

Rapport - Séries Temporelles

Philippe Real

12/01/2020

Contents

1	Partie I - Exemple de modélisation appliqué au trafic voyageur	2
1.1	Introduction	2
2	Partie II - Tentative de modélisation d'un indice boursier de type action à l'aide de processus ARIMA	2
2.1	Introduction	2
2.2	Lecture des données et premières analyses	2
2.2.1	Traitement des données	3
2.2.2	conversion des données en objet <i>time series</i>	3
2.3	Analyse des séries temporelles obtenues	3
2.3.1	Graphique des séries temporelles - valeur observée Prix à la fermeture (Close)	4
2.3.2	Représentations graphiques : month-plot et lag-plot	5
2.3.3	Etude de la stationarité	6
2.3.4	Decomposition des séries temporelles	8
2.4	Détermination des modèles ARIMA	8
2.4.1	Processus ARMA(p,q)	8
2.4.2	Processus AR(p)	9
2.4.3	Processus MA(q)	9
2.4.4	Differentiation des séries	9
2.4.5	Détermination des paramètres p et q, études des corrélogrammes et autocorrélations partiels	11
2.4.6	Méthode automatique de calibration d'un modèles ARIMA	12
2.5	Validation des modèles obtenus	13
2.5.1	Blancheur des résidus	13
2.5.2	Graphiques de résidus obtenus à partir des différents modèles	14
2.5.3	Normalité des résidus	14
2.5.4	Prévisions à partir des modèles obtenus	15
2.6	Alternative au modèle de type ARIMA, les modèles GARCH	15
2.6.1	Graphique des résidus	19
2.6.2	Prévision obtenus à partir du modèle GARCH(1,1)	19
2.7	Références	19

1 Partie I - Exemple de modélisation appliqué au trafic voyageur

1.1 Introduction

2 Partie II - Tentative de modélisation d'un indice boursier de type action à l'aide de processus ARIMA

2.1 Introduction

On cherche dans cette partie à modéliser par des processus de type ARIMA ou assimilé (SARIMA) l'évolution du prix d'indice boursier de type action. En fait un portefeuille d'actions, ou indice type CAC40, DAX, Eurostoxx, SP500... Pour cette première étude on va se baser sur l'indice CAC40. A partir de la modélisation (présupposée possible) obtenue on va chercher à prévoir l'évolution de l'indice en question, à horizon 3 mois, 6 mois voire 1 an.

Notre but est double :

- La définition de stress test de type action : A partir de la modélisation obtenue, et de l'intervalle de confiance sous jacent, on va chercher à déterminer une valeur de choc absolue à la hausse et à la baisse. Cette méthodologie de définition d'un choc absolue associé à un niveau de confiance devrait nous aider à définir un scénario économique plausible (avec un certain seuil de confiance) à horizon 3mois, 6mois et 1an. Ainsi cette modélisation devrait pouvoir nous guider dans la détermination de stress test de type action. Pour être complet quant à la définition de stress test de type financier, il faudrait parvenir à définir une méthodologie équivalente pour les produits de types taux ou courbes de taux d'intérêt. Ce dernier cas est plus complexe dans la mesure où on cherche à modéliser une surface et les séries temporelles ne sont peut être pas appropriées. Plus précisément on cherche à modéliser un faisceau de courbes aléatoires qui dépendent les unes des autres. Le mécanisme de dépendance étant en parti connu, ou plutôt des modèles existent.
- Elaboration d'un portefeuille simplifié : Une fois les principaux indices modélisés, on va chercher à décomposer nos portefeuilles sur ces indices et ainsi constituer un portefeuille simplifié. Ce portefeuille simplifié serait la base d'un indice benchmark du portefeuille étudié.

Dans un premier temps on va étudier la série temporelle associée à l'évolution du prix de l'indice étudié le CAC40 : représentation graphique, saisonnalité, tendance, stationarité... Pour entrer dans le cadre d'un modèle ARMA(p,q), on va dans un premier temps, étudier la stationarité de notre série. Et la rendre stationnaire le cas échéant. A partir de là on cherchera à déterminer les paramètres p et q du processus auto régressif AR(p) et moyenne mobile MA(q) sous jacent à partir des graphiques ACF et PACF. Enfin on ajustera les coefficients pour obtenir notre modèle. On terminera l'étude en validant le modèle: blancheur des résidus, indépendance, normalité. On pourra alors après validation l'utiliser pour nos prédictions.

2.2 Lecture des données et premières analyses

Les données ont été récupérées sur le site Yahoo Finance. Ticker “^FCHI” pour les données de l'indice CAC40. On considère un jeu de données quotidienne et un autre mensuel. Avec dans les 2 cas un historique de Janvier 1998 à Janvier 2020. A partir de cet historique de 22 ans on va construire différentes séries de profondeur d'historique différente. Après avoir analysé ces séries on essaiera de construire un modèle de type ARIMA pour chacune d'elles.

```
dataCAC40_raw_d <-read.table("Daily_Data_CAC40_1997-2019.csv", sep=";", dec=".",header=T, na.strings =  
dataCAC40_raw_m <-read.table("Monthly_Data_CAC40_1997-2019.csv", sep=";", dec=".",header=T, na.strings
```

2.2.1 Traitement des données

Dans le cas des données journalières, il y a des données manquantes. On va les supprimer. La variable Date est aussi convertit en structure date.

```
##           Date      Open      High      Low
## 1997-11-12: 1  Min.    :2453  Min.    :2518  Min.    :2401
## 1997-11-13: 1  1st Qu.:3695  1st Qu.:3724  1st Qu.:3667
## 1997-11-14: 1  Median :4341  Median :4374  Median :4309
## 1997-11-17: 1  Mean    :4392  Mean    :4424  Mean    :4358
## 1997-11-18: 1  3rd Qu.:5106  3rd Qu.:5135  3rd Qu.:5074
## 1997-11-19: 1  Max.    :6929  Max.    :6945  Max.    :6839
## (Other)    :5660  NA's    :54    NA's    :54    NA's    :54
##           Close      Adj.Close      Volume
## Min.    :2403  Min.    :2403  Min.    :      0
## 1st Qu.:3697  1st Qu.:3697  1st Qu.:      0
## Median :4341  Median :4341  Median : 90349500
## Mean    :4392  Mean    :4392  Mean    : 83842758
## 3rd Qu.:5106  3rd Qu.:5106  3rd Qu.:129391650
## Max.    :6922  Max.    :6922  Max.    :531247600
## NA's    :54    NA's    :54    NA's    :54
```

```
##           Date      Open      High      Low      Close      Adj.Close      Volume
## 1 1997-11-12 2688.8 2701.0 2649.5 2694.5      2694.5      0
## 2 1997-11-13 2691.6 2712.2 2681.8 2700.7      2700.7      0
## 3 1997-11-14 2735.9 2751.4 2691.9 2698.9      2698.9      0
## 4 1997-11-17 2772.1 2779.6 2760.1 2773.0      2773.0      0
## 5 1997-11-18 2787.2 2793.6 2762.6 2782.6      2782.6      0
## 6 1997-11-19 2753.0 2792.3 2753.0 2790.6      2790.6      0
```

Dans le cas des données mensuelles ont n'a pas de problème de données manquantes.

