ordre 1 a été effectuée sur la série brute, elle semble être stationnaire au niveau de la moyenne. Afin de vérifier ceci, il sera effectué un Test de Dickey-Fuller puis à l'aide du corrélogramme qui seront étudiés dans la partie suivante.

La figure 1 permet de dire que la série différenciée à l'ordre 1 est stationnaire puisque les p.value sont toutes inférieures à 0.01, donc on rejette H_0 et accepte H_1 au seuil de risque 1%, de plus le fait d'ajouter des indicatrices saisonnières permet d'annoncer qu'un processus SARIMA est à mettre en place par la suite.

Figure 1 : Test de Dickey-Fuller sur la série SPGSCI différencié à l'ordre 1 avec ajout d'indicatrice saisonnière

```
Test de Dickey-Fuller augmenté pour d_SPGSCI
test à reculons à partir de 14 retards, suivant le critère AIC
taille de l'échantillon 186
hypothèse nulle de racine unitaire : a = 1

test avec constante avec des indicatrices saisonnières
avec 0 retards de (1-L)d_SPGSCI
modèle: (1-L)y = b0 + (a-1)*y(-1) + e
valeur estimée de (a - 1): -0,993408
statistique de test: tau_c(1) = -13,3056
p. critique 5,637e-022
Coeff. d'autocorrélation du 1er ordre pour e: 0,001

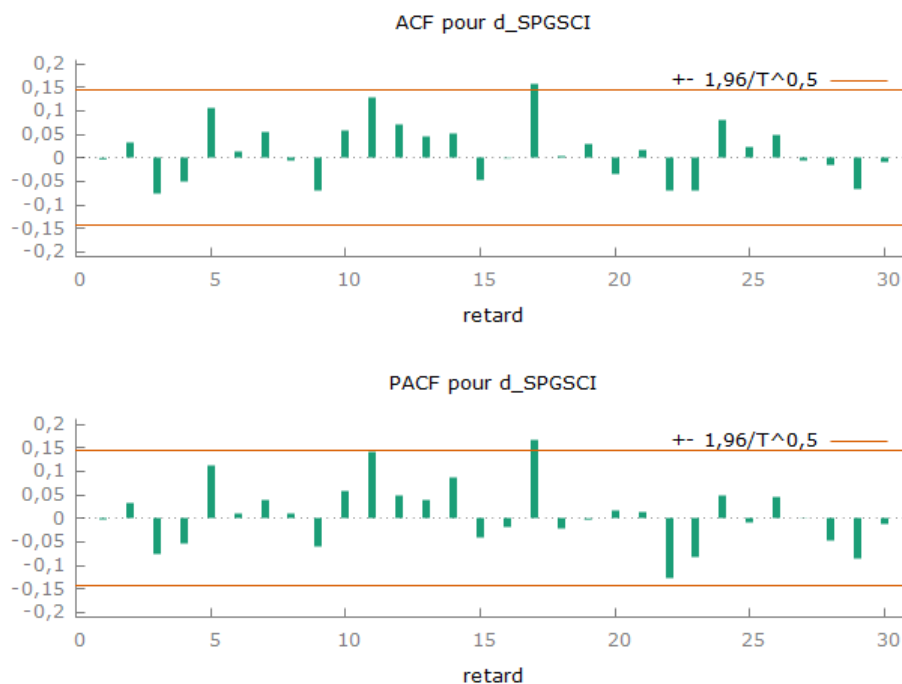
avec constante et tendance temporelle avec des indicatrices saisonnières
avec 0 retards de (1-L)d_SPGSCI
modèle: (1-L)y = b0 + b1*t + (a-1)*y(-1) + e
valeur estimée de (a - 1): -0,994465
statistique de test: tau_ct(1) = -13,2867
p. critique 2,135e-021
Coeff. d'autocorrélation du 1er ordre pour e: 0,001
```

Source : Dossier

Identification, Estimation, Vérification

A l'aide du corrélogramme suivant, aucune décroissance n'est remarquée et aucune rupture n'est remarqué. Pour la partie classique ARIMA, aucun processus que cela soit AR(p) soit MA(q) n'est à mettre en place.

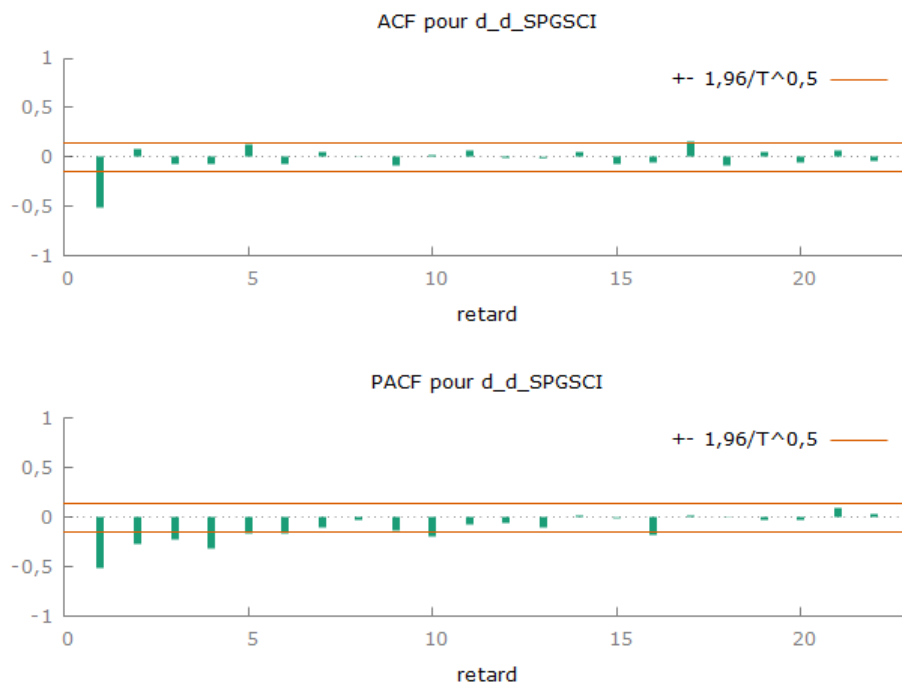
Graphique 11 : Corrélogramme de la série SPGSCI différencié à l'ordre 1



Source : Dossier

On peut donc envisager d'effectuer une deuxième différenciation afin de travailler sur la partie classique comme le montre le graphique suivant. Avec une rupture au premier retard dans l'ACF et une décroissance dans la PACF, un processus MA(1) est envisagé.

Graphique 12 : Corrélogramme de la série SPGSCI différencié à l'ordre 2



Source : Dossier

Plusieurs modèles sont envisagés à partir de la transformation logarithmique car au vu de plusieurs tests effectués les modèles en log sont de meilleurs qualités (Critère AIC et R^2) : un modèle $ARIMA(0,1,0)(1,0,0)S=5$, $ARIMA(0,2,1)(0,0,0)S=5$ et $ARIMA(0,2,1)(1,0,0)S=5$; aucun de ces modèles ne comportent une constante. On comparera les résultats obtenus de ces trois modèles à partir de l'écart-type, le R^2 et le critère AIC.

$ARIMA(0,1,0)(1,0,0)S=5$

Le tableau suivant montre le modèle ARIMA estimé, le paramètre estimé est significatif au seuil de risque 10%. Ce modèle à un R^2 de 0.964, un écart-type de 0.0108 et un critère AIC de -1155.828.

