



Projet STA202

Étude de la paire GBP-JPY du marché FOREX

Abderrahim Abdellaoui - Hamza Amnay

Mars 2021

Contents

1	Introduction	2
2	Pré-traitement, mise en forme et description de la source	2
3	Analyse descriptive des données	3
4	Modélisation des données et modèle prédictif	6
4.1	Étude de la décomposition en tendances et saisonnalité	6
4.1.1	Choix du modèle	6
4.1.2	Estimation de la tendance par régression linéaire	7
4.1.3	Estimation de la tendance par noyaux gaussiens	7
4.1.4	Estimation de la tendance par polynômes locaux	8
4.1.5	Estimation de la tendance sur bases de splines	9
4.1.6	Estimation de la tendance par moyenne mobile	9
4.1.7	Election du modèle sans tendance (detrended)	10
4.2	Etude de la saisonnalité des données	11
4.3	Choix du modèle	14
5	Simulation de prévision sur un échantillon test	16
6	Conclusion	19

1 Introduction

Dans le cadre du cours de séries chronologiques, nous avons réalisé le projet de séries chronologiques sur la paire de devises GBP/JPY (Pound britannique contre Yen japonais). La paire de devises GBP/JPY est une paire de devises majeures du Forex, le marché des changes flottants. Dans ce marché, les devises convertibles s'échangent l'une contre l'autre à un taux de change qui varie continuellement. La devise de gauche est appelée la devise de base (ici GBP), celle de droite est appelée la devise de contre-partie ou devise de contre-valeur (ici JPY). Et la cotation se lit toujours dans le même ordre (on exprime la devise de base en fonction de la devise de contre-partie). La plus petite différence de cotation possible est le pip. Pour la grande majorité des devises, il s'agit de la quatrième décimale de la cotation. Par exemple, si le cours passe de 1,0001 à 1,0003, alors le cours est monté de 2 pips. Le Yen est une exception en ce qui concerne la valeur du pip: dans ce cas particulier, le pip est la deuxième décimale de la cotation.

2 Pré-traitement, mise en forme et description de la source

Jeu de données et pré-traitement

Dans ce projet, on commence en premier lieu par l'importation des données en utilisant l'API yfinance, malheureusement cette Api n'est pas utilisable sur R, on a donc importé les données sur python puis on les a enregistrés sous format csv pour les exploiter plus tard sur R. On a importé les données journalières du cours du GBP/JPY, pour les dates comprises entre le 17 décembre 2012 et le 13 août 2015. Dans cette période de temps, le GBP a performé contre le JPY.

Une première description fondamentale des données

En effet, d'un point de vue fondamental, la livre sterling avait des raisons d'augmenter et le yen avait des raisons de baisser.

En décembre 2013, une autre donnée solide du Royaume-Uni a poussé le GBP/JPY vers de nouveaux sommets pluriannuels. Avec l'accélération de l'activité des secteurs de la fabrication et de la construction au mois de novembre, les perspectives pour le Royaume-Uni s'amélioraient de jour en jour. La Banque d'Angleterre avait déjà admis que son objectif de taux de chômage pourrait être atteint un an plus tôt que prévu et vu que les rapports PMI du pays continuaient de surprendre à la hausse, non seulement les investisseurs se sont positionnés pour un resserrement encore plus précoce de la Banque d'Angleterre.

Plus la baisse du taux de chômage était rapide, plus la livre sterling devenait forte. La surperformance de l'économie britannique s'est poursuivie au cours de l'année 2014, en grande partie grâce au soutien continu que le gouvernement a apporté au marché du logement.

En revanche, les politiques fiscales et monétaires japonaises ont maintenu le yen dans une tendance baissière. Les perspectives de la paire GBP/JPY dépendent largement de celles de la paire USD/JPY. Il y avait un écart entre les rendements à 10 ans américains et japonais s'élargir et cette tendance devrait largement se poursuivre au cours de l'année 2014. Plus important encore, lorsqu'elle a annoncé son plan de réduction des achats mensuels d'obligations, elle a inclut un calendrier pour ramener les achats d'actifs à zéro.

Mise en forme et description de la source

Les données téléchargées contiennent plusieurs colonnes:

	Date	Open	High	Low	Close	Adj Close	Volume
1:	2012-12-17	136.038	136.095	135.430	136.108	136.108	0
2:	2012-12-18	135.973	136.480	135.815	135.952	135.952	0
3:	2012-12-19	137.042	137.804	136.930	137.050	137.050	0
4:	2012-12-20	136.840	137.463	136.220	136.876	136.876	0
5:	2012-12-21	137.380	137.396	136.059	137.377	137.377	0

690:	2015-08-07	193.496	193.697	191.986	193.500	193.500	0
691:	2015-08-10	192.555	193.620	192.480	192.520	192.520	0
692:	2015-08-11	194.190	194.907	193.960	194.194	194.194	0
693:	2015-08-12	194.900	195.021	193.490	194.860	194.860	0
694:	2015-08-13	194.032	194.714	193.720	193.970	193.970	0

[figure 1] Aperçu de la base de données utilisée.

Vu que ce qui nous intéresse est le prix de clôture, on a fait le choix de ne garder que les colonnes Date et Close (prix de clôture) lors de la construction de la série temporelle. Notre choix s'est porté sur une construction de série temporelle par la bibliothèque XTS, vu que la fréquence d'échantillonnage de nos données n'est pas régulière, en effet on a pas de données pour le samedi et le dimanche.

3 Analyse descriptive des données

On commence par une visualisation de notre série temporelle du GBP/JPY.

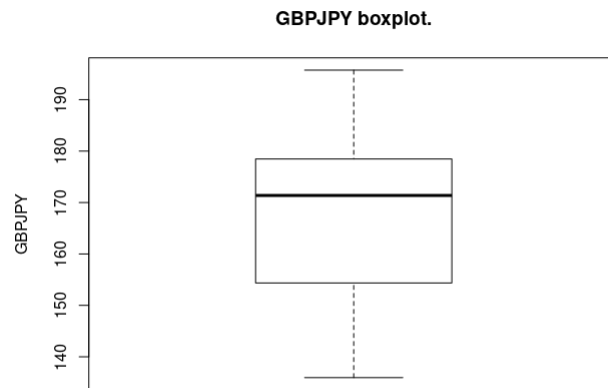


[figure 2] Aperçu de de l'évolution temporelle du GBP/JPY.