2.2.2 conversion des données en objet *time series*

Ici on convertit les données en objet R *ts* (time series) Dans le cas des données journalières on utilise pour le paramètre de fréquence (nb jours dans l'année) la valeur 256 Ce qui correspond au nombre de jours par an (jours ouvrés sans les jours de fermeture) que l'on obtient une fois les NA supprimés.(On remarque que cette valeur de fréquence influe la vitesse de traitement lors de l'appel de la fonction *auto.arima*)

2.3 Analyse des séries temporelles obtenues

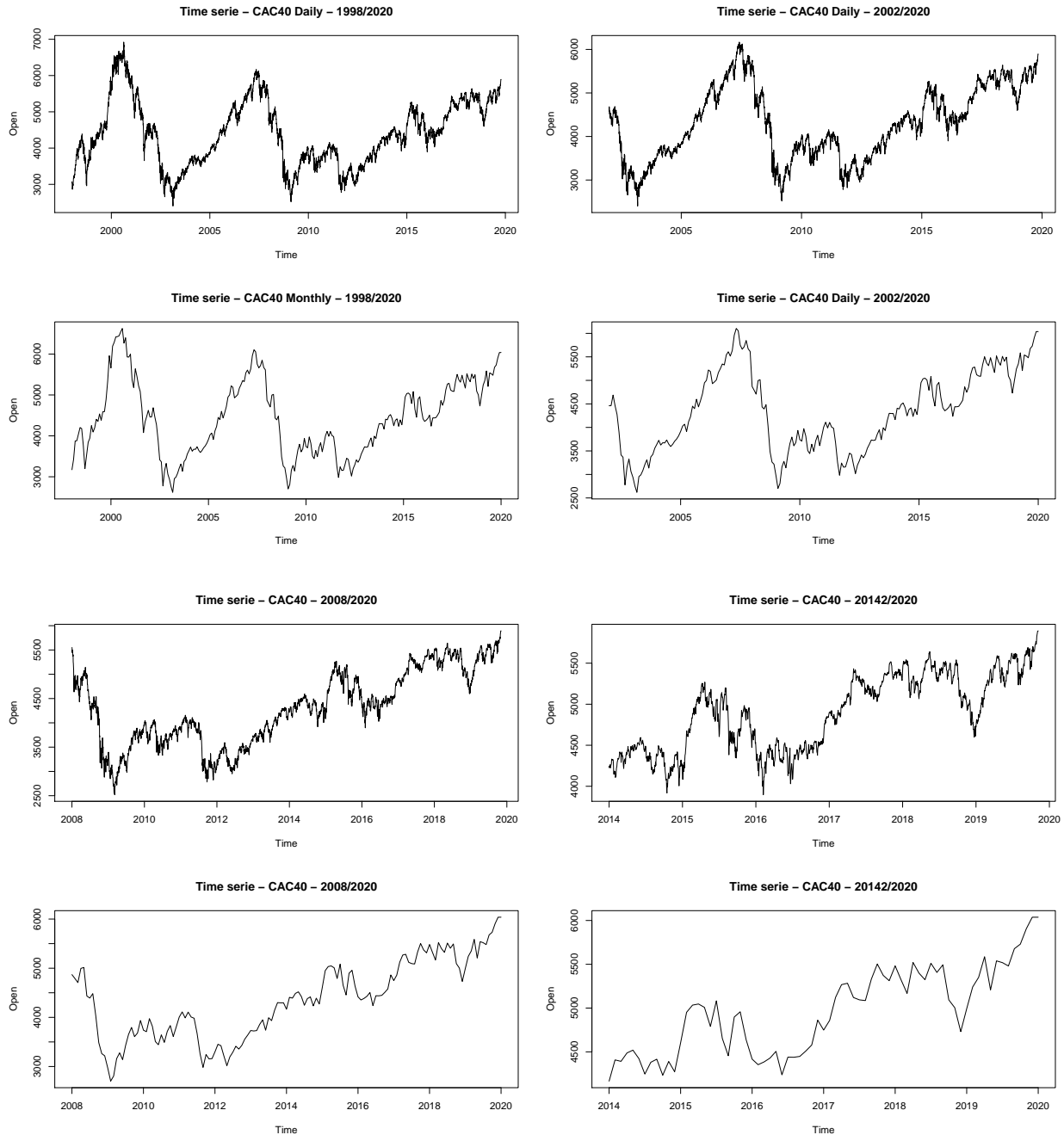
Comme déjà énoncé on va étudier plusieurs profondeur d'historique.

- Toute la série de janvier 1998 à janvier 2020 soit 22 années de profondeur d'historique.
- A partir de Janvier 2002 jusqu'à janvier 2020 soit 18 années de profondeur d'historique.
- A partir de Janvier 2008 jusqu'à janvier 2020 soit 12 années de profondeur d'historique.
- A partir de Janvier 2014 jusqu'à janvier 2020 soit 5 années de profondeur d'historique.

Et on considère 2 jeux de données, avec une fréquence quotidienne et mensuelle.

2.3.1 Graphique des séries temporelles - valeur observée Prix à la fermeture (Close)

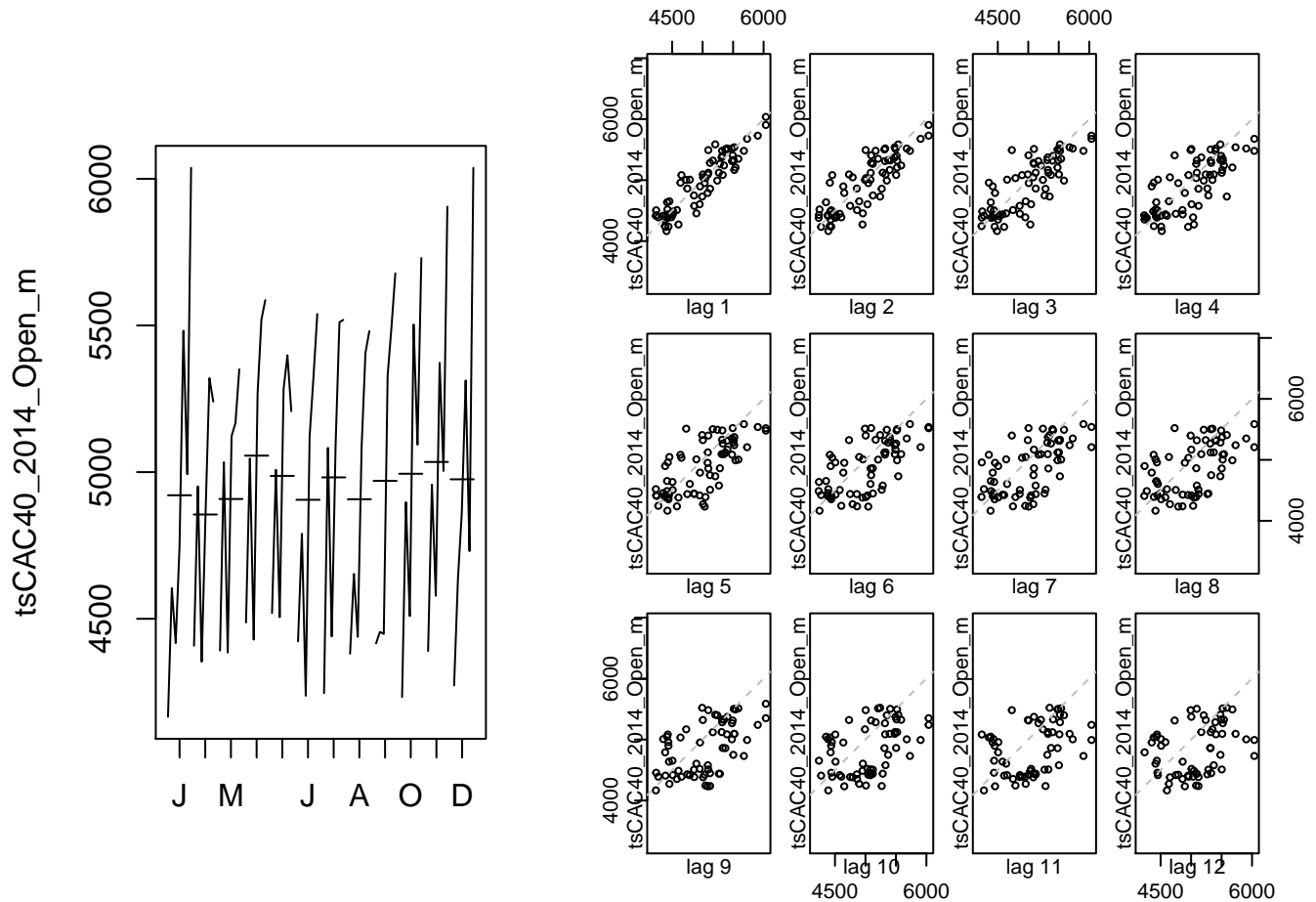
On a 4 séries temporelles possibles en fonction du choix de la quantité observée (High, Low, Open, Close, Volume). On va s'intéresser à la valeur à la fermeture pour la cotation de l'indice CAC40 (Close).



Dans le cas des données mensuelles On retrouve biensûr la forme globale de la série mais moins bruitée.

