

Analyse des données financières

Abderrahmane FEKHARI

Encadré par:Mr.Khalil SAID

Institut National de Statistique et d'Economie Appliquée

Filière:ROAD

April 2020

Contents

1	Introduction	4
1.1	Présentation de l'indice boursier BIST100	4
2	Analyse statistique	4
2.1	Présentation de la base de données	4
2.2	Statistiques descriptives	4
2.3	Rappel théorique	5
2.4	Implémentation sous R	6
2.4.1	Représentation de la série temporelle	6
2.5	Vérification des propriétés des données financières de la série	6
2.5.1	Stationnarité	6
2.5.2	Auto-corrélation	7
2.6	Vérification des propriétés des données financières du rende- ment géométrique	8
2.6.1	Le rendement géométrique (Log-Rendement)	8
2.6.2	Représentation de la série Log-Rendements de l'indice BIST100	9
2.6.3	Stationnarité	9
2.6.4	Auto-corrélation	10
2.7	Décomposition de la série	11
2.7.1	Tendance	11
2.7.2	Saisonnalité de la série	13
2.8	Prévision	13
2.8.1	Résultats de la prévision avec lissage exponentiel sim- ple ANN	14
2.8.2	Résultats de la prévision avec lissage exponentiel dou- ble MMN	15
2.9	Quelques mesures des performances de la prévision	15
3	Conclusion	16
4	Annexe	17
5	Webographie	20

List of Figures

1	Variation du prix l'indice BIST100	6
2	autocorrelogramme de la série temporelle de l'indice BIST100	8
3	Représentation de la série Log-Rendements de l'indice BIST100	9
4	L'autocorrelogramme du Log-Rendement de la série temporelle de l'indice BIST100	10
5	Technique de lissage "LOESS" avec paramétrisation de 0.1 . .	11
6	Technique de lissage "LOESS" avec paramétrisation de 0.4 . .	12
7	Décomposotion de la série en tendance,saisonnalité et résidu .	13
8	Prévision des prix avec modèle ANN	14
9	Prévision des prix de l'indice avec modèle MNN	15
10	Code source 1	17
11	Code source 2	18
12	Code source 3	19

1 Introduction

1.1 Présentation de l'indice boursier BIST100

La Borsa İstanbul (abrégée en tant que BIST) est la seule entité d'échange de Turquie associant l'ancienne Bourse d'Istanbul (ISE) (Turkish: İstanbul Menkul Kıymetler Borsası, IMKB), l'Istanbul Gold Exchange (Turkish: İstanbul Altın Borsası, İAB) et le Échange de produits dérivés de la Turquie (turc: Vadeli İşlem Opsiyon Borsası, VOB) sous un parapluie. Il a été créé en tant que société incorporée avec un capitale d'environ 240 millions de dollars le 3 avril 2013 et a commencé à fonctionner le 5 avril 2013.

Les actionnaires de Borsa İstanbul sont : 49 % Gouvernement de la Turquie, 41 % IMKB, 5 % VOB, 4 % membres IMKB, 1 % courtiers IMKB et 0,3 % membres İAB. Il est prévu que toutes les actions appartenant au gouvernement seront offertes en vente. Parmi les dirigeants du conseil d'administration de neuf membres, présidé par le président Himmet Karadağ, sont l'ancien vice-président de l'ISE Osman Akyüz, ancien responsable de VOB İşinsu Kestelli, directeur général de la Banque d'investissement Merrill Lynch Hüseyin Keleşoğlu et président de l'Association turque de Instituts intermédiaires du marché des capitaux Attila Köksal.

En juin 2017, c'est la première fois que l'indice BIST 100 de la Bourse d'Istanbul dépasse le cap des 100 000 points, mais au cours de la journée, l'indice est redescendu à 99 135 pour se stabiliser aux alentours de 99 700 points.

2 Analyse statistique

2.1 Présentation de la base de données

Les prix de clôtures de l'indice de 5/7/2007 jusqu'à le 5/8/2019(sauf les weekends : Samedi et Dimanche) avec un total de 3545 lignes, telecharge depuis le site web www.Kaggle.com

2.2 Statistiques descriptives

Min	1st Qu	Median	Mean	3rd Qu	Max
21228	46244	66335	66221	81938	120845

2.3 Rappel théorique

Les deux modèles les plus utilisés sont les suivants :

1. Le modèle additif :

C'est le modèle classique de décomposition dans le traitement des modèles d'ajustement.