Tableau 1 : Résultat du modèle ARIMA (0,1,0)(1,0,0)S=5

Evaluations de la fonction : 13

Évaluations du gradient : 3

Modèle 8: ARIMA, utilisant les observations 2012-01-04:2012-09-20 (T = 187)

Estimation par AS 197 (MV exact)

Variable dépendante: (1-L) 1_SPGSCI

Écart-types basés sur la matrice hessienne

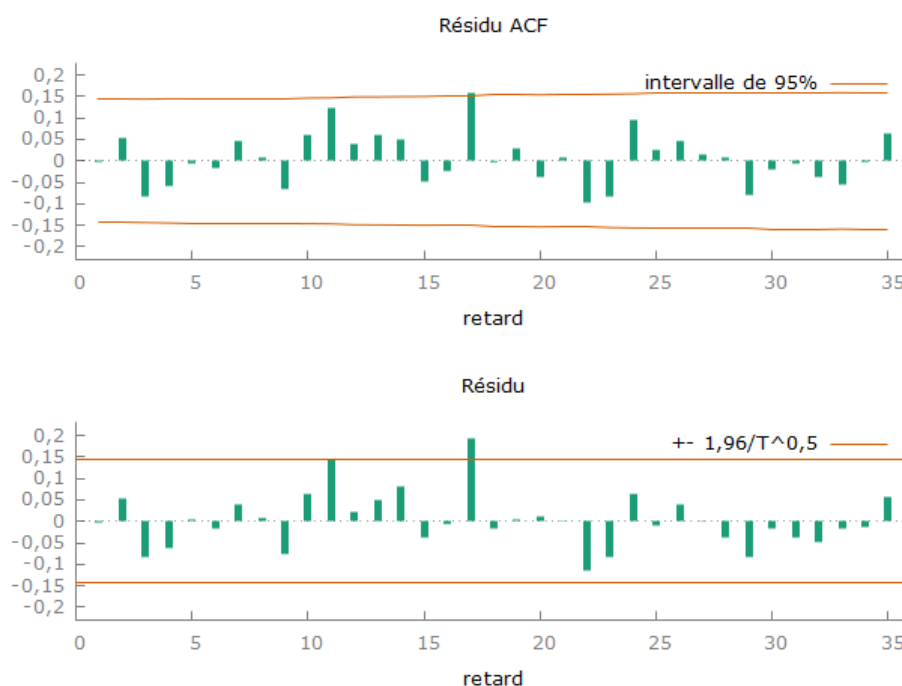
	coefficient	éc. type	z	p. critique	
-----	-----	-----	-----	-----	-----
Phi_1	0,124095	0,0726935	1,707	0,0878	*
Moyenne var. dép.	5,98e-06	Éc. type var. dép.		0,011002	
Moyenne innovations	6,92e-06	Éc. type innovations		0,010886	
R2	0,964908	R2 ajusté		0,964908	
Log de vraisemblance	579,9142	Critère d'Akaike		-1155,828	
Critère de Schwarz	-1149,366	Hannan-Quinn		-1153,210	
	Réel	Imaginaire	Modulo	Fréquence	
-----	-----	-----	-----	-----	-----
AR (saisonnier)					
Ordre 1	8,0584	0,0000	8,0584	0,0000	
-----	-----	-----	-----	-----	-----

Source : Dossier

Il faut maintenant vérifier si ce modèle est viable à l'aide d'une analyse des résidus de ce modèle.

Le corrélogramme des résidus ci-dessous montre que pour le retard 17, il y a un risque d'autocorrélation puisqu'il apparaît comme étant significatif. Cependant la Q stat du 35ème retard est de 0.862, il est donc possible de garder ce modèle mais si un meilleur modèle est possible il sera accepté.

Graphique 13 : Corrélogramme des résidus du modèle ARIMA (0,1,0)(1,0,0)S=5



Source : Dossier

Figure 2 : Q Statistique des résidus du modèle ARIMA (0,1,0)(1,0,0)S=5

33	-0,0564	-0,0181	24,3149	[0,833]
34	-0,0041	-0,0137	24,3188	[0,863]
35	0,0629	0,0583	25,2377	[0,862]

Source : Dossier

ARIMA(1,2,0)(0,0,0)S=5

Le tableau suivant montre le modèle ARIMA estimé, le paramètre estimé est significatif au seuil de risque 1%, il est donc hautement significatif. Ce modèle à un R^2 de 0.9648, un écart type de 0.011 et un critère AIC de -1140.551.

Tableau 2 : Résultat du modèle ARIMA (1,2,0)(0,0,0)S=5

Modèle 15: ARIMA, utilisant les observations 2012-01-05:2012-09-20 (T = 186)
 Estimation par AS 197 (MV exact)
 Variable dépendante: $(1-L)^2 1_SPGSCI$
 Écart-types basés sur la matrice hessienne

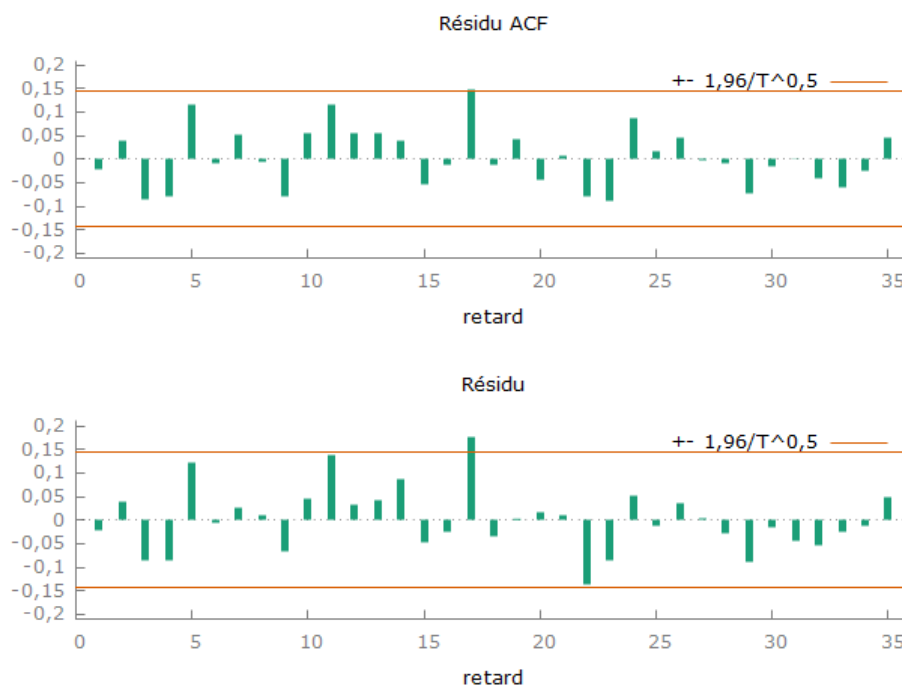
	coefficient	éc. type	z	p. critique	
theta_1	-0,975590	0,0583743	-16,71	1,06e-062	***
Moyenne var. dép.	0,000015	Éc. type var. dép.	0,015678		
Moyenne innovations	0,000138	Éc. type innovations	0,011066		
R2	0,964840	R2 ajusté	0,964840		
Log de vraisemblance	572,2754	Critère d'Akaike	-1140,551		
Critère de Schwarz	-1134,099	Hannan-Quinn	-1137,936		

	Réel	Imaginaire	Modulo	Fréquence
MM				
Ordre 1	1,0250	0,0000	1,0250	0,0000

Source : Dossier

Afin de vérifier la viabilité de ce modèle, la mise en place du corrélogramme des résidus permet d'observer un risque d'autocorrélation pour le retard 17, il faut donc étudier dans la globalité la qualité du modèle, en analysant les résidus à l'aide du test de Portmanteau, la Q statistique qui apparaît dans la figure 3 est de 0.819, ce modèle est donc viable.

Graphique 14 : Corrélogramme des résidus du modèle ARIMA (0,2,1)(0,0,0)S=5



Source : Dossier

Figure 3 : Q Statistique des résidus du modèle ARIMA (0,2,1)(0,0,0)S=5

33	-0,0596	-0,0251	25,7982	[0,772]
34	-0,0254	-0,0133	25,9466	[0,804]
35	0,0464	0,0493	26,4453	[0,819]

Source : Dossier

ARIMA(0,2,1)(1,0,0)S=5

Le tableau suivant montre le modèle ARIMA multiplicatif estimé, le paramètre estimé SAR(1) est significatif au seuil de risque 10%, le paramètre estimé MA(1) est significatif au seuil de risque 1% il est hautement significatif contrairement au paramètre cité précédemment. Ce modèle à un R^2 de 0.9651, un écart type de 0.0109 et un critère AIC de -1141,661.