2.3.2 Représentations graphiques : month-plot et lag-plot

Un lag-plot ou diagramme retarde est le diagramme de dispersion des points ayant pour abscisse la série retardée de k instants et pour ordonnée la série non retardée. Si le diagramme retardée suggère une corrélation entre les deux séries, on dit que la série présente une autocorrélation d'ordre k . Ce diagramme permet de comprendre la dépendance de la série par rapport à son passée. Il donne une vision locale de la série, si y a une corrélation entre la série a un instant et la série 1, 2... instants avant.



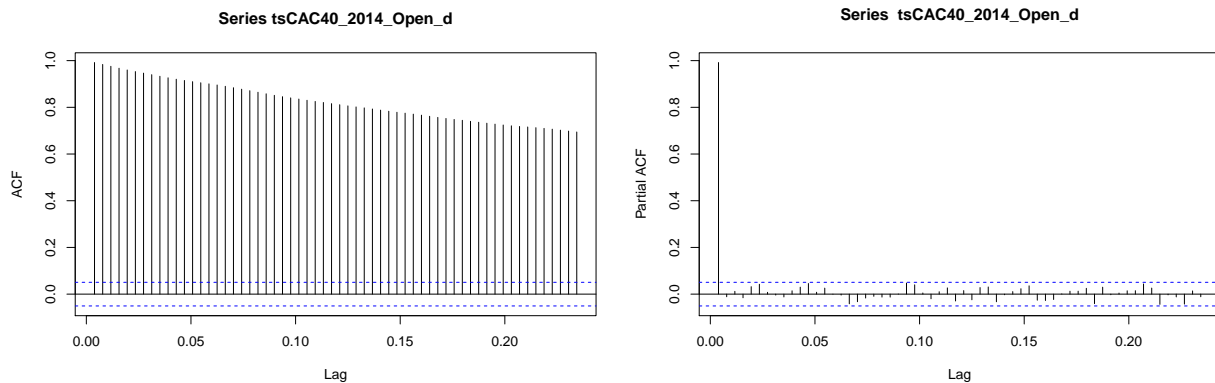
2.3.3 Etude de la stationarité

La stationnarité est la stationnarité du processus au sens faible. Un tel processus doit avoir les propriétés suivantes :

- La moyenne et la variance ne varient pas au cours du temps.
- Le processus n a pas de tendance.

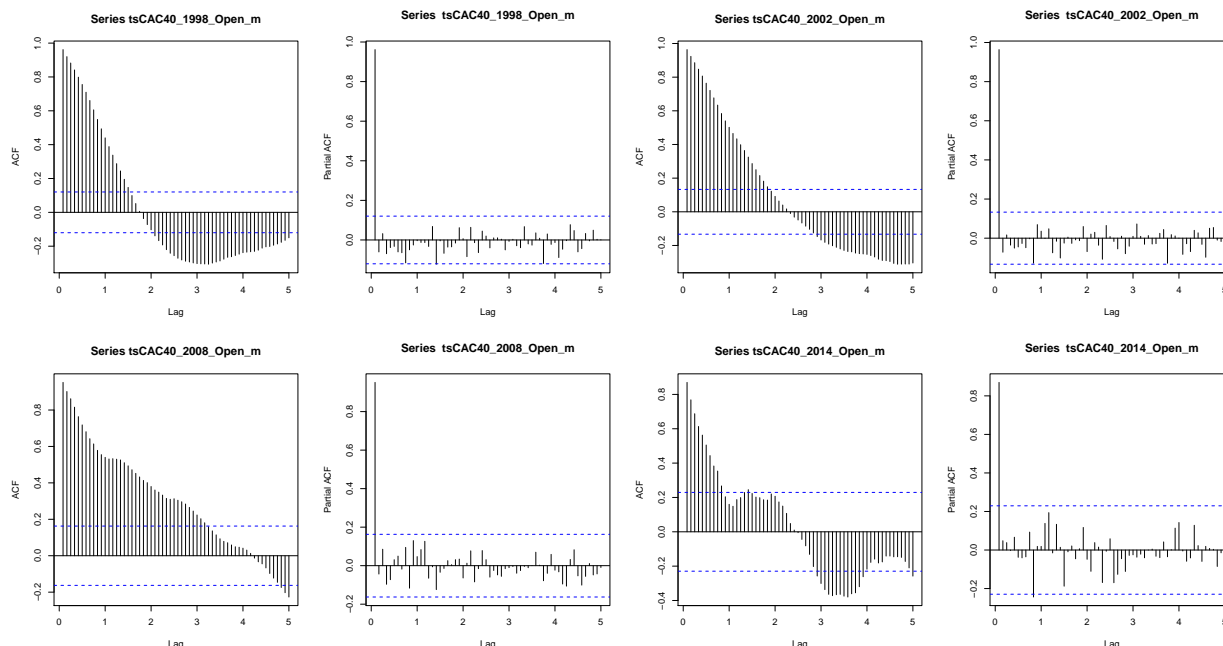
Pour vérifier ces hypothèses, on s'appuiera sur une analyse des graphique d'autocorrélation ACF et d'autocorrélation partielle PACF ainsi que sur le test de *Dickey – Fuller*.

2.3.3.1 Fonction d'autocorrélation ACF et PACF



On constate que les variables sont liées entre elles, i.e. les données ne semblent pas être stationnaires. On confirme cette hypothèse à l'aide d'un test de *Dickey – Fuller*.

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: tsCAC40_2014_Open_d  
## Dickey-Fuller = -2.7425, Lag order = 11, p-value = 0.264  
## alternative hypothesis: stationary
```



Là aussi on constate que les variables sont liées entre elles, i.e. les données ne semblent pas être stationnaires. On confirme cette hypothèse à l'aide d'un test de *Dickey – Fuller*.

```
##
## Augmented Dickey-Fuller Test
##
## data: tsCAC40_1998_Open_m
## Dickey-Fuller = -2.3093, Lag order = 6, p-value = 0.4457
## alternative hypothesis: stationary

##
## Augmented Dickey-Fuller Test
##
## data: tsCAC40_2002_Open_m
## Dickey-Fuller = -2.1857, Lag order = 5, p-value = 0.4981
## alternative hypothesis: stationary

## Warning in adf.test(tsCAC40_2008_Open_m): p-value smaller than printed p-
## value

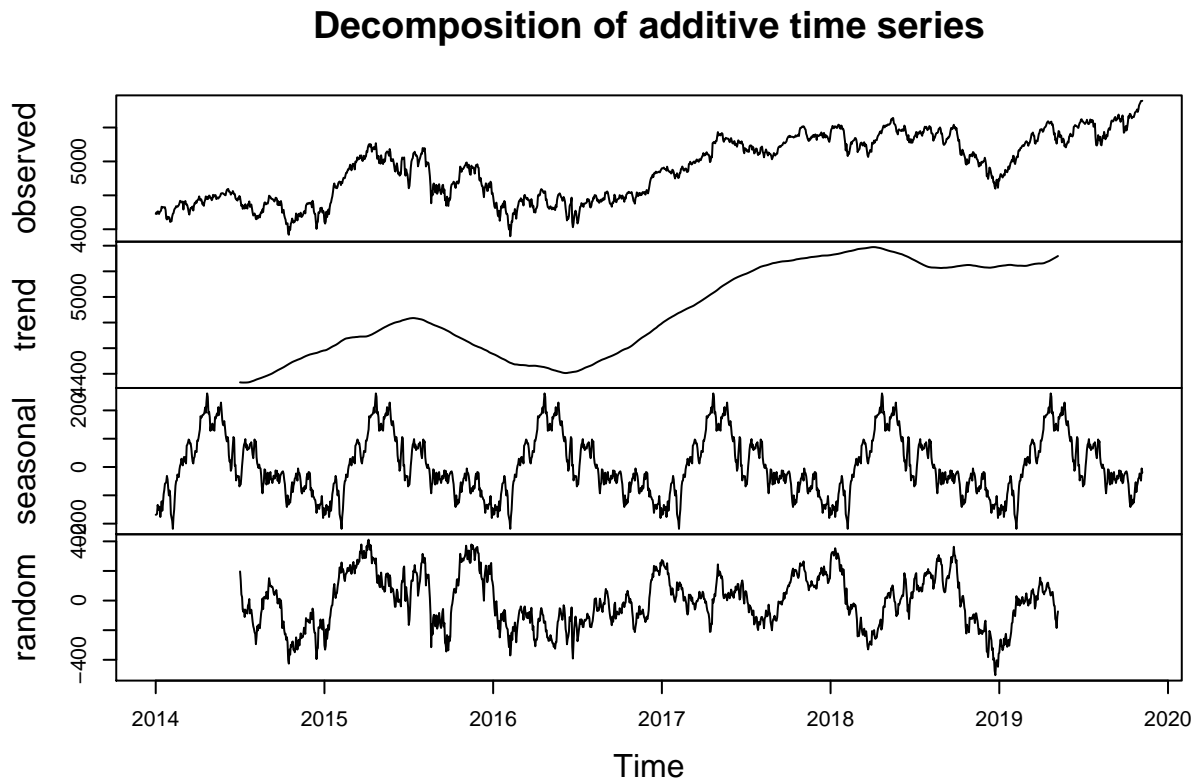
##
## Augmented Dickey-Fuller Test
##
## data: tsCAC40_2008_Open_m
## Dickey-Fuller = -4.2979, Lag order = 5, p-value = 0.01
## alternative hypothesis: stationary