La variable X_t s'écrit comme la somme de trois termes :

$$X_t = Z_t + S_t + e_t$$

où

Z_t représente la tendance qui est une composante déterministe variant lentement dans le temps,

S_t représente la saisonnalité (déterministe aussi) qui est une fonction périodique de période donnée,

et e_t représente l'erreur aléatoire du modèle (qui doit être de moyenne nulle et non corrélé (iid)).

2. Le modèle multiplicatif:

La variable X_t s'écrit au terme d'erreur près comme le produit de la tendance et d'une composante de saisonnalité :

$$X_t = Z_t(1 + S_t)(1 + e_t)$$

.

2.4 Implémentation sous R

Utilisation du package de R **TSeries** (Voir code source en annexe)

2.4.1 Représentation de la série temporelle

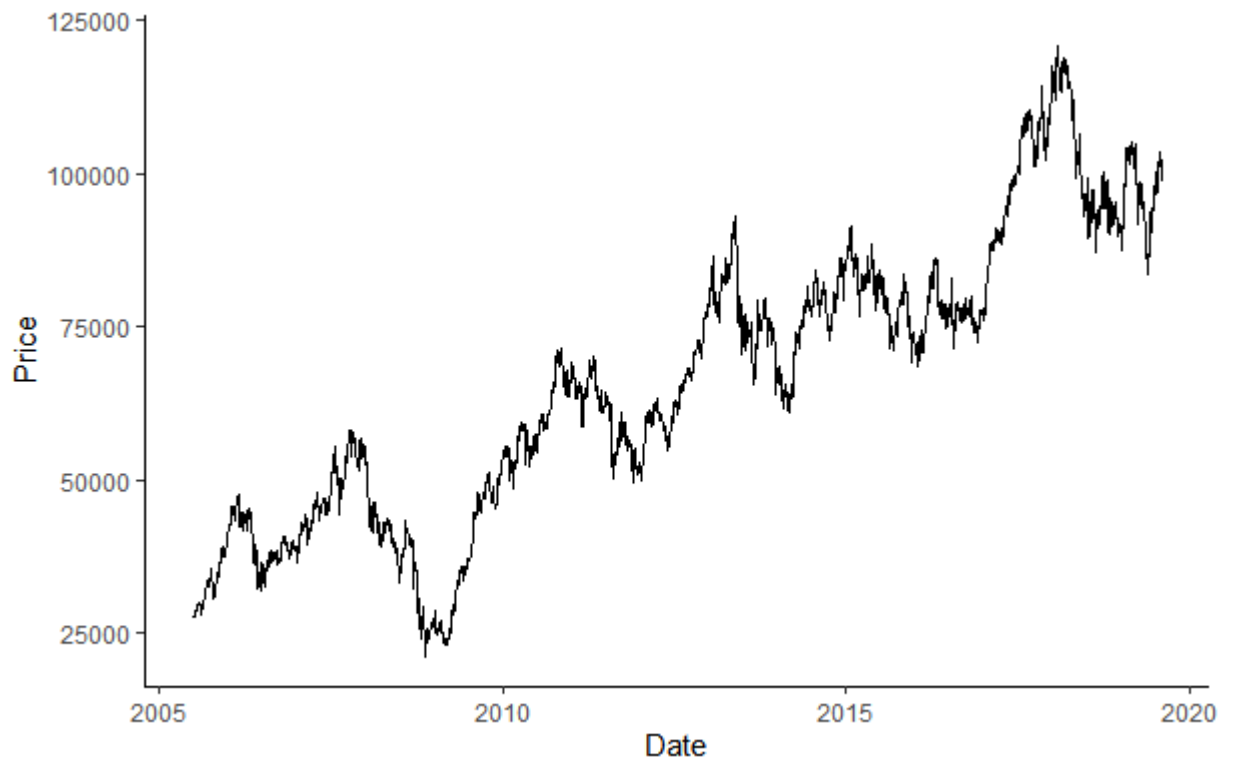


Figure 1: Variation du prix l'indice BIST100

2.5 Vérification des propriétés des données financières de la série

2.5.1 Stationnarité

En vue d'étudier la stationnarité des deux séries, nous allons réaliser le test de Dickey-Fuller Augmenté (ADF). Ce test a été développé en 1981 en se basant sur l'hypothèse de corrélation des résidus et sur l'estimation par la méthode des moindres carrés ordinaires (MCO).

Augmented Dickey-Fuller Test

Dickey-Fuller	-3.3372
Lag order	15
p-value	0.06422

Alternative hypothesis: stationary

D'après le test de Dickey-Fuller augmenté la p-value de valeur 0.06422 est supérieur à 0.05 Donc on accepte l'hypothèse du test de non stationnarité de la série à un niveau de confiance de 5%.