Tableau 3 : Résultat du modèle ARIMA (1,2,0)(0,0,0)S=5

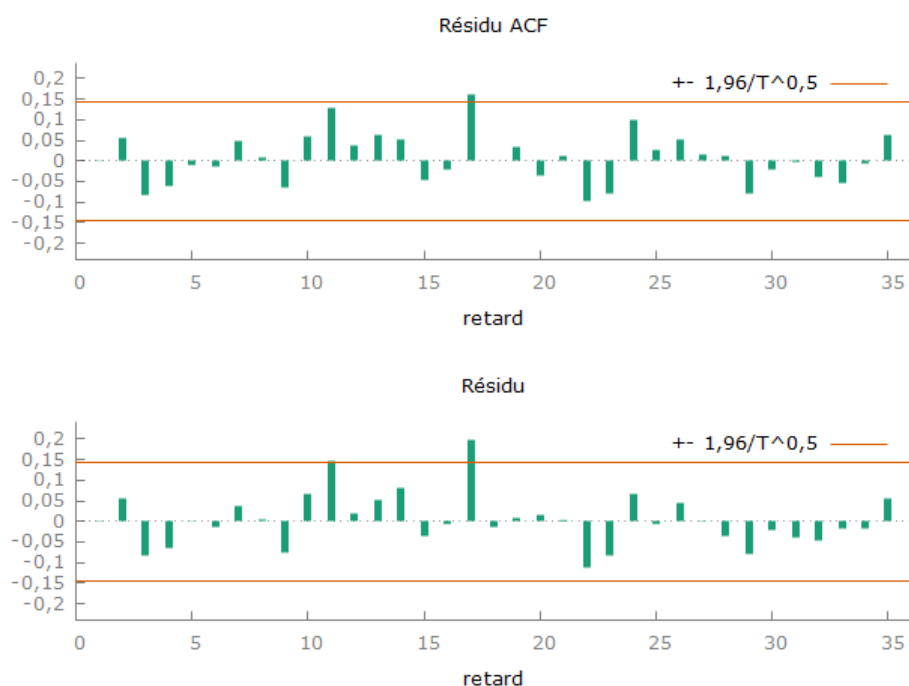
Modèle 17: ARIMA, utilisant les observations 2012-01-05:2012-09-20 (T = 186)
 Estimation par AS 197 (MV exact)
 Variable dépendante: (1-L)^2 1_SPGSCI
 Écart-types basés sur la matrice hessienne

	coefficient	éc. type	z	p. critique	
Phi_1	0,129995	0,0731199	1,778	0,0754	*
theta_1	-0,999999	0,0281561	-35,52	2,76e-276	***
Moyenne var. dép.	0,000015	Éc. type var. dép.	0,015678		
Moyenne innovations	0,000128	Éc. type innovations	0,010915		
R2	0,965333	R2 ajusté	0,965145		
Log de vraisemblance	573,8307	Critère d'Akaike	-1141,661		
Critère de Schwarz	-1131,984	Hannan-Quinn	-1137,740		
	Réel	Imaginaire	Modulo	Fréquence	
AR (saisonnier)					
Ordre 1	7,6926	0,0000	7,6926	0,0000	
MM					
Ordre 1	1,0000	0,0000	1,0000	0,0000	

Source : Dossier

Afin de vérifier la viabilité de ce modèle, la mise en place du corrélogramme des résidus permet d'entrevoir un risque d'autocorrélation pour le retard 17, il faut donc étudier dans la globalité la qualité du modèle, en analysant les résidus à l'aide du test de Portmanteau, la Q statistique qui apparaît dans la figure 3 est de 0.812, ce modèle est donc viable.

Graphique 14 : Corrélogramme des résidus du modèle ARIMA (0,2,1)(0,0,0)S=5



Source : Dossier

Figure 4 : Q Statistique des résidus du modèle ARIMA (0,2,1)(1,0,0)S=5

33	-0,0549	-0,0188	24,8410	[0,775]
34	-0,0050	-0,0188	24,8469	[0,812]
35	0,0617	0,0556	25,7290	[0,812]

Source : Dossier

Choix du Modèle

Maintenant il faut choisir entre les trois modèles estimés antérieurement à l'aide des trois propriétés suivantes : R^2 , écart-type d'innovation, la Q statistique et critère Akaike (AIC) ; les deux derniers critères sont à minimiser contrairement aux deux autres qui doit être maximal. Pour cela, il faut donc les comparer.

Tableau 4 : Comparaison des Modèles Estimés

Modèle ARIMA	R^2	Ecart-Type	Critère AIC	Q Stat
(0,1,0)(1,0,0)S=5	0.9649	0.01088	-1155	0.862
(0,2,1)(0,0,0)S=5	0.9648	0.01106	-1140	0.819
(0,2,1)(1,0,0)S=5	0.9653	0.01091	-1141	0.812

Au vu des différents résultats, les modèles sont plus ou moins semblables même s'il apparaît que le premier modèle peut être préféré ; puisque son critère AIC est le plus faible tout comme son écart-type d'innovation et sa Q Stat est la plus grande, les résidus sont donc le moins auto corrélés.

C'est donc le premier modèle qui sera utiliser pour faire nos prévisions.

Prévision à une période

Le 28/09/2012 le SPGSCI vaut 5054.80\$, dans la partie qui suit il sera calculé à l'aide du modèle ARIMA choisis précédemment avec $Y_t = 5013.3032$.

En appliquant la fonction forecast sur R, la prévision pour Y_{t+1} est de 5017.751 et est comprise entre 4882.194 et 5153.307.

On a donc un taux d'erreur de 0.82%, la prévision est donc de très bonne qualité.

En appliquant les formules étudiées en cours :

$$Z_t = 63.7412$$

$$Z_{t+1} = 0.124095 * 63.7412$$

$$Z_{t+1} = 7.91$$

$$Y_{t+1} = 7.91 + 4993.6411$$

$$Y_{t+1} = 5001.55$$

Les résultats trouvés de manières écrite et avec le logiciel Gretl diffèrent des résultats trouvés avec le logiciel R, l'erreur étant plus grande. Je pense ne pas avoir utilisé la bonne méthode manuscrite pour calculer cette prévision.

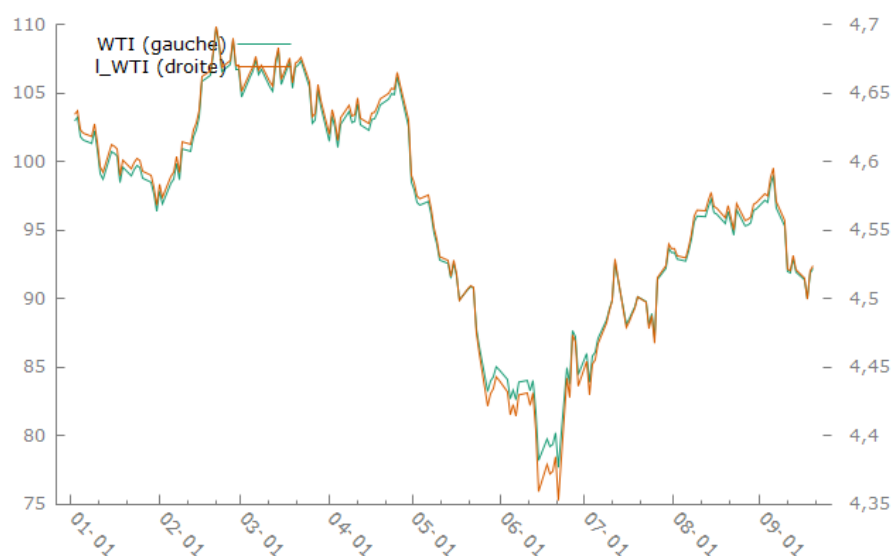
Analyse des Variables WTI et S&P 500

Dans cette partie, il sera étudié les variables WTI et S&P 500 afin d'établir pour ces séries chronologiques un ou des modèles ARIMA.

WTI

Au vu du graphique 15, la transformation logarithmique ne change rien aux mouvements de la série chronologique, elle est donc stationnaire au niveau de la variance. Il reste donc à la différencier afin que la série soit stationnaire au niveau de la moyenne puisqu'un test AFD ou un regard sur le corrélogramme confirme la non stationnarité à l'espérance (cf : annexe 1 et 2).

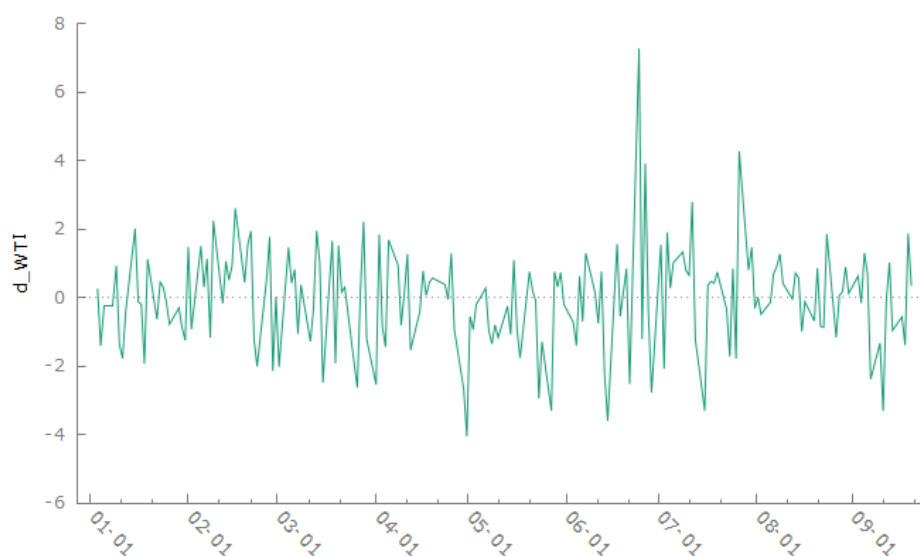
Graphique 15 : Représentation graphique de la série brute et de son logarithme



Source : Dossier

Lorsqu'une différenciation d'ordre 1 est effectuée sur la série brute, la série apparaît comme étant stationnaire au niveau de la moyenne (graphique 16), afin de vérifier ceci il faut effectuer un Test AFD (cf : annexe 3), qui nous indique par ailleurs que celle-ci est stationnaire lorsqu'elle est différenciée à l'ordre 1.