##
## Augmented Dickey-Fuller Test
##
## data: tsCAC40_2014_Open_m
## Dickey-Fuller = -2.1515, Lag order = 4, p-value = 0.5136
## alternative hypothesis: stationary
```

La p-valeur de ce test est de 0.5052 et 0.3092 Ce test confirme donc que les données ne sont pas stationnaires.

2.3.4 Décomposition des séries temporelles

Ici on va décomposer la série temporelles en utilisant la fonction `decompose` de R de façon à avoir une idée générale de la tendance (trend) saisonnalité et bruit.



On retrouve les formes générales mais mon bruitées. La saisonnalité ne semble pas très nette.

On va essayer de rendre stationnaire nos séries. C'est un prérequis pour pouvoir effectuer une modélisation de type ARMA. En utilisant la différentiation on va essayer de se ramener à un processus ARMA. Ainsi on va essayer de modéliser l'évolution du prix de l'indice CAC40 par un processus ARIMA. On commence donc par différencier les séries. Le facteur utilisé est de 1.

2.4 Détermination des modèles ARIMA

Un processus X_t $t \in \mathbb{Z}$ est un processus ARIMA(p,d,q) si $\Delta^d X$ est un processus ARMA(p,q). Si l'on est bien dans le cadre d'un modèle ARIMA, après l'opération de différentiation on se ramènera à l'étude d'un processus ARMA(p,q). Après avoir rappelé la définition d'un processus ARMA(p,q) on cherchera à déterminer ses paramètres p et q.

2.4.1 Processus ARMA(p,q)

Les processus ARMA(p,q) font parti d'une famille très large des processus stationnaires. Ces processus sont composés des processus auto-régressifs AR(p) et de moyennes mobiles ("moving average") MA(q).

2.4.2 Processus AR(p)

Processus auto régressif d'ordre p

$$\forall t = 1, \dots, n \quad X_t = \beta + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \varepsilon_t \text{ (où } \varepsilon_t \text{ est un bruit blanc)}$$

Avec l'opérateur retard cette équation se réécrit :

$$\forall t = 1, \dots, n \quad X_t = \beta + \varepsilon_t + (\alpha_1 B^1 + \alpha_2 B^2 + \dots + \alpha_p B^p) X_t$$

La fonction d'autocorrélation (ACF) d'un AR(p) montre une décroissance exponentielle avec ou sans oscillations vers 0. La fonction d'autocorrélation partielle (PACF) d'un AR(p) est nulle à partir de l'ordre p+1.

2.4.3 Processus MA(q)

Processus auto régressif d'ordre q Ils sont construits à partir de l'idée que l'observation au temps t s'explique linéairement par les observations d'un bruit blanc.

$$\forall t = 1, \dots, n \quad X_t = \beta + \alpha_1 \varepsilon_{t-1} + \alpha_2 \varepsilon_{t-2} + \dots + \alpha_q \varepsilon_{t-q} \text{ (où les } \varepsilon_t \text{ sont des bruits blanc centré)}$$

Avec l'opérateur retard cette équation se réécrit :

$$\forall t = 1, \dots, n \quad X_t = \mu + (\alpha_1 B^1 + \alpha_2 B^2 + \dots + \alpha_q B^q) \varepsilon_t$$

Un processus MA(q) est toujours stationnaire quelles que soient les valeurs des α_i , il est de plus de moyenne μ . L'ACF d'un processus MA(q) est nulle à partir de l'ordre q + 1. si une ACF empirique semble nulle à partir d'un certain ordre q + 1, on peut penser qu'il s'agit de l'ACF d'une série MA(q).

Un processus ARMA est la combinaison des processus autorégressifs et moyennes mobiles. Ainsi, avec les notations précédentes, (X_t) est un processus ARMA(p,q), si :

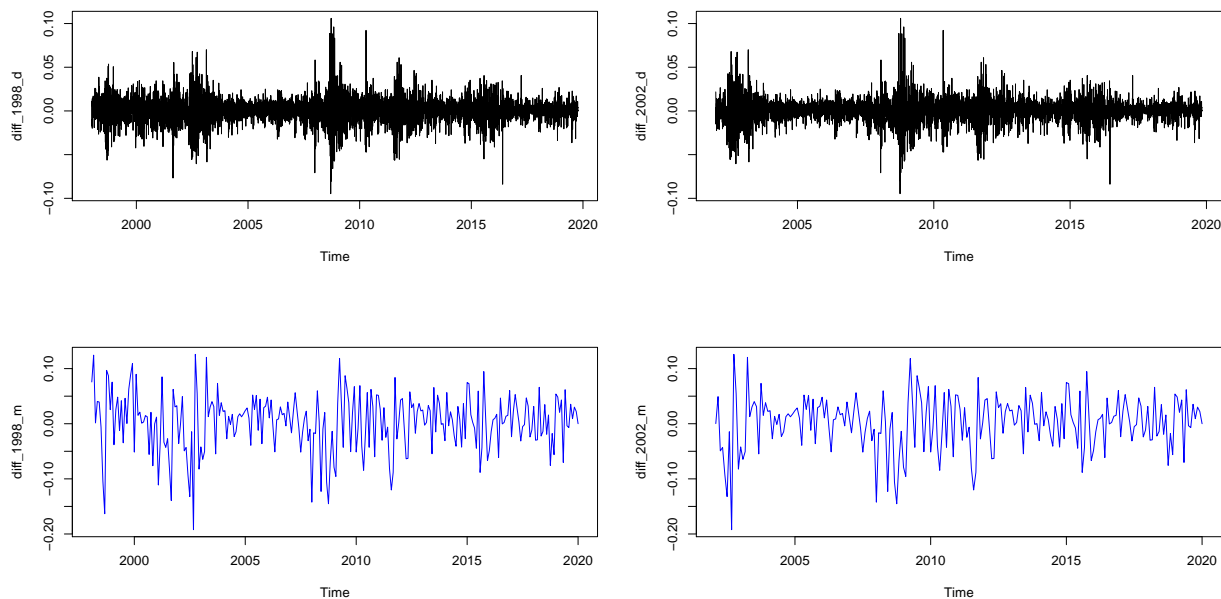
