

Introduction

Dans le monde de l'économie, il est impératif d'effectuer des études statistiques analysant le comportement des séries temporelles les plus importantes, pour pouvoir effectuer des prévisions utiles pour la conjoncture économique.

Au cours de ce stage, j'ai eu l'occasion de manipuler diverses séries temporelles à savoir la série de la production, la consommation, d'importation et de subvention des produits pétroliers...

Pour plus de rigueur, j'ai choisi de présenter l'étude relative à l'une des séries abordées à savoir **la série temporelle de la consommation des produits pétroliers au cours d'une période de donnée.**

Problématique : La demande d'énergie traverse-t-elle une phase d'expansion rapide, ralentie, de stagnation, de récession ?

Objectif

Correction des variations saisonnières de la série et détection du modèle adéquat pour effectuer une prévision et analyse des facteurs influant sur l'évolution dans le temps.

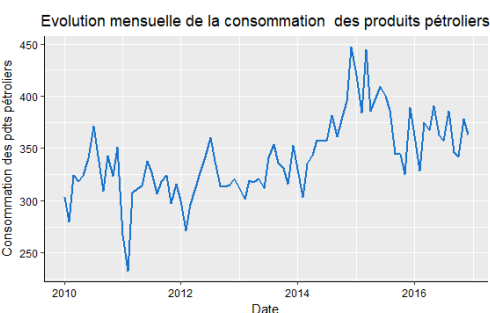
Pour la réalisation de ce travail on a utilisé les logiciels Excel et R.

Présentation de la série

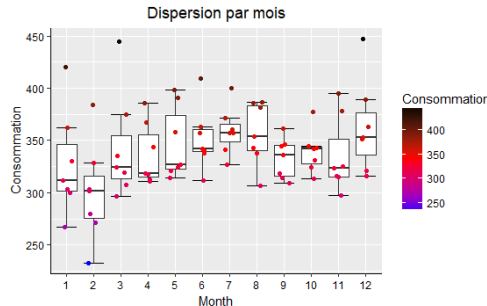
Il s'agit d'une série mensuelle représentant la consommation totale des principaux produits pétroliers (Essences, Pétrole lampant, Fuel, Gasoil...) en Kt de janvier 2010 à décembre 2016.

La source de cette série le site officiel du Ministère de l'Industrie, de l'Energie et des Mines (Open Data).

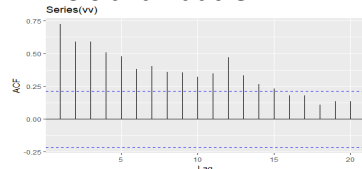
Visualisation



On remarque une allure de la courbe similaire chaque année. Cela indique l'existence de saisonnalité.

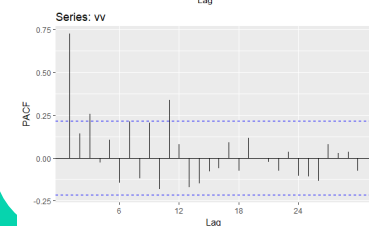


La dispersion de la consommation diffère d'un mois à un autre.



• La forme géométrique apparente au niveau du corrélogramme, plus claire dans la partie autocorrélation simple, confirme l'aspect saisonnier de la série

• On remarque la présence d'un pic au niveau du 1^{er} et 12^{ème} retard.



Etude de stationnarité

D'après l'ADF test la série n'est pas stationnaire. On applique la différence première.

```
> library(tseries)
> adf.test(vv, alternative = c("stationary", "explosive"))

Augmented Dickey-Fuller Test

data: vv
Dickey-Fuller = -2.519, Lag order = 4, p-value = 0.363
alternative hypothesis: stationary

> adf.test(diff(vv), alternative = c("stationary", "explosive"))

Augmented Dickey-Fuller Test

data: diff(vv)
Dickey-Fuller = -4.3969, Lag order = 4, p-value = 0.01
alternative hypothesis: stationary
```

On a l'ADF test statistique est supérieure à la t-statistic relative au seuil 5% en valeur absolue, confirmé par un p-value nul (< 0.05).

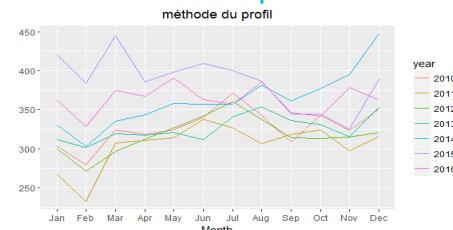
D'où la stationnarité de la série différenciée pour un seuil de 5%.

Il s'agit donc d'un **processus DS**.

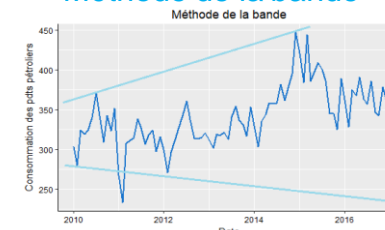
Décomposition de la série

On commence par déterminer si le modèle est additif ou multiplicatif

Méthode du profil



Méthode de la bande



On ne peut pas conclure quant au parallélisme des courbes.

On essaie la méthode de la bande.

Les 2 droites ne sont pas parallèles (entonnoir). Il s'agit a priori d'un modèle multiplicatif.