Graphique 16 : Série WTI différencié à l'ordre 1

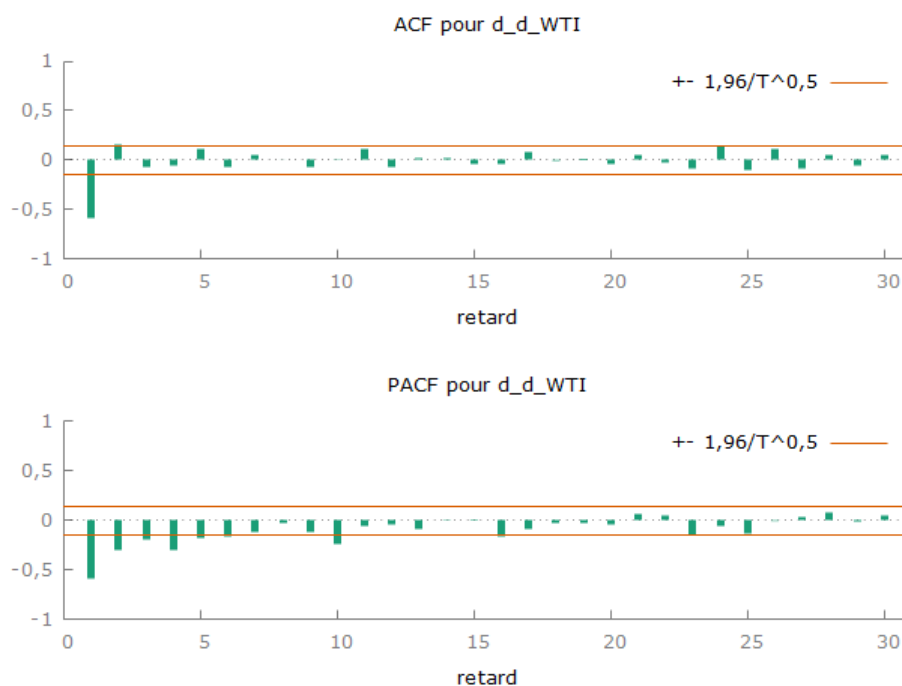


Source : Dossier

Cependant lors de l'étude du corrélogramme de la série différencié (cf : annexe 4), on ne voit apparaître aucune corrélation c'est pourquoi, une deuxième différenciation est nécessaire afin de pouvoir étudier correctement le corrélogramme et d'estimer un modèle.

Le graphique 17 permet de remarquer une décroissance dans la PACF et une rupture au retard 1 sur l'ACF, c'est pourquoi un processus MA(1) est envisagé.

Graphique 17 : Corrélogramme de la série WTI différencié à l'ordre 2



Source : Dossier

Le tableau 5 permet de remarquer que le paramètre MA estimé est parfaitement significatif, la constante a été enlevé car elle n'est pas significative et produit une augmentation du critère AIC or il faut le minimiser.

Tableau 5 : Résultat du Modèle ARIMA(0,2,1) pour la série WTI

Modèle 24: ARIMA, utilisant les observations 2012-01-05:2012-09-20 (T = 186)
 Estimation par AS 197 (MV exact)
 Variable dépendante: (1-L)^2 WTI
 Écart-types basés sur la matrice hessienne

	coefficient	éc. type	z	p. critique	
theta_1	-1,00000	0,0220408	-45,37	0,0000	***
Moyenne var. dép.	0,000430	Éc. type var. dép.	2,216615		
Moyenne innovations	0,023421	Éc. type innovations	1,486014		
R2	0,964533	R2 ajusté	0,964533		
Log de vraisemblance	-340,2122	Critère d'Akaike	684,4245		
Critère de Schwarz	690,8760	Hannan-Quinn	687,0389		
	Réel	Imaginaire	Modulo	Fréquence	
MM					
Ordre 1	1,0000	0,0000	1,0000	0,0000	

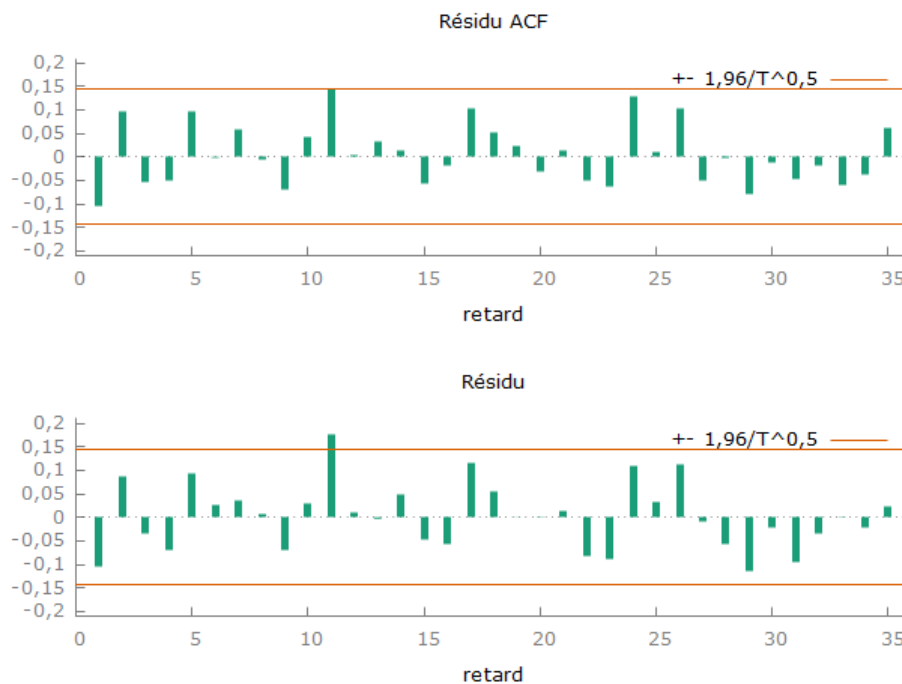
Source : Dossier

Il faut maintenant étudier le corrélogramme des résidus afin de vérifier si le modèle peut être utilisé pour faire des estimations.

Le graphique 18 permet d'observer que seulement au retard 11 il y a un problème d'autocorrélation, en étudiant d'un point de vue global (Test de Portmanteau) il apparaît, sur la figure 5, une Q statistique de 0.712 le modèle est donc globalement viable.

Ce modèle est donc gardé afin de faire des prévisions sur la série WTI.

Graphique 18 : Corrélogramme des résidus du modèle ARIMA (0,2,1) pour la série WTI



Source : Dossier

Figure 5 : Q Statistique des résidus du modèle ARIMA (0,2,1) pour la série WTI

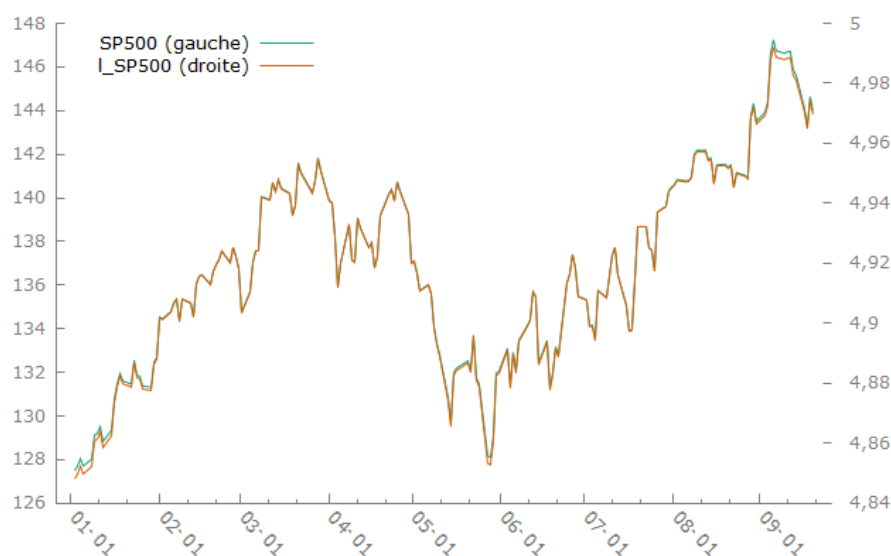
33	-0,0608	0,0010	27,8126	[0,679]
34	-0,0374	-0,0212	28,1337	[0,708]
35	0,0604	0,0236	28,9799	[0,712]

Source : Dossier

S&P 500

Au vu du graphique 19, la transformation logarithmique ne change rien aux mouvements de la série chronologique, elle est donc stationnaire au niveau de la variance. Il reste donc à la différencier afin que la série soit stationnaire au niveau de la moyenne puisqu'un test AFD ou un regard sur le corrélogramme confirme la non stationnarité à l'espérance (cf : annexe 1 et 2).