On remarque que le cours suit une tendance haussière, par la suite on étudiera cette tendance. Par ailleurs, on avait déjà évoqué les raisons fondamentales derrière cette tendance haussière au début de la section 2. On remarque également des fluctuations autour de la tendance. En particulier, on peut dire que vu que la variable n'est pas stationnaire (la moyenne change avec le temps), elle est stationnaire autour de la tendance, on

explicitera les remarques évoquées dans ce paragraphe quantitativement dans de prochaines sections.

On affiche la boîte à moustache correspondante à notre base de données:



[figure 3] Boîte à moustache du GBP/JPY.

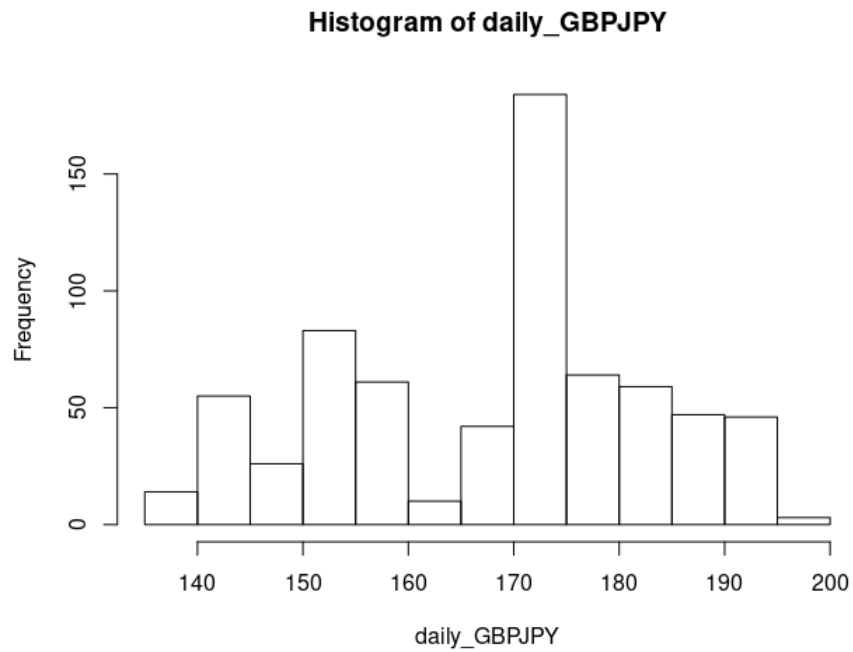
On vérifie qu'on a les bonnes valeurs dans la boîte à moustache à savoir : Dans la boîte à moustaches on a une représentation graphique des quantiles qu'on peut obtenir par la fonction `summary` : la première barre horizontale correspond au maximum, ensuite le 3ème quartile, puis la médiane et le premier quartile puis le minimum.

En exécutant la commande `summary`, on obtient le résultat suivant :

Index	daily_GBPJPY
Min. :2012-12-17 00:00:00	Min. :136.0
1st Qu.:2013-08-15 06:00:00	1st Qu.:154.4
Median :2014-04-15 12:00:00	Median :171.4
Mean :2014-04-15 07:39:09	Mean :168.0
3rd Qu.:2014-12-14 06:00:00	3rd Qu.:178.4
Max. :2015-08-13 00:00:00	Max. :195.7

[figure 4] `summary` de la base de données du GBP/JPY.

Lorsqu'on affiche l'histogramme de notre série temporelle, on retrouve la figure suivante :



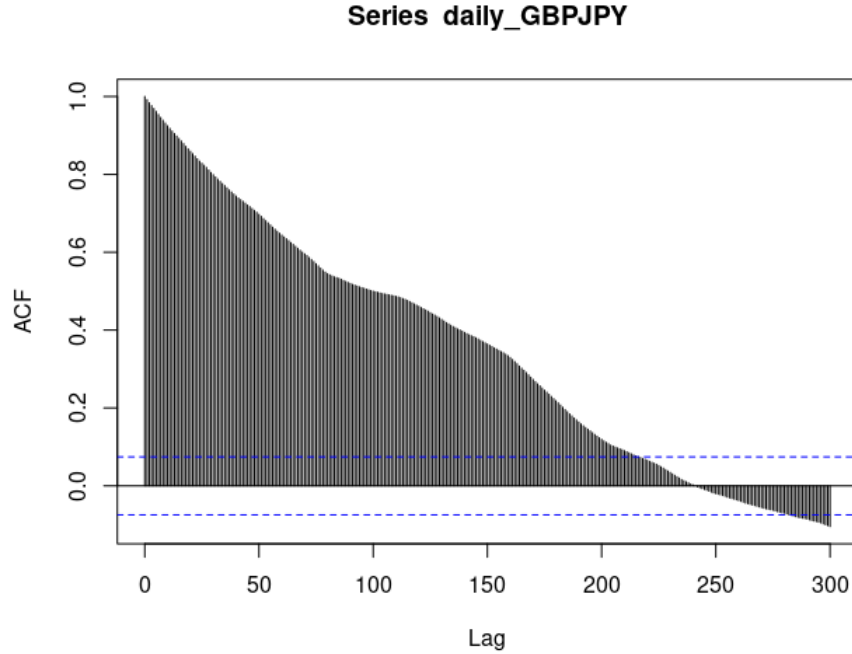
[figure 5] Histogramme des prix de cloture du GBP/JPY.

On remarque que la plus grande fréquence est entre 170 et 175, ceci s'explique par le fait que le marché évoluait en range (il faisait du surplace) entre novembre 2013 et septembre 2014. La figure suivante montre comment les prix sont restés coincés dans un canal pendant 11 mois, ce qui est très significatif sur une période de 30 mois.