$$\forall t = 1, \dots, n \quad X_t = \beta + \alpha_1 \varepsilon_{t-1} + \alpha_2 \varepsilon_{t-2} + \dots + \alpha_q \varepsilon_{t-q} \text{ (où les } \varepsilon_t \text{ sont des bruits blanc centré)}$$

On cherche maintenant à se ramener à un processus stationnaire en utilisant la différentiation.

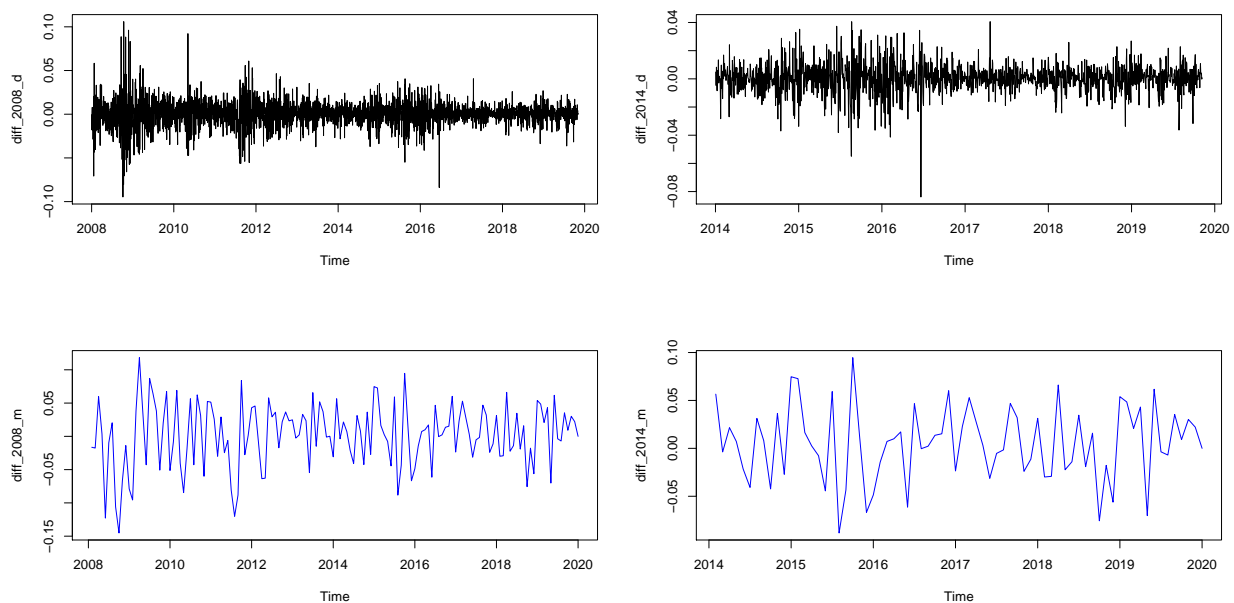
2.4.4 Différentiation des séries

Pour tenter de rendre la série stationnaire, on applique la méthode de différentiation. On utilise la fonction R diff. En paramètre on passe un facteur de 1 pour la différence et de 0 pour la saisonnalité. On retrouve l'idée de la transformation Log-return R_t des prix X_t où $R_t = \text{Log}(X_t/X_{t-1})$ qui est classique en finance.

2.4.4.1 Séries obtenues par différentiation pour les 2 jeux de données quotidien et mensuel à partir de 1998 et 2002



Au sens de la stationnarité faible, les séries semblent bien stationnaire. On retrouve bien une moyenne nulle et cosntante dans le temps. Ainsi qu'aussi une variance constante bien que cet aspect soit moins évident. On affaiblira cetre dernière hypothèse loes de l'étude des modèles GARCH au dernier paragraphe. ##### Séries obtenues par différentiation pour les 2 jeux de données quotidien et mensuel à partir de 2008 et 2014



L'hypothèse de stationnarité du processus ainsi transformé, étant vérifié on peut envisager une modélisation ARIMA. On passe maintenant à la détermination des paramètres p et q du mopdèle ARMA(p,q).

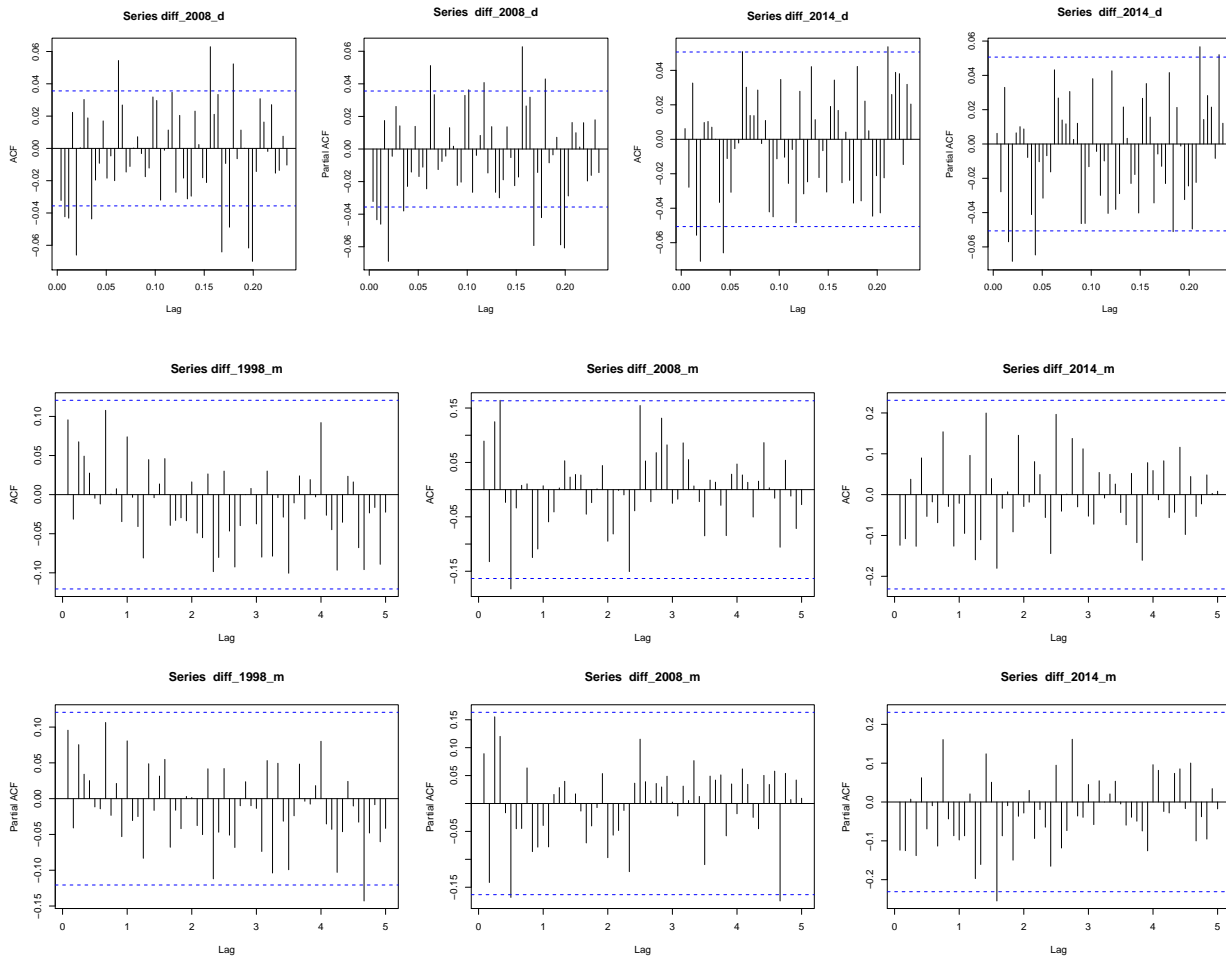
L'estimation de p et de q se fait simplement en lisant le graphe des fonctions d'auto-corrélation et d'auto-corrélation partielle. Le graphe de la fonction d'auto-corrélation nous fournit la valeur de q . Le graphe de la fonction d'autocorrélation partielle nous donne la valeur de p . On peut alors essayer d'améliorer le modèle en prenant des valeurs de p et q plus petites que les valeurs obtenues précédemment, en utilisant notamment

les critères d'AIC ou de BIC.

Les autres paramètres α_i et α_i se font pas minimisation / regression

2.4.5 Détermination des paramètres p et q, études des corrélogrammes et autocorrélations partiels

The next step is to select appropriate ARIMA model, which means finding the most appropriate values of p and q for an ARIMA(p,d,q) model. You usually need to examine the correlogram and partial correlogram of the stationary time series for this. To plot a correlogram and partial correlogram, we can use the `acf()` and `pacf()` functions in R, respectively.



Cependant pour certains processus, ni la fonction d'autocorrélation, ni la fonction d'autocorrélation partielle ne possèdent de point de rupture. Dans de tels cas, il faut construire un modèle mixte.

La PACF d'un processus qui a une composante moyenne mobile a une décroissance exponentielle. Ainsi la PACF d'un ARMA(p,q), $q > 0$ présente une décroissance exponentielle.

=> Ici la PACF ne décroît pas exponentiellement, et rien de très net ne ressort des différents graphiques.

Now, we could compare the sample ACF and PACF to those of various theoretical ARMA models. Use properties of ACF & PACF as a guide to estimate plausible models and select appropriate p, q and d. Alternative to this is discussed next.

2.4.6 Méthode automatique de calibration d'un modèles ARIMA

R provides a function `auto.arima`, which returns best ARIMA model according to either AIC, AICc or BIC value. The function conducts a search over possible model within the order constraints provided.