Graphique 19 : Représentation graphique de la série brute S&P 500 et de son logarithme

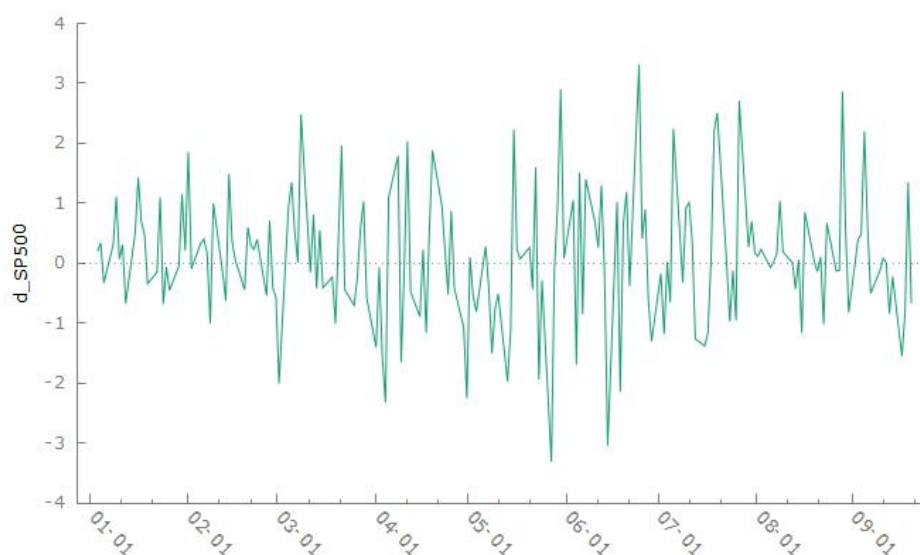


Source : Dossier

Lorsqu'une différenciation d'ordre 1 est effectuée sur la série brute, la série apparaît comme étant stationnaire au niveau de la moyenne (graphique 20), afin de vérifier ceci il faut effectuer un Test AFD (cf : annexe 5), qui nous indique par ailleurs que celle-ci est stationnaire lorsqu'elle est différenciée à l'ordre 1.

La série différenciée à l'ordre 1 sera donc utiliser par la suite afin d'étudier les corrélogrammes et estimer un modèle à partir de ces derniers.

Graphique 20 : Série S&P500 différencié à l'ordre 1

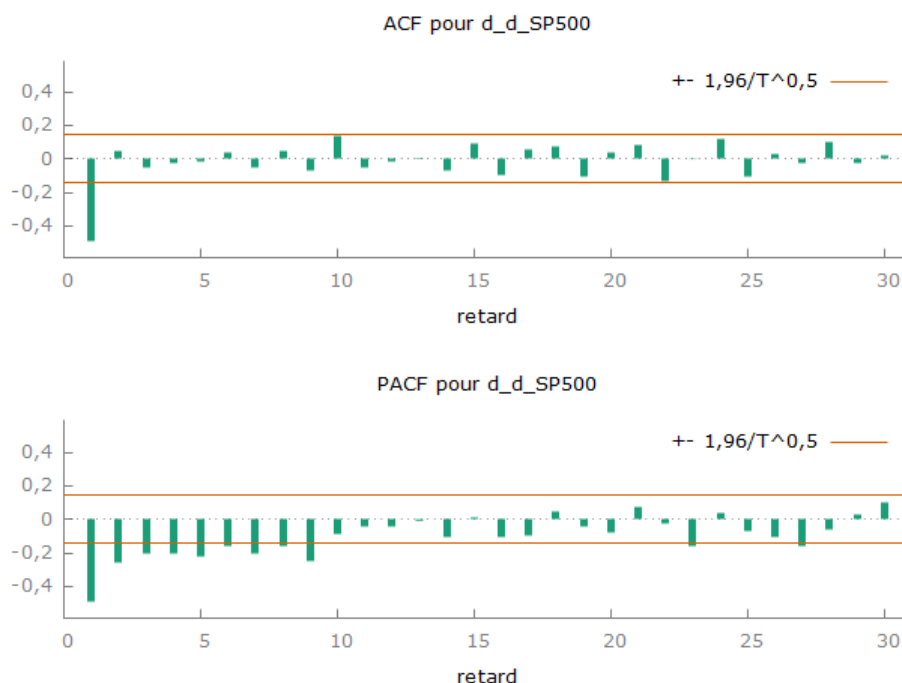


Source : Dossier

Cependant lors de l'étude du corrélogramme de la série différencié (cf : annexe 7), on ne voit apparaître aucune corrélation c'est pourquoi, une deuxième différenciation est nécessaire afin de pouvoir étudier correctement le corrélogramme et d'estimer un modèle.

Le graphique 21 permet de remarquer une décroissance dans la PACF et une rupture au retard 1 sur l'ACF, c'est pourquoi un processus MA(1) est envisagé.

Graphique 21 : Corrélogramme de la série S&P 500 différencié à l'ordre 2



Source : Dossier

Le tableau 6 permet de remarquer que le paramètre MA estimé est parfaitement significatif, la constante a été enlevée car elle n'est pas significative et produit une augmentation du critère AIC or il faut le minimiser, tout comme l'écart-type d'innovation.

Tableau 6 : Résultat du Modèle ARIMA(0,2,1) pour la série WTI

Modèle 1: ARIMA, utilisant les observations 2012-01-05:2012-09-20 (T = 186)
 Estimation par AS 197 (MV exact)
 Variable dépendante: (1-L)^2 SP500
 Écart-types basés sur la matrice hessienne

	coefficient	éc. type	z	p. critique	
theta_1	-1,00000	0,0192677	-51,90	0,0000	***
Moyenne var. dép.	-0,004677	Éc. type var. dép.		1,529554	
Moyenne innovations	-0,052781	Éc. type innovations		1,087649	
R2	0,941651	R2 ajusté		0,941651	
Log de vraisemblance	-282,1656	Critère d'Akaike		568,3313	
Critère de Schwarz	574,7828	Hannan-Quinn		570,9456	
	Réel	Imaginaire	Modulo	Fréquence	
MM					
Ordre 1	1,0000	0,0000	1,0000	0,0000	

Source : Dossier

Il faut maintenant étudier le corrélogramme des résidus afin de vérifier si le modèle peut être utilisé pour faire des estimations.

Le graphique 22 permet d'observer que seulement au retard 11 il y a un problème d'autocorrélation, en étudiant d'un point de vue global (Test de Portmanteau) il apparaît, sur la figure 5, une Q statistique de 0.712 le modèle est donc acceptable au seuil de risque 5% cependant un doute peut être émis concernant sa qualité de prédiction.

Ce modèle est donc gardé afin de faire des prévisions sur la série S&P 500.

Graphique 22 : Corrélogramme des résidus du modèle ARIMA (0,2,1) pour la série S&P 500



Source : Dossier

Figure 5 : Q Statistique des résidus du modèle ARIMA (0,2,1) pour la série S&P 500

33	-0,0167	0,0147	41,0165	[0,132]
34	0,0702	0,0360	42,1510	[0,132]
35	0,1315	0,1096	46,1552	[0,080]

Source : Dossier

Test de cointégration selon Engle-Granger

Dans leur papier fondateur de 1987 Engle et Granger définissent la cointégration entre série temporelle par :

Les éléments du vecteur de variables aléatoires $X_t = (x_{1,t}, \dots, x_{k,t})$ sont dits cointégrés d'ordres (δ, b) , noté $CI \sim (\delta, b)$ si

a) tous les éléments du vecteur X_t sont $I(\delta)$

b) il existe un vecteur $\beta \neq 0$ tel que $\varepsilon_t = \beta' X_t \sim I(\delta - b)$ avec $b > 0$

Source : Cours de Gilles de Truchis

Cette relation permet de déterminer une relation de long terme et court terme non fallacieuse entre des séries temporelles. C'est pourquoi dans cette partie, il sera étudié les trois séries chronologiques du dossier par couple afin de déterminer une relation entre chacun.