2.5.2 Auto-corrélation

Quand la série est non stationnaire, le coefficient d'autocorrélation d'ordre 1 (empirique) est très élevé(proche de 1), l'autocorrélogramme simple décroît alors lentement.

Définition de la fonction d'autocorrélation ρ :

$$\rho(X_t, X_{t+k}) = \frac{\text{cov}(X_t, X_{t+k})}{\sigma(X_t)\sigma(X_{t+k})} = \rho(X_t, X_{t+k}) = \frac{\gamma(k)}{\gamma(0)}$$

Un estimateur de cette fonction est:

$$\hat{\rho}(X_t, X_{t+k}) = \frac{1}{T} \sum_{j=1}^{T-k} (x_j - \bar{x}_T)(x_{j+k} - \bar{x}_T)$$

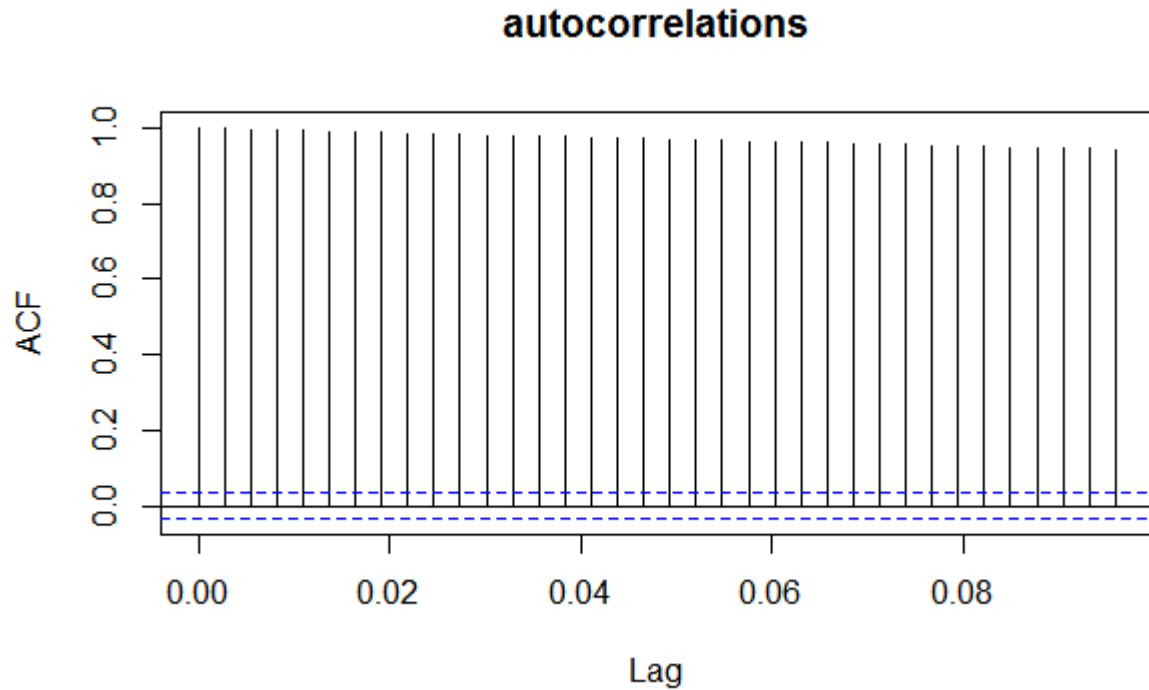


Figure 2: autocorrelogramme de la série temporelle de l'indice BIST100

2.6 Vérification des propriétés des données financières du rendement géométrique

2.6.1 Le rendement géométrique (Log-Rendement)

La rentabilité géométrique:

$$r_t = \ln(1 + R_t) = \ln(P_t + D_t) - \ln(P_{t-1})$$

avec:

P_t : Le prix d'un actif financier à l'instant t,

$P_t = \ln(P_t)$: Le log-prix l'instant t,

D_t : Le flux payé par l'actif à l'instant t,

R_t : La rentabilité simple: $R_t = \frac{P_t + D_t - P_{t-1}}{P_{t-1}}$

2.6.2 Représentation de la série Log-Rendements de l'indice BIST100

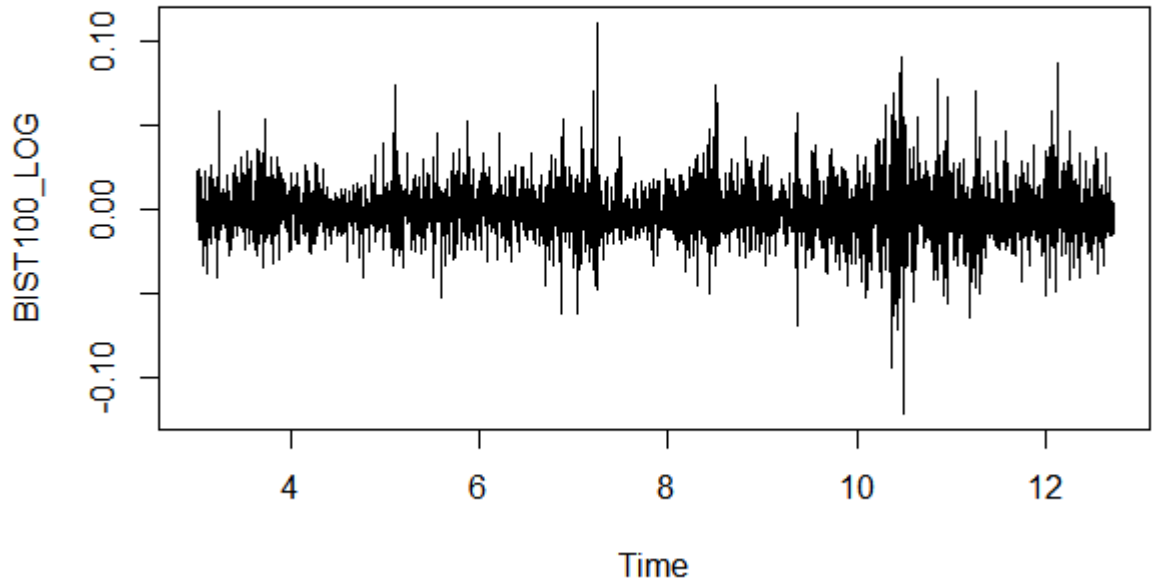


Figure 3: Représentation de la série Log-Rendements de l'indice BIST100

2.6.3 Stationnarité

Les résultats du test Dickey-Fuller Augmenté sont les suivants:

Augmented Dickey-Fuller Test

Dickey-Fuller	-14.075
Lag order	15
p-value	0.01

Alternative hypothesis: stationary

Remarque: p-value smaller than printed p-value.

D'après le test de Dickey-Fuller augmenté la p-value de valeur 0.01 est inférieur à 0.05 Donc on rejette l'hypothèse nulle du test de non stationnarité de la série à un niveau de confiance de 5%(voir meme 1%) et on accepte l'hypothèse de stationnarité.

2.6.4 Auto-corrélation

Quand la série est stationnaire, l'autocorrélogramme associé présente une décroissance rapide:

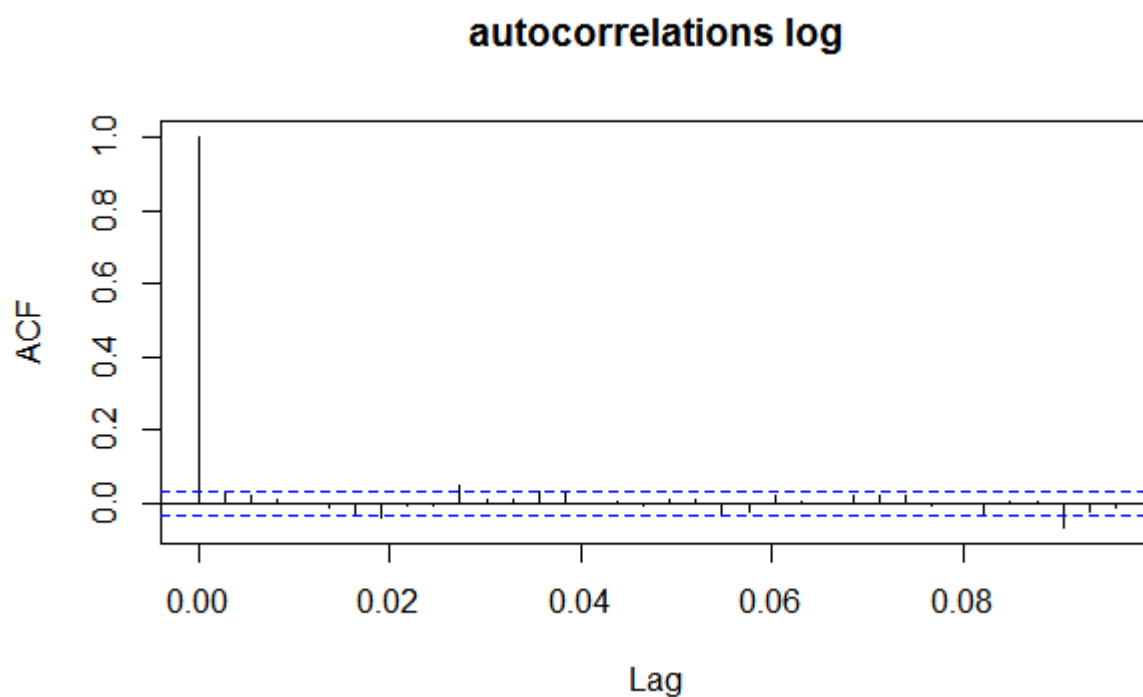


Figure 4: L'autocorrelogramme du Log-Rendement de la série temporelle de l'indice BIST100