[figure 6] Graphique de l'auto-corrélation edité depuis TradingView

Ci-dessous un corrélogramme de notre série temporelle :



[figure 7] Graphique de l'auto-corrélation.

On remarque au niveau du corrélogramme qu'on a une décroissance assez lente au niveau du corrélogramme ce qui témoigne de l'effet mémoire de notre série temporelle. La série présente une tendance additive linéaire par morceau, donc ça nous confirme que la série ne représente pas un mouvement stationnaire.

4 Modélisation des données et modèle prédictif

4.1 Étude de la décomposition en tendances et saisonnalité

4.1.1 Choix du modèle

Une première analyse des données nous a permis d'élire un modèle candidat. On peut observer qu'on a une tendance linéaire, avec des fluctuations autour de cette tendance c'est le bruit. On se doute qu'il y ait de la saisonnalité, notamment pour des raisons fondamentales : le marché du Forex de par sa nature n'est pas saisonnier sinon il serait prédictible et donc facile à trader ce qui n'est pas le cas, le marché du Forex est l'un des marchés les plus efficaces et difficiles à trader.

Lorsqu'on isolera (detrend) le signal de sa tendance, on s'attend à ce que le résidu soit stationnaire, il ne dépend pas du temps. On vérifiera ces suppositions par la suite.

Nous modélisons notre série temporelle, qu'on note ici Y_t comme fonction de :

1. La tendance T_t qu'on a supposé linéaire (on vérifiera ceci dans la section prochaine), la tendance représente l'évolution de la série à long terme.
2. La saisonnalité S_t représente un phénomène périodique qui se reproduit au long de la série temporelle.
3. Le bruit ϵ_t qui représente la partie aléatoire.

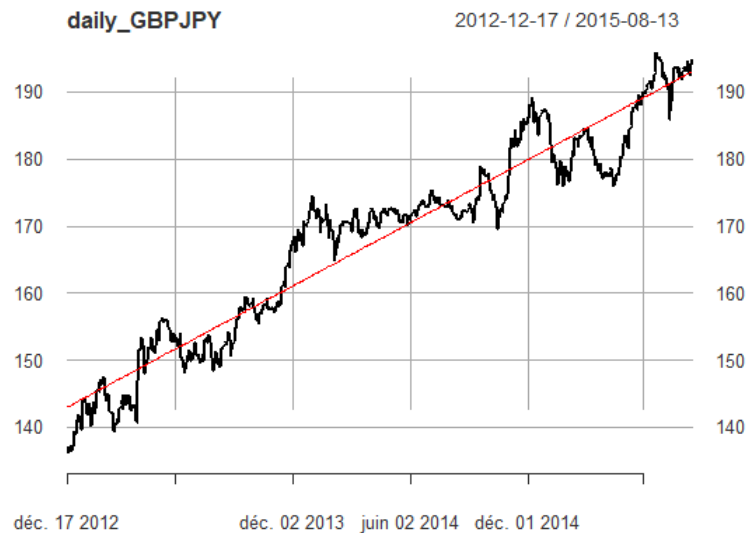
Cette décomposition peut-être additive $Y_t = T_t + S_t + \epsilon_t$ ou multiplicative $Y_t = T_t * S_t * \epsilon_t$. Il est également possible de combiner ces deux décompositions: $Y_t = (T_t + S_t) * \epsilon_t$ ou $Y_t = (T_t * S_t) + \epsilon_t \dots$

4.1.2 Estimation de la tendance par régression linéaire

On commence par estimer la tendance par régression linéaire :

```
reg<-lm(daily_GBPJPY~df_daily_GBPJPY$Date)
summary(reg)
ychap.lm <- reg$fitted
plot(daily_GBPJPY)
lines(ychap.lm,col='red',lwd='1')
summary(daily_GBPJPY - reg$fitted)
```

[figure 8] Code de l'estimation de la tendance par régression linéaire.

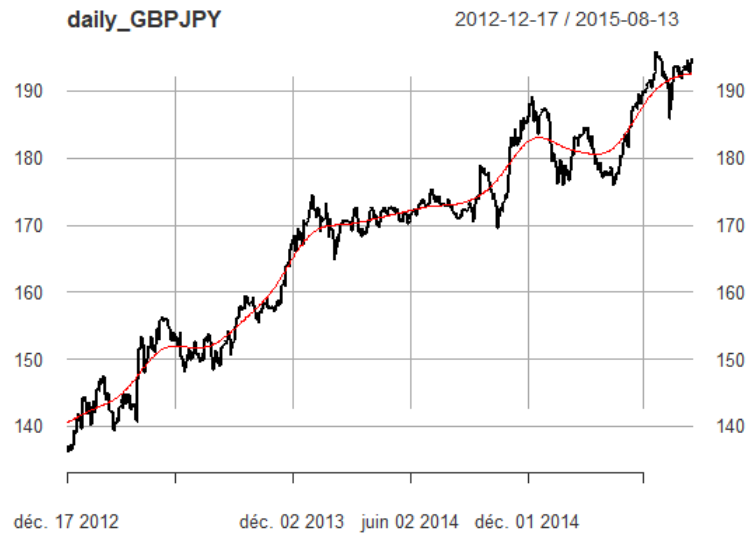


[figure 9] Estimation de la tendance par régression linéaire.

On remarque qu'on a donc une tendance linéaire, ce qui confirme notre supposition émise précédemment .

4.1.3 Estimation de la tendance par noyaux gaussiens

Pour confirmer notre résultat sur l'estimation de la tendance, on estime cette fois la tendance par noyaux gaussiens.