Tout d'abord, il est nécessaire de travailler ayant le même nombre de différenciation mais pas stationnaire, car pour qu'une relation ne soit pas fallacieuse il est nécessaire que le test de racine unitaire ADF soit accepté sur les variables mais rejeté sur les résidus de la relation. Puisque la cointégration permet de rendre stationnaires les résidus de la relation entre les deux variables si elles suivent tendance stochastique commun.

Le travail qui suit sera effectué couple par couple.

Couple SPGSCI et WTI

Relation de corrélation à long terme

$$Y_t = \alpha + \beta X_t + \varepsilon_t$$

Afin d'obtenir la relation de long terme entre ce couple de variable, la méthode des Moindres Carrés Ordinaires (MCO) est utilisée sur nos séries brutes afin de créer la relation, puis il sera vérifié dans un premier temps la significativité de la relation mais surtout si les résidus de la relation sont stationnaires, ce qui permettra de conclure sur la cointégration de long terme du couple de variables.

Tableau 7 : Résultat de la cointégration entre le SPGSCI et le WTI

Régression de cointégration -
MCO, utilisant les observations 2013-01-03:2013-09-23 (T = 188)
Variable dépendante: SPGSCI

	coefficient	éc. type	t de Student	p. critique	
const	1809,78	100,064	18,09	7,55e-043	***
WTI	32,6914	1,03718	31,52	1,59e-076	***
Moyenne var. dép.	4953,401	Éc. type var. dép.	278,6289		
Somme carrés résidus	2289380	Éc. type régression	110,9437		
R2	0,842303	R2 ajusté	0,841455		
Log de vraisemblance	-1151,051	Critère d'Akaike	2306,103		
Critère de Schwarz	2312,576	Hannan-Quinn	2308,725		
rho	0,994606	Durbin-Watson	0,026894		

Source : Dossier

Le R^2 est de 0.84, puis la variable explicative (WTI) et la constante sont significatives au seuil de risque 1%, le modèle paraît de bonne qualité, il est nécessaire maintenant de vérifier la stationnarité des résidus afin de vérifier si la relation est fallacieuse ou non.

Test de Dickey-Fuller Augmenté sur les résidus

Tout d'abord il est obligatoire de regarder le nombre de retard du VAR qui est la condition pour laquelle le critère AIC est minimisé pour ce modèle, c'est à dire le meilleur modèle possible.

Dans le cas de cette relation de long terme, le nombre de retard est de 1, pour nos variables étudié l'ADF test est accepté cependant concernant les résidus il est accepté aussi, cela convient de dire que la relation entre ce couple de variables est fallacieuse. On ne peut donc pas étudier le modèle et l'utilisé afin d'appliquer le modèle à correction d'erreur (MCE), et donc de l'utiliser afin de faire une prédiction.

Tableau 8 : Résultats du Test de Racine Unitaire ADF sur les résidus

Test de Dickey-Fuller augmenté pour uhat
test à reculons à partir de 5 retards, suivant le critère AIC
taille de l'échantillon 187
hypothèse nulle de racine unitaire : $a = 1$

test sans constante
avec 0 retards de $(1-L)uhat$
modèle: $(1-L)y = (a-1)*y(-1) + e$
valeur estimée de $(a - 1)$: -0,00539388
statistique de test: $\tau_c(2) = -0,443535$
p. critique 0,9667
Coeff. d'autocorrélation du 1er ordre pour e: 0,085

Source : Dossier

Couple SPGSCI/SP500

Comme vu précédemment, il est nécessaire de ne pas rendre stationnaire les variables utilisées afin d'appliquer les MCO. En appliquant le test de Dickey-Fuller augmenté sur chacune des variables utilisées, l'hypothèse nulle est acceptée, les séries ne sont donc pas stationnaire et les MCO peuvent donc être appliqué.

Tout d'abord, il est nécessaire de sélection le bon nombre de retard, grâce à la sélection de retard du VAR, d'après l'annexe 9 le bon nombre de retard est de 1. La constante n'étant pas significative, elle a été enlevée du modèle estimé.

Tableau 9 : Résultat de la cointégration entre le SPGSCI et le S&P 500

Régression de cointégration - MCO, utilisant les observations 2013-01-03:2013-09-23 (T = 188) Variable dépendante: SPGSCI				
	coefficient	éc. type	t de Student	p. critique
SP500	36,1903	0,128656	281,3	1,24e-247 ***
Moyenne var. dép.	4953,401	Éc. type var. dép.	278,6289	
Somme carrés résidus	10909981	Éc. type régression	241,5412	
R2 non-centré	0,997642	R2 centré	0,248498	
Log de vraisemblance	-1297,823	Critère d'Akaike	2597,645	
Critère de Schwarz	2600,882	Hannan-Quinn	2598,957	
rho	0,976376	Durbin-Watson	0,032024	

Le R2 est de 0.24 et la variable explicative (S&P 500) est significative au seuil de risque 1%, cependant le R² est vraiment faible. De plus en appliquant l'ADF sur les résidus, l'hypothèse nulle est acceptée, donc les résidus ne sont pas stationnaires. On ne peut donc appliquer le MCE pour cette relation.

Tableau 10 : Test ADF sur les résidus du couple SPGSCI/S&P 500

```
Étape 4 : test de racine unitaire dans uhat

Test de Dickey-Fuller augmenté pour uhat
test à reculons à partir de 1 retards, suivant le critère AIC
taille de l'échantillon 187
hypothèse nulle de racine unitaire : a = 1

test sans constante
avec 0 retards de (1-L)uhat
modèle: (1-L)y = (a-1)*y(-1) + e
valeur estimée de (a - 1): -0,0236243
statistique de test: tau_nc(2) = -1,81428
p. critique 0,3141
Coeff. d'autocorrélation du 1er ordre pour e: -0,030
```

Couple WTI/S&P 500

En appliquant la sélection de retard VAR, le retard optimal est de 2 (Annexe 10), les MCO sont donc appliqué avec un retard de 1 tout comme le test ADF.

Tableau 11 : Résultat de la cointégration entre le WTI et le S&P 500

Régression de cointégration -

MCO, utilisant les observations 2013-01-03:2013-09-23 (T = 188)

Variable dépendante: WTI

	coefficient	éc. type	t de Student	p. critique	
const	40,0771	16,7756	2,389	0,0179	**
SP500	0,409818	0,122517	3,345	0,0010	***
Moyenne var. dép.	96,16053	Éc. type var. dép.	7,822165		
Somme carrés résidus	10792,59	Éc. type régression	7,617394		
R2	0,056742	R2 ajusté	0,051671		
Log de vraisemblance	-647,4767	Critère d'Akaike	1298,953		
Critère de Schwarz	1305,426	Hannan-Quinn	1301,576		
rho	0,982985	Durbin-Watson	0,027806		

Le R^2 est très faible, et appliquant l'ADF l'hypothèse nulle est donc acceptée, les résidus ne sont pas stationnaires.

Dans aucun des couples étudiés, la possibilité d'application d'un MCE est ignorée car chacune des relations est fallacieuse malgré la sélection du retard optimal VAR et l'utilisation du critère AIC.

Conclusion

Pour revenir sur l'idée de financiarisation des matières premières et le fait de co-mouvement entre les matières premières et les indices boursiers, n'implique en rien la possibilité d'une cointégration ou d'une relation entre ces variables. Il est donc impossible de conclure sur ce sujet à partir de ce dossier et des études proposées. Il aurait été préférable d'étudier la financiarisation par d'autres modèles tel que des modèles GARCH qui permettent d'étudier la volatilité et non les prix des variables.

Cependant, il est possible de prédire ces séries boursières d'une manière très précise en journalier puisque les différents modèles ARIMA proposés dans ce dossier sont tous de bonnes qualités de part leur R^2 d'une grande valeur, qui indique donc une grande explication de la variance de la variable étudié et de part leur écart-type qui est minimisé. De plus lorsqu'il est effectué la prédiction à partir de ces modèles, l'erreur n'est que de 2%.

Ces modèles auraient pu être amélioré avec la mise en place de variables explicatives, tel que les volumes échangés par jour ou l'inflation, dans le modèle ARIMA, ce qui les transformeraient en modèle XARIMA, ou augmenté la période étudiée et donc avoir un plus grand nombre d'observation.

Bibliographie

Cours

Cours sur les Séries Temporelles et le modèle ARIMA de D. GIRARD.