2.7 Décomposition de la série

2.7.1 Tendance

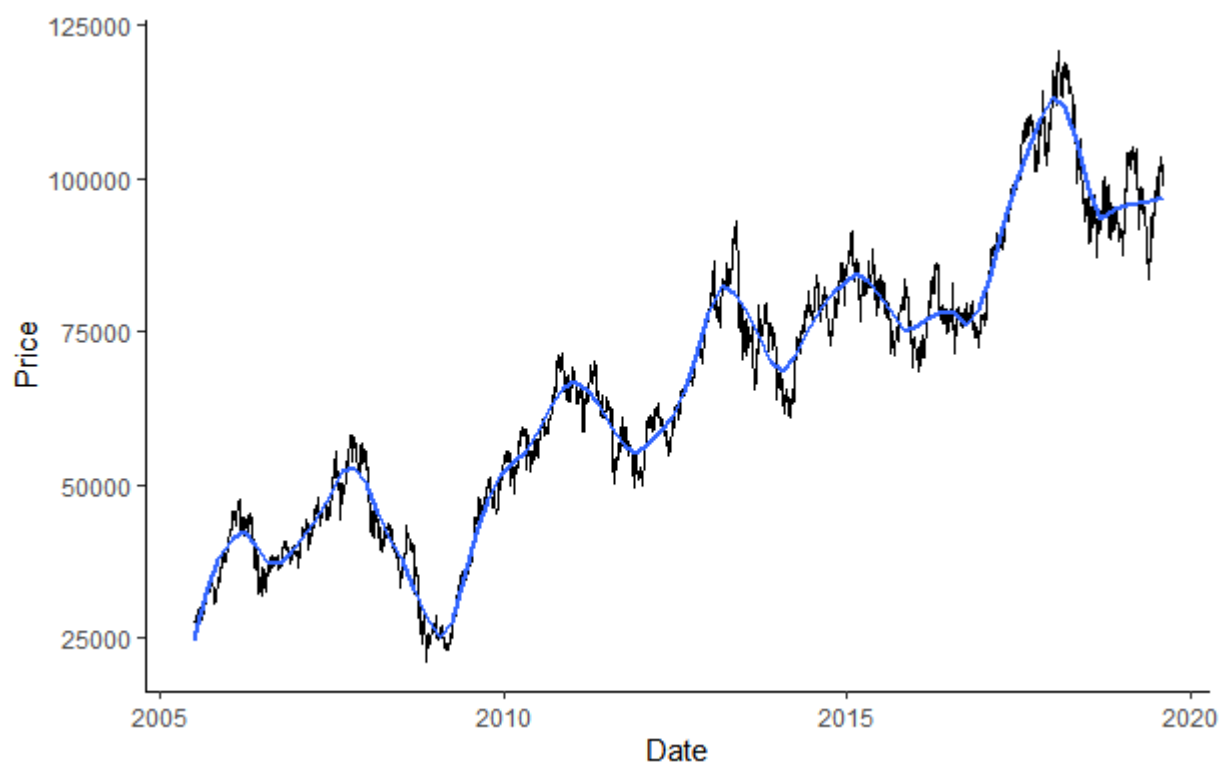


Figure 5: Technique de lissage "LOESS" avec paramétrisation de 0.1

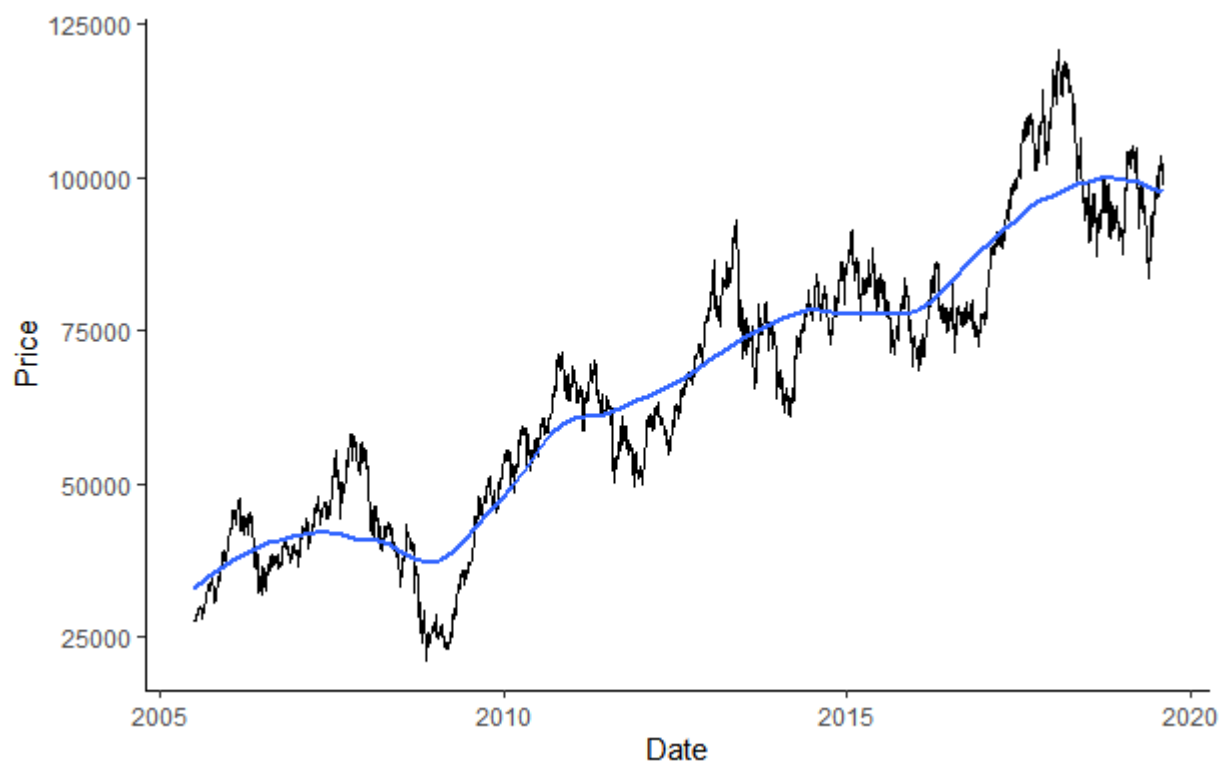


Figure 6: Technique de lissage "LOESS" avec paramétrisation de 0.4

2.7.2 Saisonnalité de la série

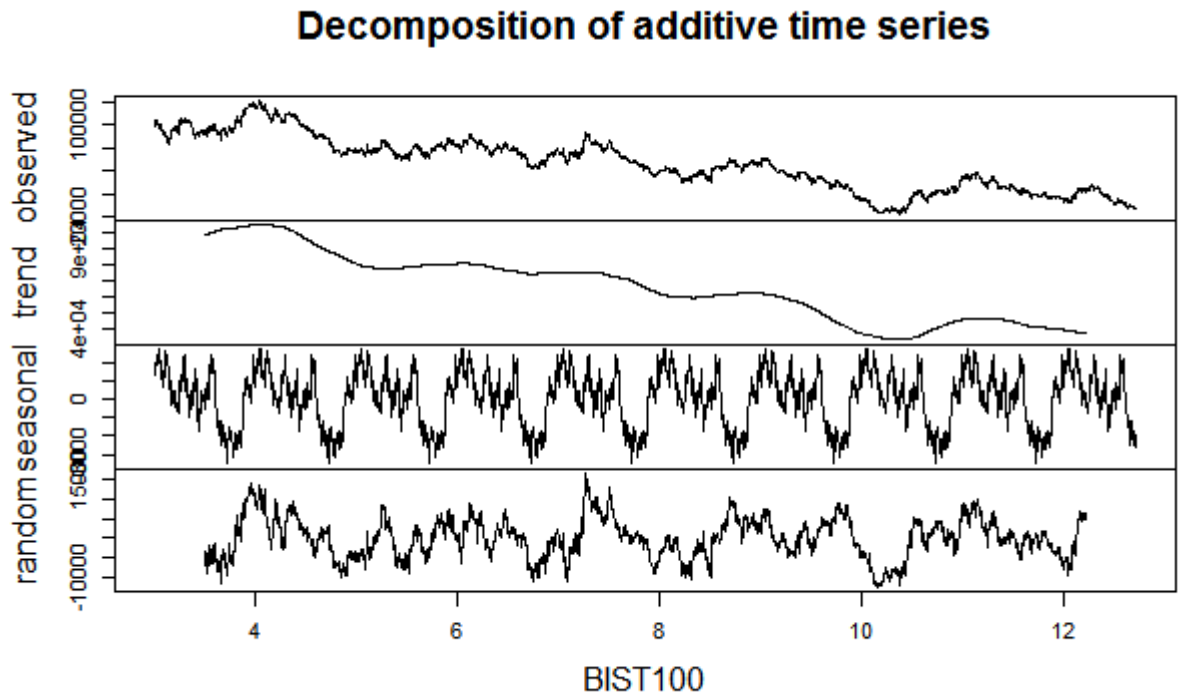


Figure 7: Décomposotion de la série en tendance,saisonnalité et résidu

2.8 Prévision

Utilisation du package **Forecast** du langage R (Voir code source en annexe).

Pour faire une bonne modélisation on divise l'ensemble de données en deux sous-ensembles : Ensemble d'apprentissage : sous-ensemble destiné à l'apprentissage d'un modèle. Ensemble d'évaluation : sous-ensemble destiné à l'évaluation du modèle. On prend par exemples 80% des données pour l'entraînement et les 20% restante pour l'évaluation du modèle.