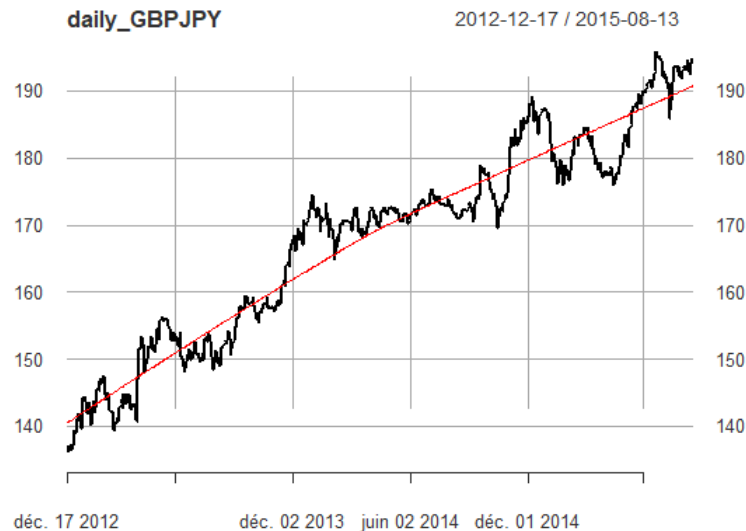
[figure 10] Estimation de la tendance par noyaux gaussiens.

La tendance cette fois n'est plus une droite oblique, mais on garde l'aspect haussier du trend.

4.1.4 Estimation de la tendance par polynômes locaux

```
daily_lo<-loess(daily_GBPJPY~t_daily, degree=1, span=0.9)
ychap.daily_lo<-xts(daily_lo$fitted,order.by=daily_Date)
plot(daily_GBPJPY,type='l')
lines(ychap.daily_lo,col='red')
```

[figure 11] Histogramme des prix de cloture du GBP/JPY.



[figure 12] Estimation de la tendance par polynômes locaux

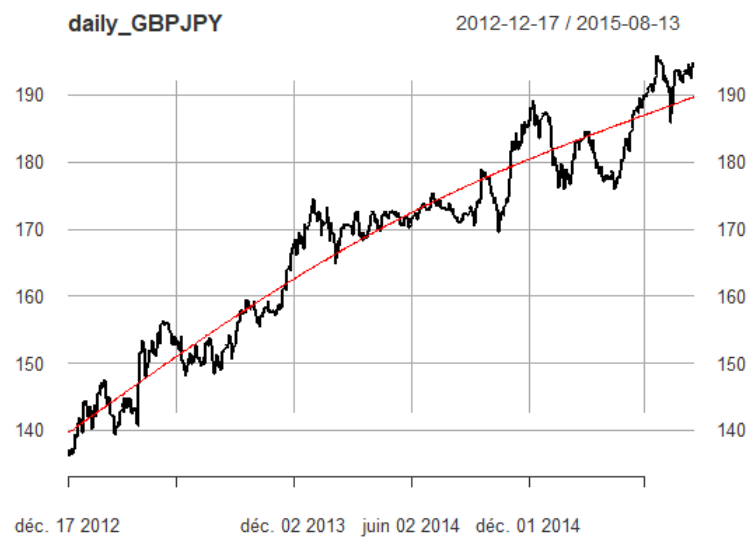
On obtient un résultat similaire à l'estimation par régression linéaire , ce qui confirme notre estimation de tendance.

4.1.5 Estimation de la tendance sur bases de splines

On effectue une estimation sur bases splines.

```
g<-gam(daily_GBPJPY~s(t_daily, k=3))
summary(g)
ychap.gam<-xts(g$fitted,order.by=daily_Date)
plot(daily_GBPJPY,type='l')
lines(ychap.gam,col='red')
summary(daily_GBPJPY - ychap.gam)
```

[figure 13] Estimation de la tendance sur bases spline



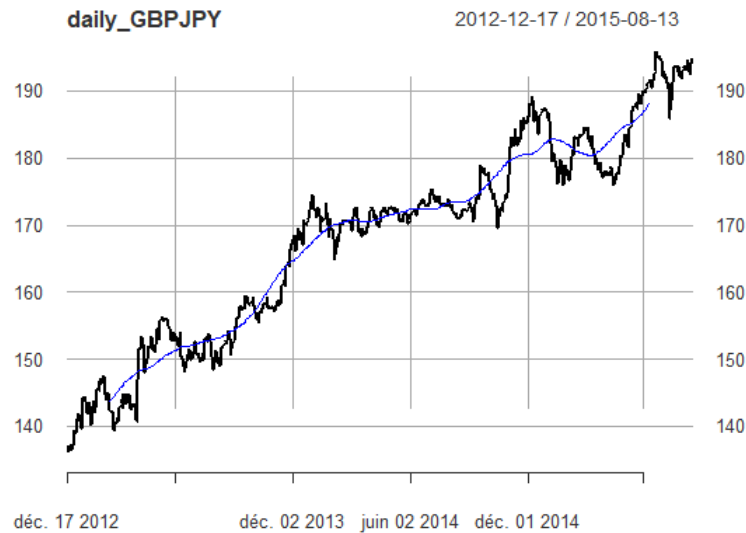
[figure 14] Estimation de la tendance sur bases spline

Même remarque que pour l'estimation précédente.

4.1.6 Estimation de la tendance par moyenne mobile

```
g<-gam(daily_GBPJPY~s(t_daily, k=3))
summary(g)
ychap.gam<-xts(g$fitted,order.by=daily_Date)
plot(daily_GBPJPY,type='l')
lines(ychap.gam,col='red')
summary(daily_GBPJPY - ychap.gam)
```

[figure 15] Estimation de la tendance par moyenne mobile.



[figure 16] Estimation de la tendance par moyenne mobile

Ce fut la dernière estimation de la tendance, dans la section suivante, on élira le meilleur modèle.