```
## Series: tsCAC40_1998_Open_m
## ARIMA(1,1,0)(2,0,2)[12]
##
## Coefficients:

## Warning in sqrt(diag(x$var.coef)): production de NaN

##          ar1      sar1    sar2     sma1     sma2
##      0.0725 -0.4722  0.206  0.5295  -0.0942
## s.e.  0.0307      NaN     NaN     NaN    0.0158
##
## sigma^2 estimated as 48253:  log likelihood=-1795.75
## AIC=3603.5   AICc=3603.83   BIC=3624.96
```

Pour les 2 jeux de données à partir de 2008 on obtient un SARIMA(0,1,0)(1,0,0) dans le cas des données quotidiennes et un SARIMA(1,1,0)(2,0,2) pour les données mensuelles

```
## Series: tsCAC40_2008_Open_m
## ARIMA(0,1,0)
##
## sigma^2 estimated as 39166:  log likelihood=-965.77
## AIC=1933.53   AICc=1933.56   BIC=1936.5
```

Pour les 2 jeux de données à partir de 2014 on obtient un ARIMA(0,1,0) soit une marche aléatoire

```
## Series: tsCAC40_2014_Open_d
## ARIMA(0,1,0)
##
## sigma^2 estimated as 2486:  log likelihood=-7976.35
## AIC=15954.7   AICc=15954.7   BIC=15960.01
```

```
## Series: tsCAC40_2014_Open_m
## ARIMA(0,1,0)
##
## sigma^2 estimated as 37198:  log likelihood=-481.03
## AIC=964.06   AICc=964.11   BIC=966.33
```

On peut utiliser une approche empirique et regarder le critère AIC AICc et BIC des différents modèle obtenues on prendra celui qui minimise ces critères. C'est ce qui est fait dans le document en référence D2.

modèle	AIC	AICc	BIC
212	1.5959274×10^4	1.5959274×10^4	1.5959274×10^4
212	1.5959274×10^4	1.5959314×10^4	1.5985827×10^4

2.5 Validation des modèles obtenus

2.5.1 Blancher des résidus

```
##
## Box-Pierce test
##
## data:  arima_2014_d$residuals
## X-squared = 28.717, df = 20, p-value = 0.09348

##
## Box-Pierce test
##
## data:  arima_1998_m$residuals
## X-squared = 12.61, df = 20, p-value = 0.8935

##
## Box-Ljung test
##
## data:  arima_1998_m$residuals
## X-squared = 13.268, df = 20, p-value = 0.8656

##
## Box-Pierce test
##
## data:  arima_2008_m$residuals
## X-squared = 14.892, df = 20, p-value = 0.7826

##
## Box-Ljung test
##
## data:  arima_2008_m$residuals
## X-squared = 15.886, df = 20, p-value = 0.7237

##
## Box-Pierce test
##
## data:  arima_2014_m$residuals
## X-squared = 17.294, df = 20, p-value = 0.6338

##
## Box-Ljung test
##
## data:  arima_2014_m$residuals
## X-squared = 21.41, df = 20, p-value = 0.3734
```

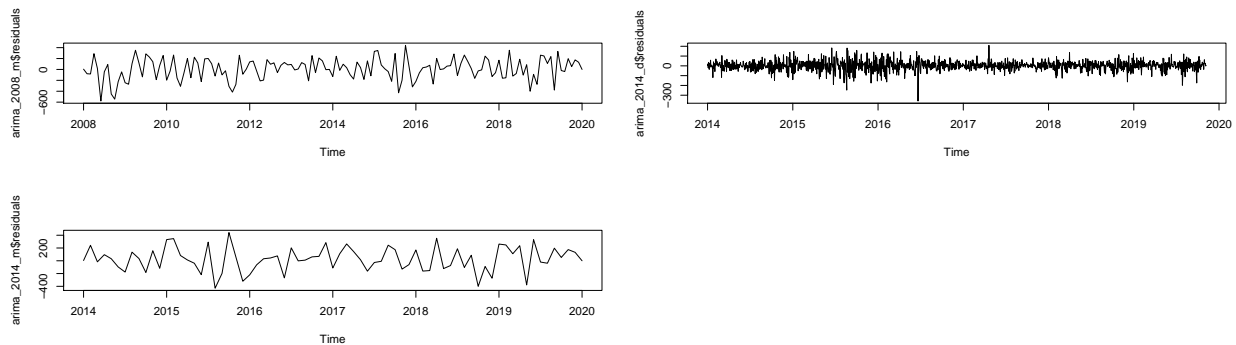
Critère	Type de selection	Variables explicatives	R^2	R^2_{ajus}
AIC	Descendante	x_1, x_3, x_5	0.67	0.54
BIC	Ascendante	x_1, x_5	0.61	0.49

Dans le cas des modèles sur données mensuelles le test de balçeur des résidus rejette le modèle, et ce dans tous les cas. Dans l'autre cas la modélisation est acceptable La profondeur de l'historique et ainsi la forme de la courbe semble aussi jouer un rôle.

```
##
## p.val

##
## [1,]
```

2.5.2 Graphiques de résidus obtenus à partir des différents modèles



Dans le cas des données quotidiennes on a une variance non constante On va utiliser les modèles GARCH comme alternative. On verra qu'ils sont mieux adaptées à la modélisation des séries financières.

2.5.3 Normalité des résidus

```
##
## Shapiro-Wilk normality test
##
## data:  arma_2014_d$residuals
## W = 0.9664, p-value < 2.2e-16

##
## Shapiro-Wilk normality test
##
## data:  arma_1998_m$residuals
## W = 0.98071, p-value = 0.001178

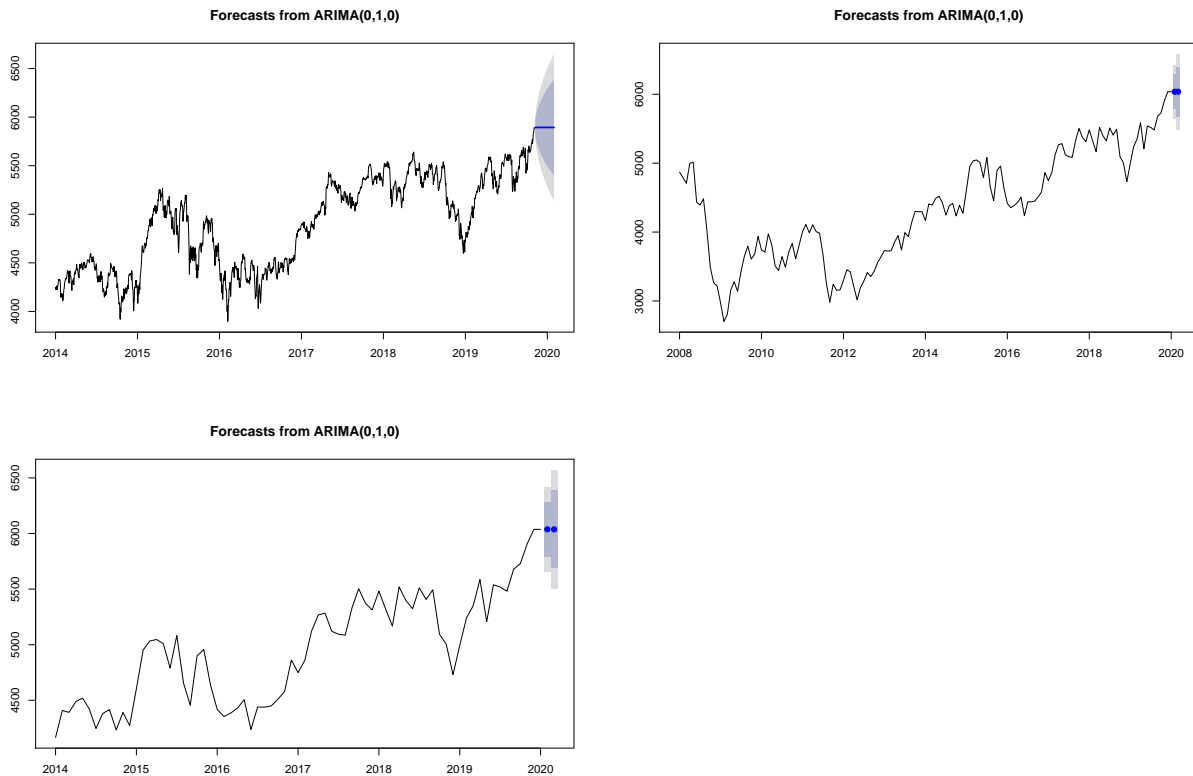
##
## Shapiro-Wilk normality test
##
## data:  arma_2008_m$residuals
## W = 0.9844, p-value = 0.09963