2.8.1 Résultats de la prévision avec lissage exponentiel simple ANN

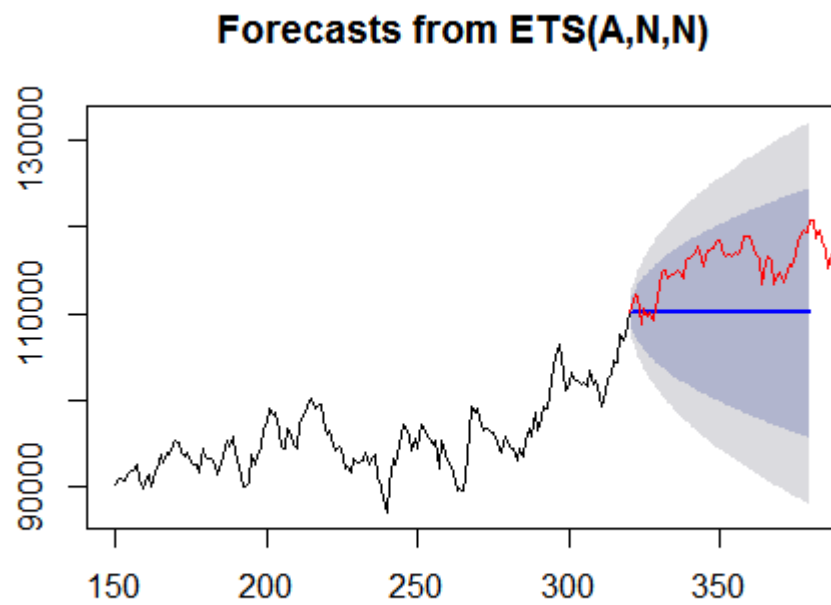


Figure 8: Prévision des prix avec modèle ANN

2.8.2 Résultats de la prévision avec lissage exponentiel double MMN

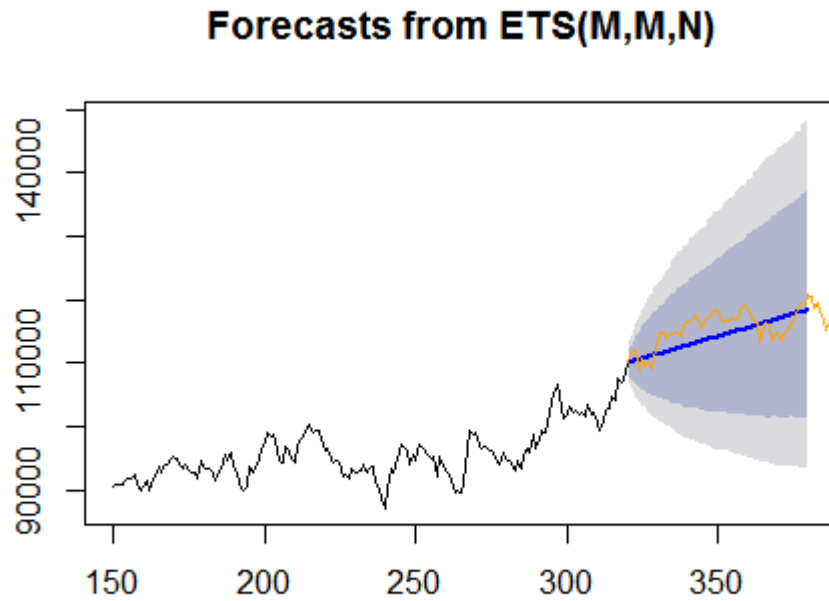


Figure 9: Prévision des prix de l'indice avec modèle MNN

2.9 Quelques mesures des performances de la prévision

$$Erreur = X_t - \hat{X}_t$$

Mean absolute error:

$$MAE = \frac{\sum_{i=1}^n |X_t - \hat{X}_t|}{n}$$

Mean Absolute Percentage Error : MAPE

Mean Absolute Scaled Error : MASE

3 Conclusion

En guise de conclusion, nous avons pu montrer les propriétés des données financières à travers l'exemple de l'indice BIST100. Et on a pu justifier le choix des rendements et non pas les prix, Ces derniers sont caractérisés une non-stationnarité et une forte auto-corrélation qui complique l'étude, contrairement aux rendements qui sont généralement mieux pour les études statistiques.

4 Annexe

Code source avec les commentaires:

```
1 library(forecast)
2 library(dplyr)
3 library(tseries)
4
5 BIST100 <- read.csv("C:\\Users\\Anonymous\\Downloads\\BIST 100 Gemi verileri")
6 head(BIST100)
7 summary(BIST100$Price)
8
9 # representation
10 (time_plot <- ggplot(BIST100, aes(x = Date, y = Price)) +
11   geom_line() +
12   scale_x_date(date_labels = "%Y") +
13   theme_classic())
14
15 barplot
16
17 #decomposition
18 #span sets the number of points used to plot each local regression in the curve
19
20 (time_plot <- ggplot(BIST100, aes(x = Date, y = Price)) +
21   geom_line() +
22   geom_smooth(method = "loess", se = FALSE, span = 0.6) +
23   theme_classic())
24
25 barplot
26
27
```

Figure 10: Code source 1

```

27
28 ## Dickey fuller
29 #alternative hypothesis: stationary
30 adf.test(BIST100_ts)
31
32 #autocorrelogramme
33 acf(BIST100_ts,
34     main = "autocorrelations")
35
36 plot(BIST100_ts)
37 ## plot decomposition
38 plot(decompose(BIST100_ts),xlab="BIST100")
39 #####
40 ##LOG
41 BIST100_LOG=diff(log(BIST100_ts)) #BIST100_LOG
42 plot(BIST100_LOG)
43
44 adf.test(BIST100_LOG)
45
46 acf(BIST100_LOG,
47     main = "autocorrelations log")
48 ## saisonalite de la serie LOG
49 BIST100_LOG_stl <- stl(BIST100_LOG, s.window = "period")
50 # Generate plots
51 plot(BIST100_LOG_stl)
52 ## plot decomposition
53 plot(decompose(BIST100_LOG),xlab="BIST100_LOG")
54

```

Figure 11: Code source 2

```

54
55 ##Modelisation
56 BIST100_LOG_ts <- ts(BIST100_LOG)#, start = c(03,02),frequency=365)
57 BIST100_LOG_ts
58 plot(BIST100_LOG_ts)
59 Train <- window(x = BIST100_ts, start = 150, end =320)
60 Test <- window(x = BIST100_ts, start = 320,end=403)
61
62 ##Moving-Average Smoothing
63 plot(Train)
64 lines(ma(Train,order=3),col="red")
65
66 ##prevision
67 ets1<- ets(Train,model="MMN")
68 fcast<- forecast(ets1,h=60)
69 plot(fcast)
70 lines(Test,col="orange")
71 accuracy(fcast,Test)
72
73
74

```

Figure 12: Code source 3

5 Webographie

<http://aujourdhuiturquie.com/fr/bourse-istanbul-a-battu-record-historique>

<https://www.kaggle.com/umar47/bist100-turkish-stock-market-dataset>

<http://www.wikipedia.org>

<http://www.rdocumentation.org>