4.1.7 Election du modèle sans tendance (detrended)

Pour choisir le modèle de la tendance , qu'on supprimera par la suite de nos données , on a évalué la moyenne des données sans tendance pour les différents estimateurs de la tendance. Pour faire , ceci on a utilisé le code ci dessous (figure 17). Clairement, on voit que le modèle d'estimation par régression linéaire est le meilleur (on compare les valeurs absolues). On va donc l'utiliser pour étudier la partie restante du signal à savoir $S_t + \epsilon_t$ (Sous l'hypothèse qu'on a une saisonnalité bien qu'on y croit pas , on confirmera plus tard qu'il n'y pas de saisonnalité)

```
> ##### Comparaison entre les modèles pour elire le modèle privé de sa tendance
> detrend_splines = daily_GBPJPY - ychap.gam
> detrend_lo = daily_GBPJPY - ychap.daily_lo
> detrend_reg = daily_GBPJPY - reg$fitted
> detrend_ma = daily_GBPJPY - mb
> detrend_kernel = daily_GBPJPY - ychap.kernel
> # On evalue leur moyennes
> mean(detrend_splines)
[1] -2.271287e-13
> mean(detrend_lo)
[1] 0.2793408
> mean(detrend_reg)
[1] 6.168051e-16
> mean(detrend_ma, na.rm=TRUE)
[1] -0.1234543
> mean(detrend_kernel)
[1] -0.01203969
```

[figure 17] comparaison entre différents modèles pour estimer la tendance.

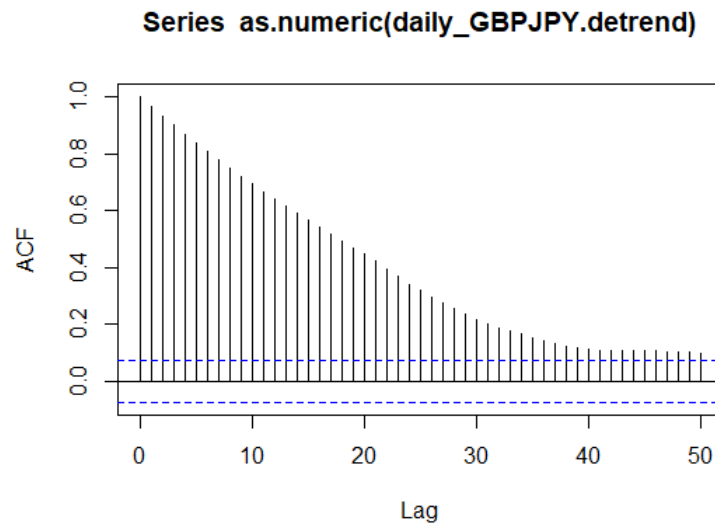
4.2 Etude de la saisonnalité des données

Lorsqu'on affiche le graphique du signal privé de sa tendance (qu'on appellera à présent le résidu) , on obtient la figure suivante :

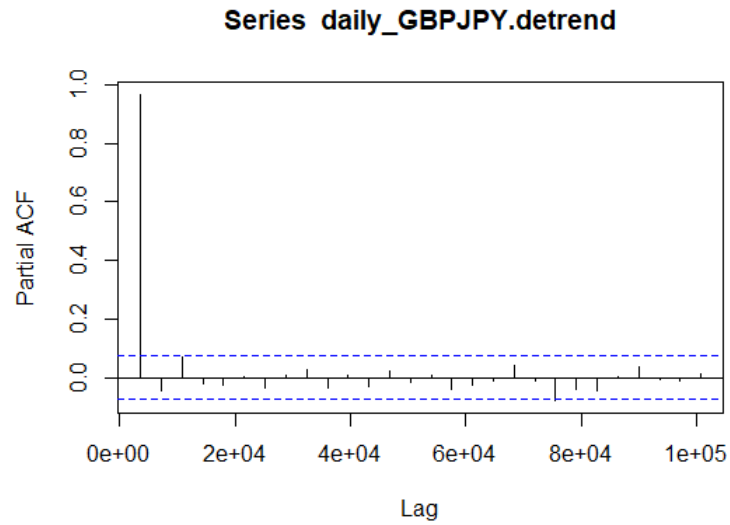


[figure 18] graphique du signal privé de sa tendance (résidu)

On peut remarquer que le résidu n'a pas de tendance, donc on a bien filtré la tendance dans le signal initial. Toutefois, on ne voit pas de motifs répétitifs et donc pas de saisonnalité par conséquent . En affichant le corrélogramme et le diagramme des corrélations partielles, on confirme qu'il n'y a pas de saisonnalité.



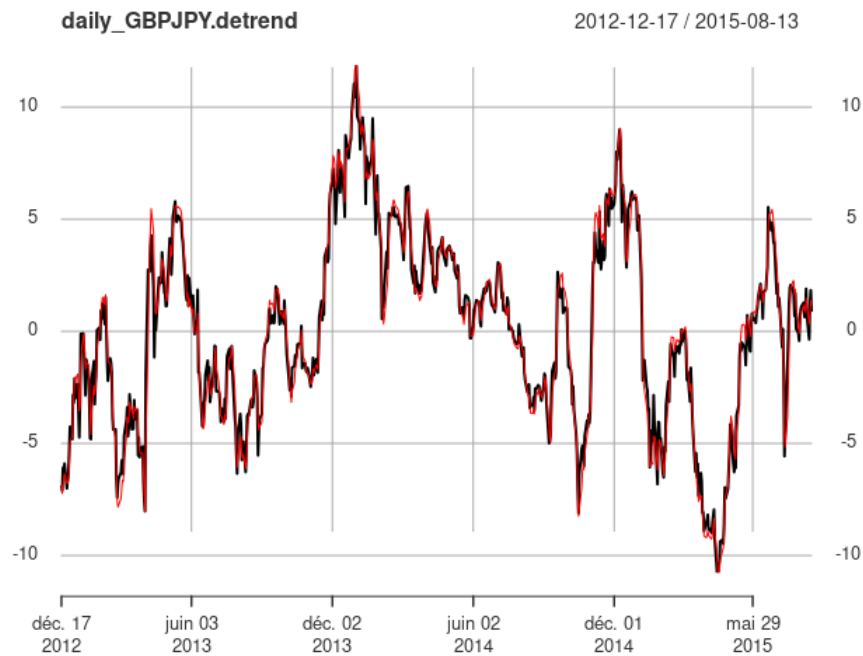
[figure 19] Corrélogramme de daily_GBP.detrend