Pour un **modèle multiplicatif** : $Y_t = T_t \times S_t \times I_t$
avec T_t : tendance, S_t : saisonnalité, I_t : irrégulier

Correction des variations saisonnières

Pour corriger les variations saisonnières, sous Excel, certains calculs sont à effectuer dans le cas d'un modèle multiplicatif :

• La moyenne mobile d'ordre 12: $MM(12)_t = [(\sum_{i=t-5}^{t+5} Y_i) + (Y_{t-6} + Y_{t+6})/2] / 12$

• Le rapport $Y / MM(12)$

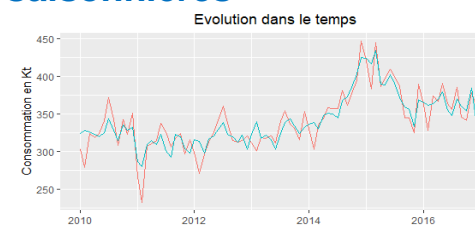
• Le coefficient de saisonnalité S_t qui est la moyenne de ce rapport respective pour chaque mois.

• On obtient la série corrigée des variations saisonnières $Y_{(cvs)t} = Y_t / S_t$

Résultats obtenus sous R

Série corrigée des variation saisonnières

Sous R, on peut corriger directement les variations saisonnières pour obtenir les courbes suivantes:



Détection du modèle

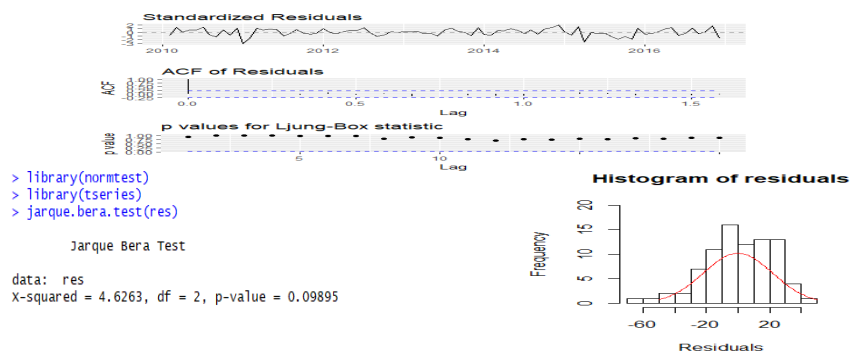
```
> fit <- auto.arima(diff(vv), d = 1, D = 0, max.p = 2, max.q = 1, max.P = 1, max.Q = 1,
+ stationary = TRUE, ic = "aic", stepwise=FALSE, trace=FALSE,
+ test = c("kpss", "adf", "pp"), seasonal.test = c("ocsb", "ch"))
> fit
Series: diff(vv)
ARIMA(2,0,0)(1,0,0)[12] with zero mean

Coefficients:
ar1      ar2      sar1
-0.2915  -0.3279  0.5173
s.e.      0.1072   0.1069  0.1015

sigma^2 estimated as 491.6: log likelihood=-375.46
AIC=758.91  AICc=759.43  BIC=768.59
```

Sous Rstudio, on a détecté le modèle saisonnier **SARIMA (2,0,0)(1,0,0)** en estimant les coefficients AR, MA et le coefficient saisonnier SMA et SAR.

Validation du modèle



```
> library(nortest)
> library(tseries)
> jarque.bera.test(res)

Jarque Bera Test

data: res
X-squared = 4.6263, df = 2, p-value = 0.09895
```

- Nous remarquons bien la normalité des résidus p-value > 5% (confirmée par l'allure normale de l'histogramme).
- Absence d'autocorrélation des résidus : p-value du test Ljung-Box = 0,7 > 5% et confirmée par l'ACF des résidus.

➡ Notre modèle est bien confirmé.

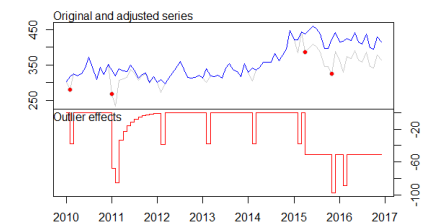
Détection des points atypiques

```
> library(tsooutliers)
> outliers <- tso(vv, types = c("ao", "ls", "tc", "sls"))
Series: vv
Regression with ARIMA(0,1,1)(0,0,1)[12] errors

Coefficients:
ma1      sma1      SLS2      TC13      L564      AO71
-0.3446  0.3415  -37.9095  -67.4638  -51.3683  -46.7775
s.e.      0.1248  0.1243   7.0618   16.4738   16.1666   13.7890

sigma^2 estimated as 324.7: log likelihood=-355.45
AIC=724.9  AICc=726.4  BIC=741.84

Outliers:
type ind time coefhat tstat
1 SLS 2 2010:02 -37.91 -5.368
2 TC 13 2011:01 -67.46 -4.095
3 LS 64 2015:04 -51.37 -3.177
4 AO 71 2015:11 -46.78 -3.392
```

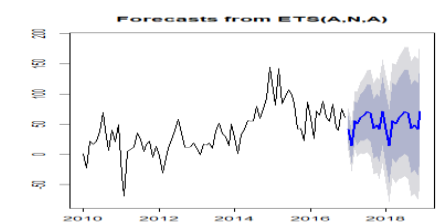


On a 4 types de points atypiques détectés:

- AO ; Additive Outlier
- TC : Temporary Change
- LS : Level Shift
- SLS : Seasonal Level Shift

Prévision

D'après le modèle détecté précédemment SARIMA(2,0,0)(1,0,0) on a pu avoir la prévision suivante pour l'année 2018:



Conclusion

Malgré qu'on a pu prévoir le comportement futur de la série, il est possible que cette prévision soit différente de la réalité à cause de perturbations dues éventuellement à des circonstances socioéconomiques particulières. Dans ce cas de nouveaux points atypiques seraient détectés et doivent être analysés. On doit donc réévaluer régulièrement les modèles d'évolution temporelle de ces séries pour plus de précision.