Cours sur les Séries Temporelles de Lardic S. & Mignon V.

Cours sur l'économétrie des séries temporelles non-stationnaires, Gilles de Truchis.

Articles

Financiarisation des marchés de matières premières et conséquences, **David Bichetti et Nicolas Maystre**

Commodity Prices and Inflation, **Fred Furlong and Robert Ingenito**

LES HAUSSES RÉCENTES DES COURS DES MATIÈRES PREMIÈRES TRADUISENT-ELLES L'ENTRÉE DANS UN RÉGIME DE PRIX PLUS ÉLEVÉS ? **Armand Colin**

Sites Internet

Régulation des Marchés de Matières Premières, **AMF**

https://www.amf-france.org/sites/default/files/contenu_simple/colloque_journee/colloque_conseil_scientifique/Colloque%20du%20Conseil%20Scientifique%20du%206%20mai%202011%20-%20Regulation%20des%20marches%20de%20matieres%20premieres%20%20quel%20cadre%20enviager%20.pdf

S&P GSCI Methodology : **S&P Dow Jones Indices : Index Methodology**

<https://www.spglobal.com/spdji/en/documents/methodologies/methodology-sp-gsci.pdf>

Définition WTI : **IG invest**

<https://www.ig.com/fr/glossaire-trading/wti-definition>

Définition SP500 : **Capital**

<https://capital.com/fr/indice-s-p-500-definition>

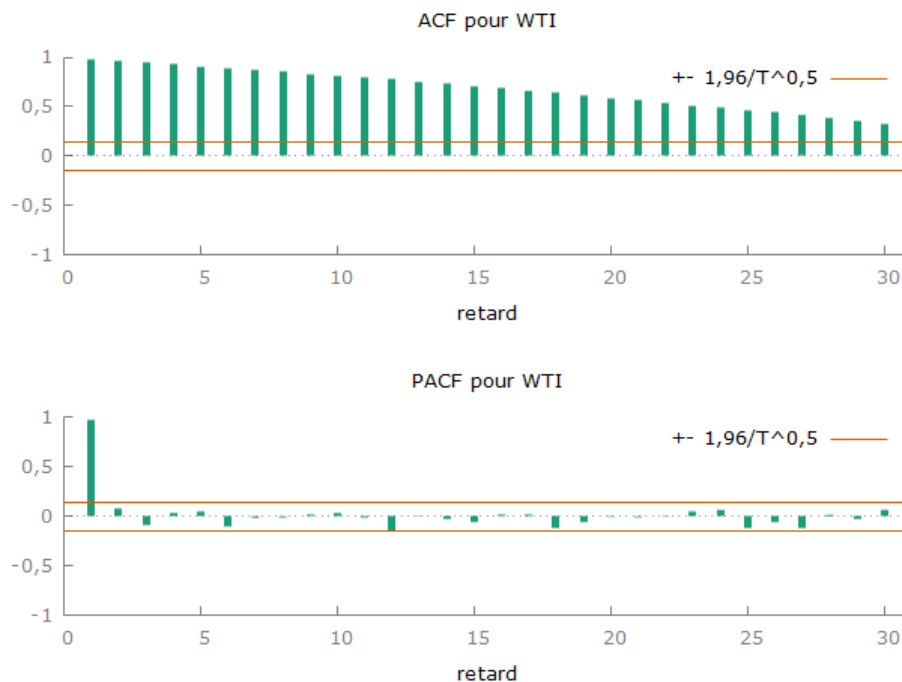
ACF et PACF : **IBM**

https://www.ibm.com/support/knowledgecenter/fr/SS3RA7_sub/modeler_mainhelp_client_ddita/components/dt/timeseries_acf_pacf.html

Annexes

Toutes les annexes présentées proviennent du dossier et ont été effectuées avec le logiciel Gretl.

Annexe 1 : Corrélogramme Série WTI non différencié



Annexe 2 : Test de Dickey-Fuller Augmenté sur la série brute WTI

Test de Dickey-Fuller augmenté pour WTI
test à reculons à partir de 14 retards, suivant le critère AIC
taille de l'échantillon 187
hypothèse nulle de racine unitaire : $a = 1$

test avec constante avec des indicatrices saisonnières
avec 0 retards de $(1-L)WTI$
modèle: $(1-L)y = b_0 + (a-1)*y(-1) + e$
valeur estimée de $(a - 1)$: -0,0188382
statistique de test: $\tau_c(1) = -1,36518$
p. critique 0,5985
Coeff. d'autocorrélation du 1er ordre pour e: -0,087

avec constante et tendance temporelle avec des indicatrices saisonnières
avec 0 retards de $(1-L)WTI$
modèle: $(1-L)y = b_0 + b_1*t + (a-1)*y(-1) + e$
valeur estimée de $(a - 1)$: -0,0254365
statistique de test: $\tau_{ct}(1) = -1,49682$
p. critique 0,8275
Coeff. d'autocorrélation du 1er ordre pour e: -0,083

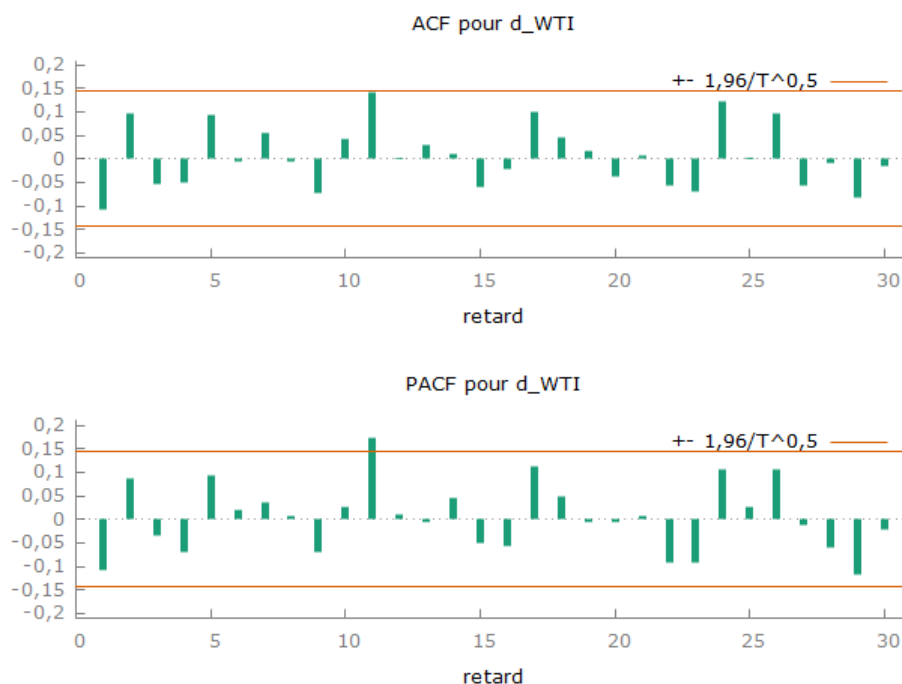
Annexe 3 : Test de Dickey-Fuller Augmenté sur la série WTI différencié à l'ordre 1

Test de Dickey-Fuller augmenté pour d_WTI
test à reculons à partir de 14 retards, suivant le critère AIC
taille de l'échantillon 186
hypothèse nulle de racine unitaire : $\alpha = 1$

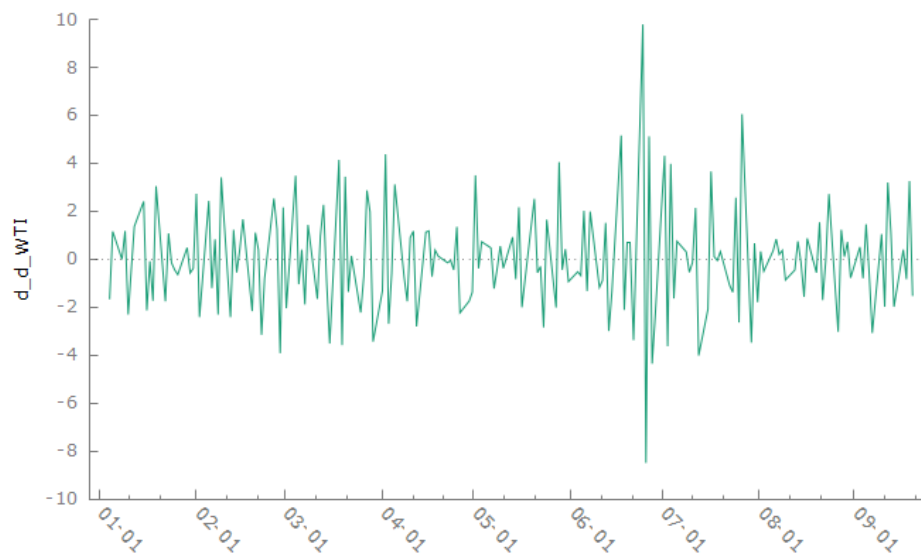
test avec constante
avec 0 retards de $(1-L)d_WTI$
modèle: $(1-L)y = b_0 + (\alpha-1)*y(-1) + e$
valeur estimée de $(\alpha - 1)$: -1,10689
statistique de test: $\tau_c(1) = -15,1$
p. critique 4,565e-024
Coeff. d'autocorrélation du 1er ordre pour e: 0,010

avec constante et tendance temporelle
avec 0 retards de $(1-L)d_WTI$
modèle: $(1-L)y = b_0 + b_1*t + (\alpha-1)*y(-1) + e$
valeur estimée de $(\alpha - 1)$: -1,10722
statistique de test: $\tau_{ct}(1) = -15,0652$
p. critique 1,05e-023
Coeff. d'autocorrélation du 1er ordre pour e: 0,010