[figure 20] pacf de daily_GBP.detrend

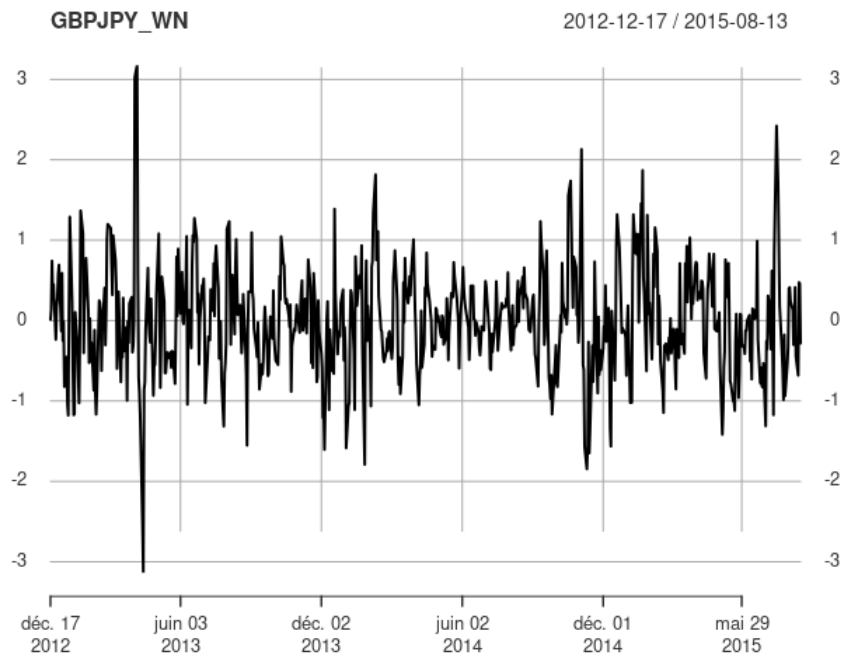
De plus, on voit que le résidu n'est pas un bruit blanc.

On procède donc à un double lissage exponentiel de notre résidu, on obtient la figure suivante :



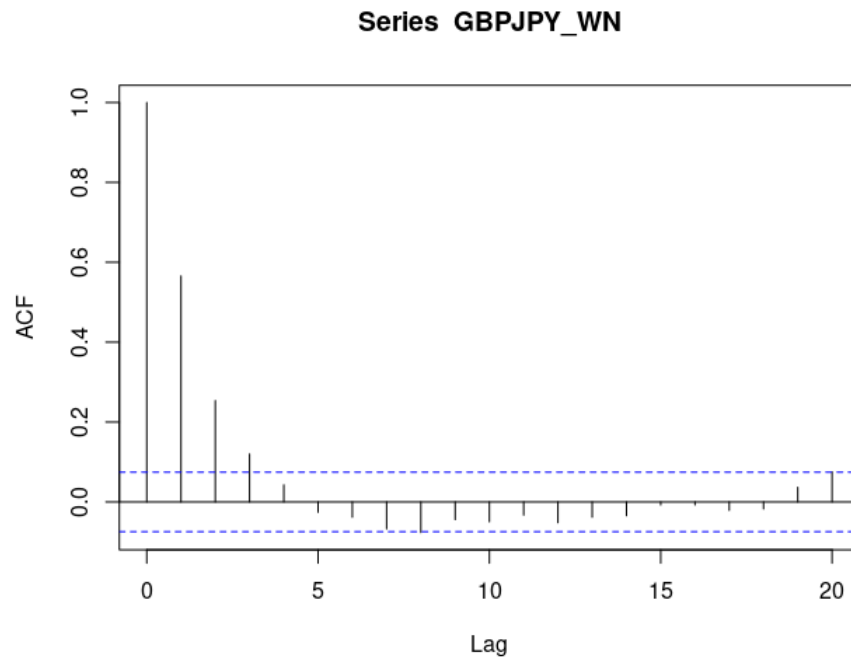
[figure 21] Double lissage exponentiel du (signal - tendance) = daily_GBP.detrend

On opère sur le résidu (detrended signal) et on lui enlève son lissage exponentiel double , on obtient la figure suivante :

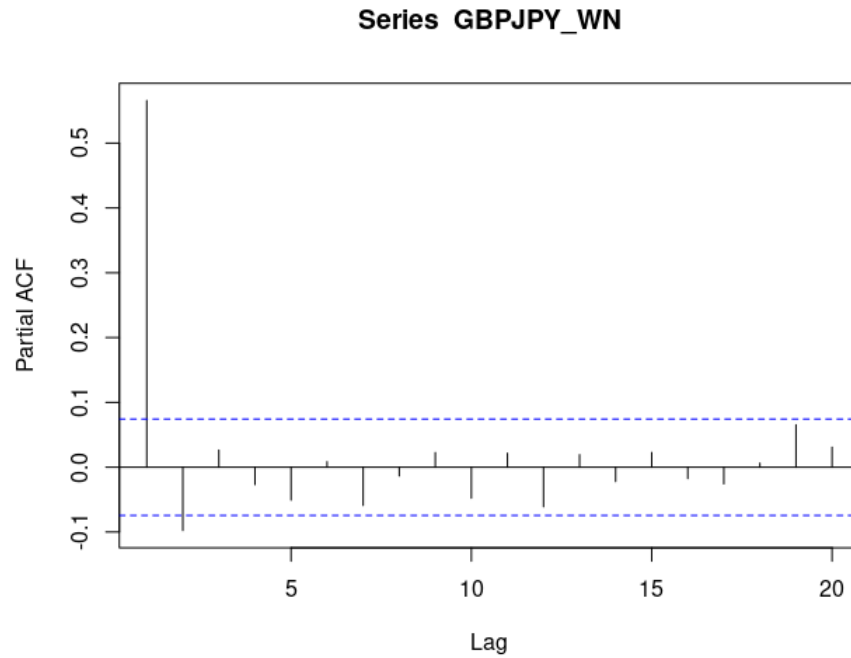


[figure 22] figure de (detrended signal - son double lissage exponentiel)

On doit donc analyser son corrélogramme et son diagramme des corrélations partielles pour pouvoir tirer des conclusions :



[figure 22] acf de (detrended signal - son double lissage exponentiel)



[figure 23] pacf de (detrended signal - son double lissage exponentiel)

Cette fois, le résultat obtenu est largement meilleur, d'après l'analyse de son acf et son pacf, on voit qu'il s'agit d'un processus auto-régressif d'ordre 2 .