##
## Shapiro-Wilk normality test
##
## data:  arma_2014_m$residuals
## W = 0.9907, p-value = 0.8734
```

Critère	Type de selection	Variables explicatives	R^2	R^2_{ajus}
AIC	Descendante	x_1, x_3, x_5	0.67	0.54
BIC	Ascendante	x_1, x_5	0.61	0.49

2.5.4 Prévisions à partir des modèles obtenus

Ces prévisions seront utilisées pour nous permettre de déterminer les valeurs possibles prise par l'indice CAC40 à horizon 1 mois, 2 mois, 6 mois. On souhaite utiliser l'intervalle de confiance obtenu pour nous aider à déterminer un scénario économique possible sur les actions. Vu que l'on est dans le cadre de stress tests, on cherche à déterminer un choc absolu plausible et non pas obtenir la valeur du CAC à horizon.



Les erreurs de prédictions obtenues semblent suivre une normale centrée et de variance assez constante. Le modèle ARIMA semble être adapté pour la prédiction.

The forecast errors seem to be normally distributed with mean zero and constant variance, the ARIMA model does seem to provide an adequate predictive model. Here we looked at how to best fit ARIMA model to univariate time series. Next thing that I'll work on is Multivariate Time Series Forecasting using neural net.

Cependant, on remarque que la

2.6 Alternative au modèle de type ARIMA, les modèles GARCH

Ces modèles prennent en compte l'hétéroscédasticité. Ils sont mieux adaptés aux séries financières.

Les séries financières comme on a pu le voir, ne sont pas stationnaires. Et on observe une tendance locale B1 p227.

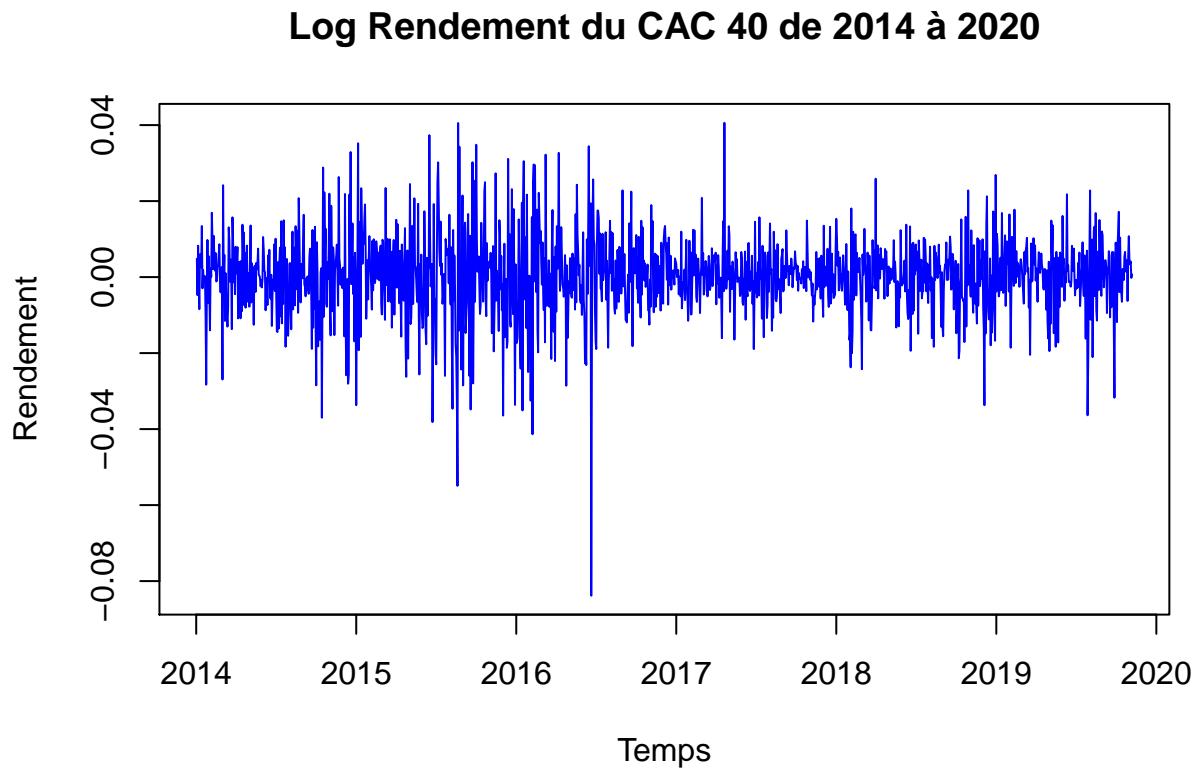
Comme on l'a fait dans le cas de la modélisation ARIMA on transforme la série originelle par différentiation. Pour obtenir une série stationnaire. Ici on va considérer comme c'est souvent le cas en finance le log-return. C'est à dire la quantité déduite du prix X_t de la manière suivante : Log-return R_t des prix X_t où $R_t = \log(X_t/X_{t-1})$. Cette approche est bien adaptée au cadre de la théorie de Black-Scholes. On pourra se reporter au Document D3 page

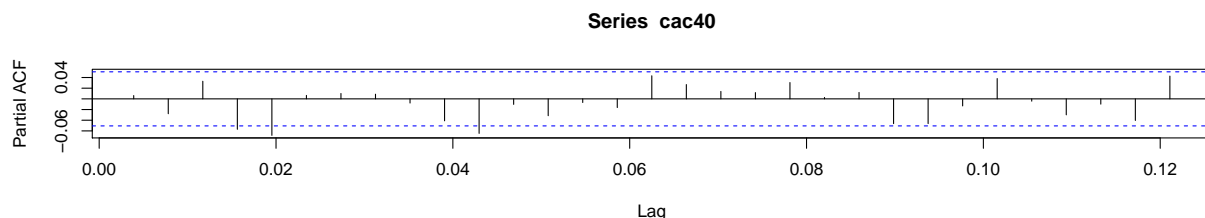
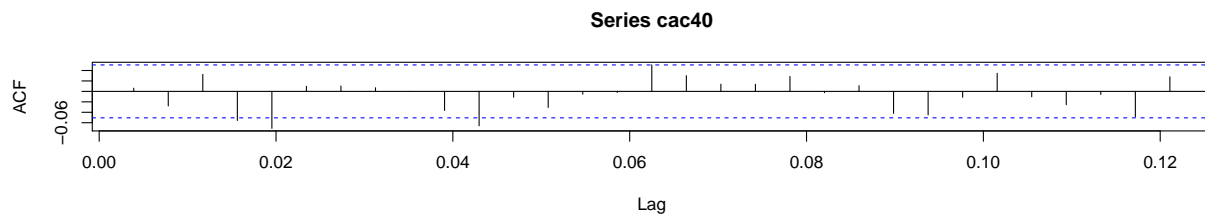
D'après D3 page 96 : Au regard de l'autocorrélogramme partiel du log return, une modélisation à l'aide d'un modèle GARCH(1, 1) semblerait possible. En effet, l'autocorrélogramme et l'autocorrélogramme partiel sont significativement nuls à partir des premiers retards (i.e. $p = q = 1$) Dans notre cas vu que l'indice et la période diffère.

- Obtention de la série des log return