Annexe 4 : Corrélogramme de la série WTI différencié à l'ordre 1



Annexe 5 : Représentation Graphique de la série WTI différencié à l'ordre 2



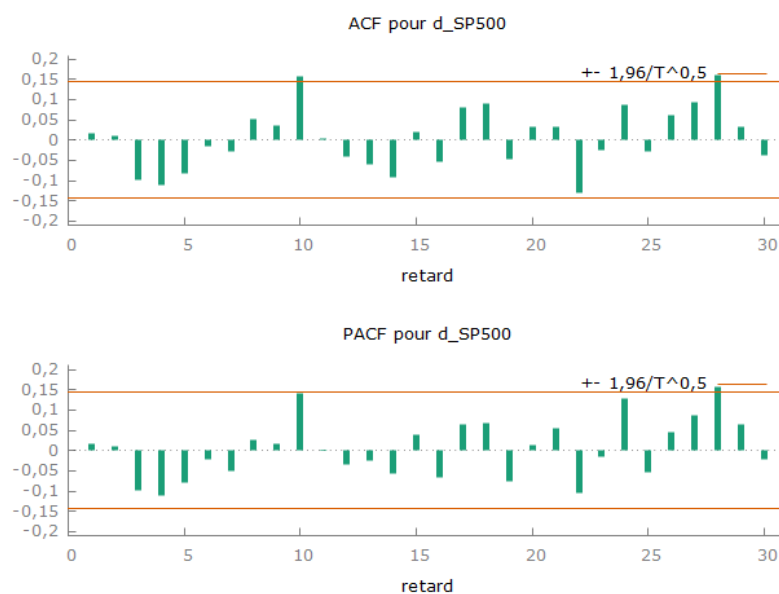
Annexe 6 : Test de Dickey-Fuller sur la série Brute S&P 500

Test de Dickey-Fuller augmenté pour SP500
test à reculer à partir de 14 retards, suivant le critère AIC
taille de l'échantillon 187
hypothèse nulle de racine unitaire : $a = 1$

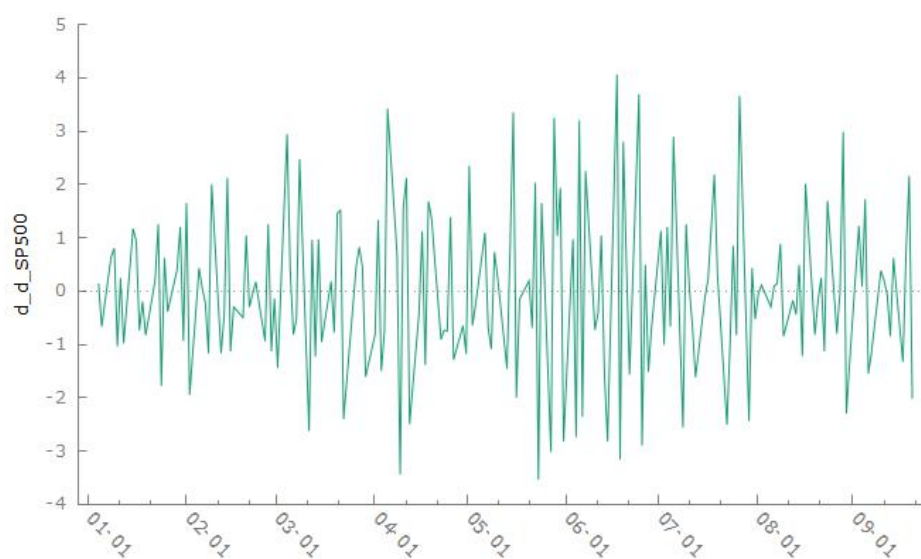
test avec constante
avec 0 retards de $(1-L)SP500$
modèle: $(1-L)y = b_0 + (a-1)*y(-1) + e$
valeur estimée de $(a - 1)$: -0,033678
statistique de test: $\tau_c(1) = -1,92631$
p. critique 0,3197
Coeff. d'autocorrélation du 1er ordre pour e: 0,029

avec constante et tendance temporelle
avec 0 retards de $(1-L)SP500$
modèle: $(1-L)y = b_0 + b_1*t + (a-1)*y(-1) + e$
valeur estimée de $(a - 1)$: -0,045925
statistique de test: $\tau_{ct}(1) = -2,16055$
p. critique 0,5083
Coeff. d'autocorrélation du 1er ordre pour e: 0,036

Annexe 7 : Corrélogramme de le série S&P 500 différencié à l'ordre 1



Annexe 8 : Graphique de la série S&P 500 différencié à l'ordre 2



Annexe 9 : Tableau de la sélection de retard VAR pour le SPGCSI et le S&P 500

Les astérisques indiquent les meilleures (donc les plus faibles) valeurs des critères d'information suivants, AIC = critère d'Akaike, BIC = critère bayésien de Schwartz et HQC = critère d'Hannan-Quinn.

retards	log-vrais.	p(LR)	AIC	BIC	HQC
1	-957,22313		10,789024*	10,842650*	10,810771*
2	-957,22045	0,94159	10,800230	10,871730	10,829225
3	-957,00750	0,51401	10,809073	10,898449	10,845317
4	-956,67925	0,41780	10,816621	10,923872	10,860114
5	-956,31587	0,39393	10,823774	10,948900	10,874516

Annexe 10 : Tableau de la sélection de retard VAR pour le WTI et le S&P 500

Système VAR, ordre des retards max 10					
Les astérisques indiquent les meilleures (donc les plus faibles) valeurs des critères d'information suivants, AIC = critère d'Akaike, BIC = critère bayésien de Schwartz et HQC = critère d'Hannan-Quinn.					
retards	log-vrais.	p(LR)	AIC	BIC	HQC
1	-321,99241		3,651600	3,705226*	3,673347*
2	-320,91626	0,14236	3,650744*	3,722245	3,679740
3	-320,05978	0,19060	3,652357	3,741733	3,688601
4	-320,00100	0,73353	3,660037	3,770100	3,706400

Table des matières

Introduction	3
Analyse Economique du Dossier	4
Analyse des 3 Variables.....	7
Le Standard & Poors and Goldman Sachs Commodity Index (SPGSCI).....	8
Futures Pétrole Brut West Texas Intermediate (WTI)	9
Indice S&P 500	9
Présentation du Modèle ARIMA.....	10
Stationnarité.....	11
Corrélogramme	13
Processus Moyenne Mobile, $MA(q)$	14
Processus AutoRegressif $AR(p)$	14
Construction du modèle Box-Jenkins	14
Modèle Multiplicatif	15
Choix du modèle ARIMA pour le SPGSCI	16
Stationnarité.....	16
Identification, Estimation, Vérification	18
$ARIMA(0,1,0)(1,0,0)S=5$	19
$ARIMA(1,2,0)(0,0,0)S=5$	21
$ARIMA(0,2,1)(1,0,0)S=5$	23
Choix du Modèle	24
Prévision à une période.....	25
Analyse des Variables WTI et S&P 500	25
WTI	26
S&P 500.....	29
Test de cointégration selon Engle-Granger	33
Couple SPGSCI et WTI.....	33
Relation de corrélation à long terme.....	33
$Y_t = \alpha + \beta X_t + \epsilon_t$	33
Test de Dickey-Fuller Augmenté sur les résidus	34
Couple SPGSCI/SP500.....	35
Couple WTI/S&P 500.....	35
Conclusion	36
Bibliographie	37
Cours	37

Articles.....	37
Sites Internet.....	37
Annexes.....	38