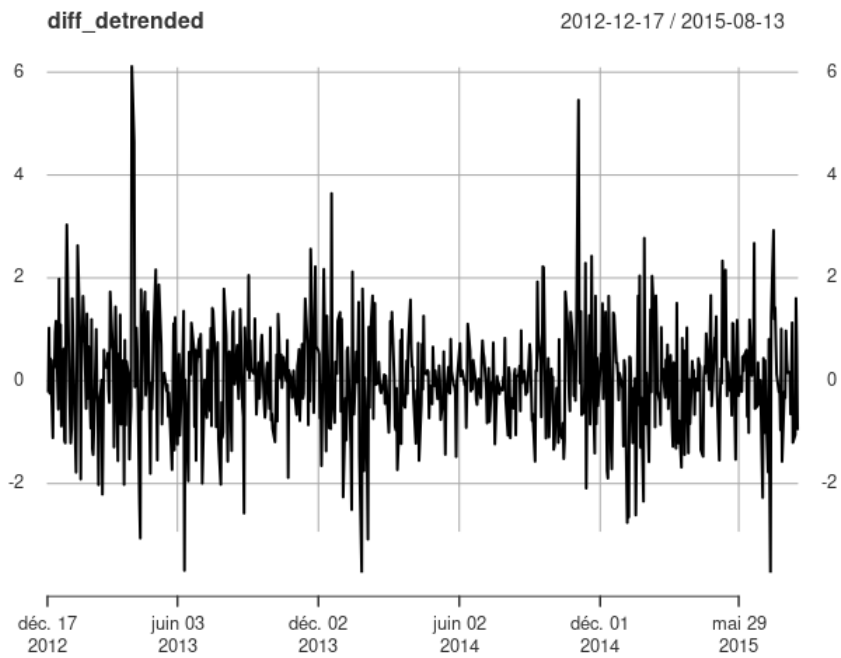
On a également essayé avec la méthode Holt-Winters, mais les résultats n'ont pas été convaincants.

4.3 Choix du modèle

On a décidé donc d'exploiter une nouvelle piste, celle d'évaluer la différence des returns détrendés. Pour ceci on exécute le code suivant :

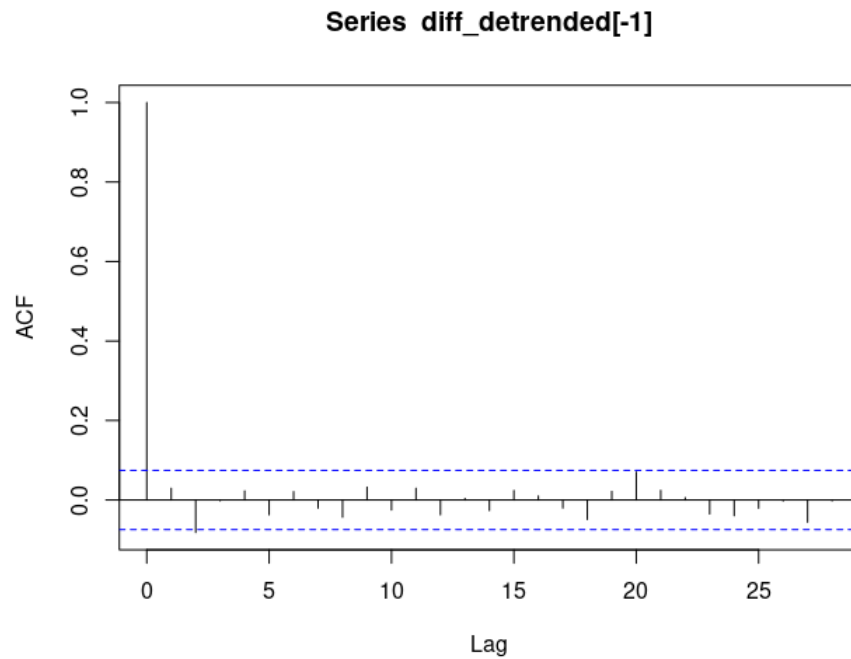
```
> diff_detrended=diff(daily_GBPJPY.detrend)
> plot(diff_detrended)
```

[figure 23] Code pour évaluer la différence des returns détrendés

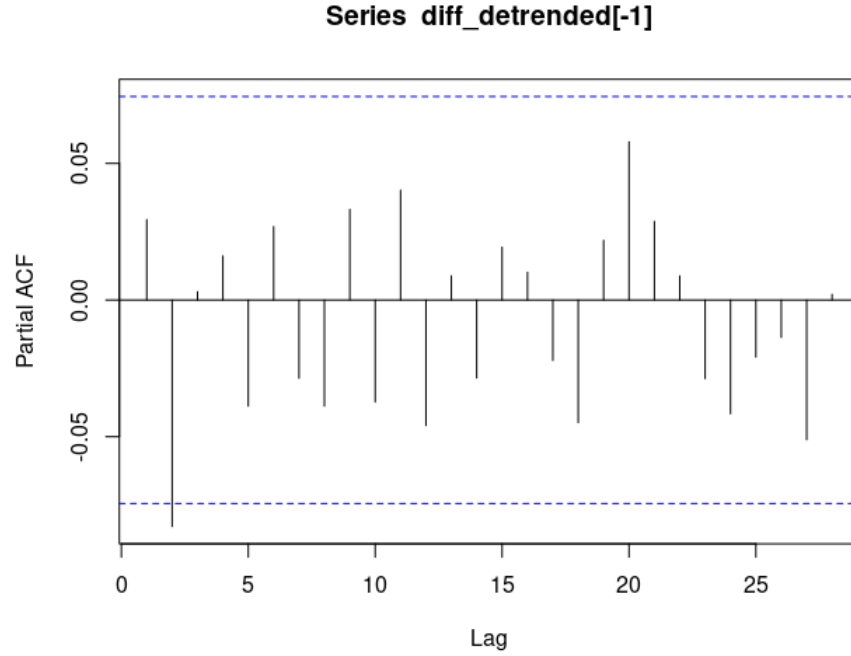


[figure 24] graphique du signal : différence des returns détrendés

On vérifie que c'est un bruit blanc, pour ce faire , on étudie son corrélogramme et son diagramme de corrélations partielles :



[figure 25] acf du signal différence des returns détrendés



[figure 26] pacf du signal différence des returns détrendés

D'après l'analyse de l'acf et pacf de diff_detrended, on voit clairement qu'il s'agit d'un bruit blanc gaussien dont les paramètres sont:

mean=0.01118511, sd =1.067107.