```
cac40 <- diff(log(tsCAC40_2014_Open_d))
```

- Graphique de la série obtenue





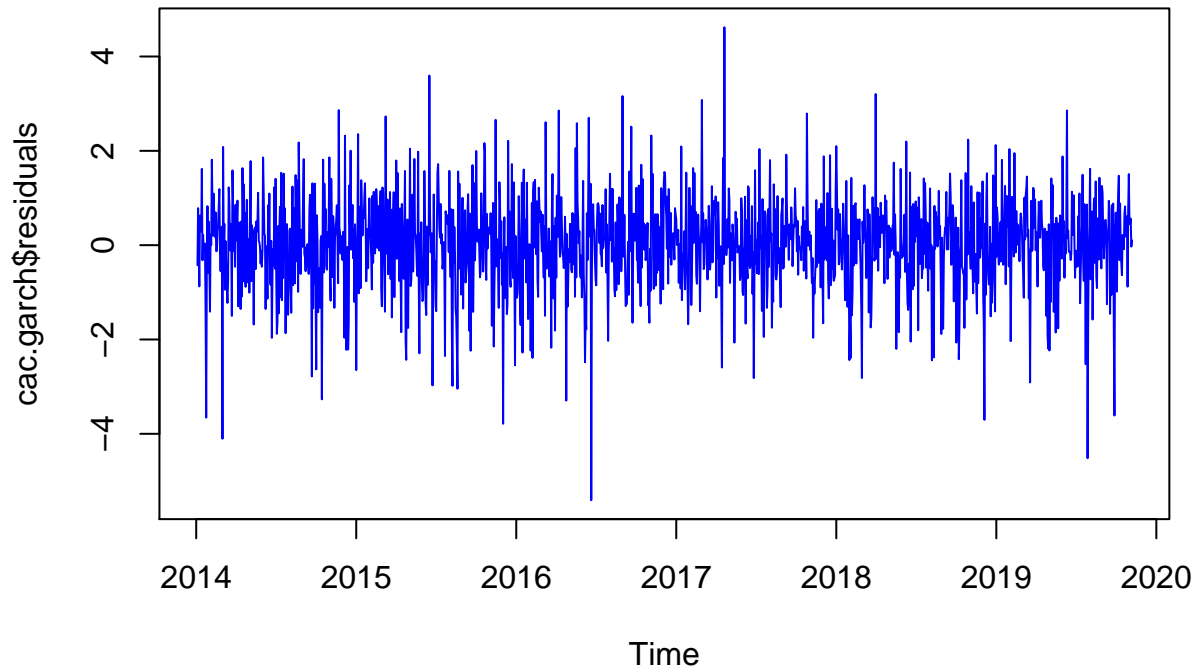
```
##
## ***** ESTIMATION WITH ANALYTICAL GRADIENT *****
##
##
##      I      INITIAL X(I)      D(I)
##
##      1      1.013071e-04      1.000e+00
##      2      5.000000e-02      1.000e+00
##      3      5.000000e-02      1.000e+00
##
##      IT      NF      F      RELDF      PRELDF      RELDX      STPPAR      D*STEP      NPRELDF
##      0      1      -6.076e+03
##      1      8      -6.077e+03      1.41e-04      2.65e-04      4.9e-05      6.7e+10      4.9e-06      8.85e+06
##      2      9      -6.077e+03      4.38e-07      4.64e-07      4.9e-05      2.0e+00      4.9e-06      9.34e+00
##      3      17      -6.095e+03      3.06e-03      4.45e-03      4.5e-01      2.0e+00      8.0e-02      9.30e+00
##      4      20      -6.130e+03      5.64e-03      4.20e-03      7.4e-01      1.9e+00      3.2e-01      5.37e-01
##      5      22      -6.140e+03      1.66e-03      1.50e-03      7.8e-02      2.0e+00      6.4e-02      1.06e+02
##      6      24      -6.162e+03      3.64e-03      3.43e-03      1.3e-01      2.0e+00      1.3e-01      4.63e+03
##      7      26      -6.167e+03      7.37e-04      7.41e-04      2.2e-02      2.0e+00      2.6e-02      3.07e+05
##      8      36      -6.168e+03      1.40e-04      2.66e-04      1.4e-06      5.1e+00      1.6e-06      6.22e+03
##      9      37      -6.168e+03      3.34e-07      3.53e-07      1.4e-06      2.0e+00      1.6e-06      4.28e+03
##      10     47      -6.180e+03      1.98e-03      2.14e-03      5.8e-02      2.0e+00      7.4e-02      4.29e+03
##      11     55      -6.180e+03      1.30e-05      2.93e-05      3.3e-07      9.7e+00      4.4e-07      1.37e-02
##      12     56      -6.180e+03      3.48e-07      3.91e-07      3.3e-07      2.0e+00      4.4e-07      1.14e-02
##      13     67      -6.187e+03      1.12e-03      1.37e-03      4.4e-02      1.2e+00      6.1e-02      1.14e-02
##      14     68      -6.190e+03      4.52e-04      6.35e-04      3.9e-02      6.5e-01      6.1e-02      8.52e-04
##      15     69      -6.193e+03      4.40e-04      4.31e-04      2.3e-02      0.0e+00      4.6e-02      4.31e-04
##      16     70      -6.195e+03      4.59e-04      5.21e-04      2.5e-02      7.9e-01      4.6e-02      7.05e-04
##      17     72      -6.196e+03      3.61e-05      5.13e-05      6.3e-03      1.1e+00      1.3e-02      7.86e-05
##      18     73      -6.196e+03      5.21e-07      2.32e-06      8.6e-04      0.0e+00      1.8e-03      2.32e-06
##      19     74      -6.196e+03      7.61e-07      7.14e-07      5.2e-04      0.0e+00      8.9e-04      7.14e-07
##      20     90      -6.196e+03      -1.10e-14      1.62e-14      1.8e-14      2.1e+04      3.0e-14      2.46e-10
##
## ***** FALSE CONVERGENCE *****
##
##      FUNCTION      -6.195664e+03      RELDX      1.786e-14
##      FUNC. EVALS      90      GRAD. EVALS      20
```

```
## PRELDF      1.617e-14      NPRELDF      2.462e-10
##
##      I      FINAL X(I)      D(I)      G(I)
##
##      1      3.544011e-06      1.000e+00      3.325e+03
##      2      1.307012e-01      1.000e+00      9.388e-02
##      3      8.438273e-01      1.000e+00      2.096e-01
```

```
summary(cac.garch)
```

```
##
## Call:
## garch(x = cac40)
##
## Model:
## GARCH(1,1)
##
## Residuals:
##      Min      1Q  Median      3Q      Max
## -5.40797 -0.52192  0.06096  0.62007  4.61860
##
## Coefficient(s):
##      Estimate Std. Error t value Pr(>|t|)
## a0 3.544e-06   6.782e-07   5.226 1.73e-07 ***
## a1 1.307e-01   1.269e-02  10.302 < 2e-16 ***
## b1 8.438e-01   1.468e-02  57.469 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Diagnostic Tests:
##  Jarque Bera Test
##
## data:  Residuals
## X-squared = 273.67, df = 2, p-value < 2.2e-16
##
##
##  Box-Ljung test
##
## data:  Squared.Residuals
## X-squared = 0.043021, df = 1, p-value = 0.8357
```

2.6.1 Graphique des résidus



Les résidus obtenus sont plus conforme à un bruit blanc.

2.6.2 Prévision obtenus à partir du modèle GARCH(1,1)

2.7 Références

Books :

- (B1) Statistics of financial Markets (J. Franke, W.K. Härdle, C. M. Hafner)
- (B2) Series-Temporelles-avec-R-methodes et cas (Y. Aragon)

Documents:

- (D1) Modèles GARCH et à volatilité stochastique (Christian. Francq)
- (D2) Time Series Analysis with ARIMA – ARCH/GARCH model in R (L-Stern Group - Ly Pham)
- (D3) Séries Temporelles et test d'adéquation d'un modèle GARCH(1,1) (Y. Djabrane)
- (D4) Rapport ISFA - Les Momentums et leur application dans le cadre des marchés boursiers (M. Adil Rahimi)

Blog/Internet:

- (I1) <https://tradingninja.com/2017/03/sp-500-exponential-garch-volatility-model-using-r/>
- (I2) <https://tradingninja.com/2016/01/financial-time-series-modelling-using-arima-plus-garch-models/>