Le modèle est donc :

$$X_t = T_t + I_t$$

I_t étant le résidu qui était de nature inconnue

On a donc : $I_t = X_t - T_t$

Or on a montré que les I_t étaient des réalisations de $N(0.01118511, 1.067107)$ i.i.d

Donc : $(X_t - T_t) - (X_{t-1} - T_{t-1}) = \epsilon_t$

où ϵ_t suit $N(0.01118511, 1.067107)$

On conclut que notre signal peut être modélisé par :

$$X_t = (X_{t-1} - T_{t-1}) + T_t + \epsilon_t$$

Cette modélisation est très puissante, d'ailleurs dans la section suivante (section 5) , elle prédira très bien l'évolution du prix (forecasting) , voir figure 32 .

5 Simulation de prévision sur un échantillon test

Dans cette section on réalise la prévision sur un échantillon test de notre dataset. On a commencé par une prévision par lissage exponentiel simple, le code utilisé est : (ici screen code pred lissage simple)

```
# forecasting with simple exponential smoothing:
n1<-94
n0<-n_daily-n1

alpha<-seq(0.05,0.95,length=100)
forecast<-lapply(alpha,expSmooth,x=as.vector(daily_GBPJPY))
erreur<-unlist(
  lapply(forecast,
    function(x){mean((tail(daily_GBPJPY,n_daily-1)-head(x,n_daily-1))^2)}))
plot(alpha,erreur,type='l')

daily_GBPJPY.smooth.simple<-expSmooth(as.vector(daily_GBPJPY),alpha=alpha[which.min(erreur)])

daily_GBPJPY.smooth.simple.forecast<-predict.expSmooth(daily_GBPJPY.smooth.simple,n0,n1,smooth.type="simple")
plot(daily_GBPJPY)
lines(xts(daily_GBPJPY.smooth.simple.forecast,order.by = daily_Date),col='red')
```

[figure 27] Code du forecasting avec lissage exponentiel simple

On obtient la figure suivante : (screen figure pred lissage simple)



[figure 28] forecasting avec lissage exponentiel simple

Dans notre cas de figure, le lissage exponentiel simple ne s'avère pas efficace, on a obtenu une droite horizontale, notons que le choix du alpha a été le plus optimal parmi les valeurs possibles. Il pourrait être efficace pour des modèles avec peu de fluctuations ou stationnaires.

On effectue une deuxième tentative de prévision, cette fois ci avec un modèle de lissage exponentiel double :

```

# forecasting with double exponential :
alpha = seq(0.05,0.95,length=100)
forecast = lapply(alpha, ExpSmooth, x=daily_GBPJPY)
erreur = sapply(forecast, mse_smoothing_simple, true_values=df_daily_GBPJPY$Close)
X.d.exp.mooHW<-DoubleExpSmooth(as.vector(daily_GBPJPY), alpha[which.min(erreur)])
prevHW<-predict.expSmooth(X.d.exp.mooHW,
                          inst=600,horizon=94, smooth.type="double")
plot(daily_GBPJPY,pch=20,ylim=range(daily_GBPJPY,prevHW),type='l')
lines(xts(prevHW,order.by = daily_Date),col='red',lwd=2)
abline(v=600,lty='dashed')

```

[figure 29] code forecasting avec lissage exponentiel double



[figure 30] graphique du forecasting avec lissage exponentiel double

On a obtenu une droite oblique pour un choix optimal de alpha, le lissage exponentiel double est insuffisant pour la prédiction .

```

# simulation:
Xt<-as.vector(daily_GBPJPY)[1:600]
Tt<-as.vector(ychap.lm)
eps=rnorm(694, mean = 0, sd = 1.067107)
for (i in seq(600,693)) {
  Xt<-append(Xt,(Xt[i-1]-Tt[i-1])+Tt[i]+eps[i])
}
plot(daily_GBPJPY)
lines(xts(Xt,order.by = daily_Date),col = 'red')

```

[figure 31] Code du modèle défini pour modéliser et prédire la paire GBP-JPY en section 4.3



[figure 32] Forecasting avec le modèle défini pour modéliser et prédire la paire GBP-JPY en section 4.3

6 Conclusion

En guise de conclusion, ce projet a été pour nous l'occasion de découvrir et mettre en oeuvre plusieurs méthodes vues en cours pour étudier des notions fondamentales des séries temporelles telles que la saisonnalité, la tendance, la stationnarité, les corrélations, le bruit blanc et les résidus. Toutefois, on s'est aperçu rapidement que notre dataset n'était pas facile à traiter vu qu'il s'agit des données du prix d'un taux de change (GBP/JPY), il appartient donc au marché Forex qui est l'un des marchés les plus difficile à trader. On a pu remarquer ceci surtout lors de l'étude de la saisonnalité où on a pas trouvé de caractère saisonnier, ceci était attendu : en effet si le marché était saisonnier, il serait facile à trader, ce qui n'est pas le cas. Les mouvements des taux de change sont très difficiles à prévoir avec précision, néanmoins on a montré que les returns du signal detrended ($\text{diff}(\text{daily_GBPJPY}.\text{detrend})$) étaient des réalisations d'un bruit blanc.

Plus important, on a réussi à construire un modèle qui prédit l'évolution du prix de la paire GBP-JPY.