Séries temporelles

Jean Gaudart
Aix-Marseille Université
UMR 912, SESSTIM (AMU, INSERM, IRD)

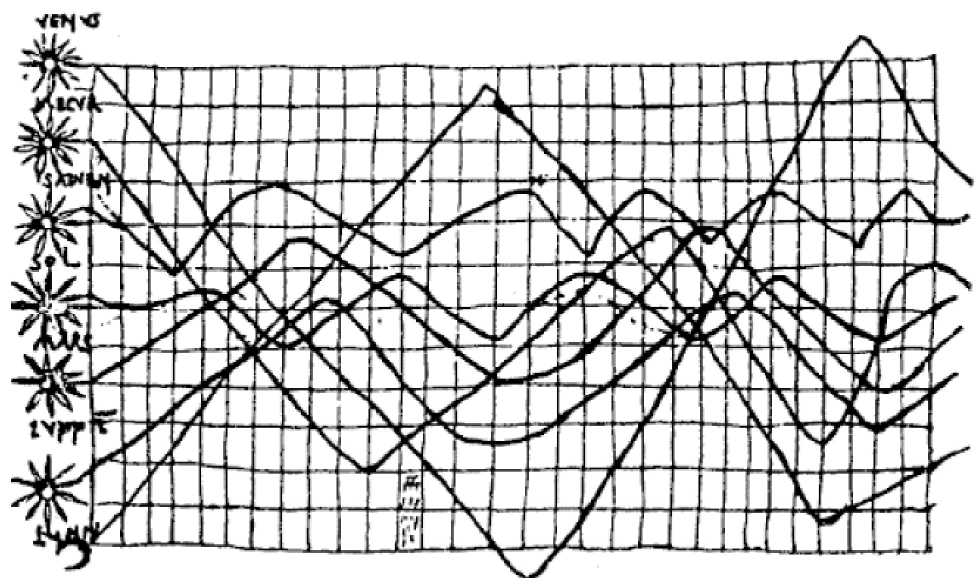
plan

1. Introduction
2. Analyse graphique
3. Décomposition d’une série
4. Corrélation
5. Cofacteur
6. Lissage exponentiel
7. Modèles Stochastiques
8. Régression
9. Séries stationnaires = ARMA
10. Séries non-stationnaires = ARIMA

I. Introduction

- séries temporelles (ou chronologiques) => analyse statistique d’observations régulièrement espacées dans le temps
- Objectif: comprendre le passé et prévoir l’avenir
- => étudier les principales caractéristiques des données et les variations aléatoires
- Histoire: astronomie

Xe siècle, inclinaisons de l’orbite des planètes en fonctions du temps => loi sur les mouvements des planètes



1ères prédictions: Halley 18^{e}
Observations de la comètes 1531, 1607, 1682 => prédiction 1758

- Premières approches: détecter des saisonnalités cachées au sein des données => approche « fréquence » = analyse harmonique (Fourier): estimer par une somme pondérée de fonctions sinus/cosinus

Modèle:

$$Y_t = \sum_j \left[\alpha_j \cos(\omega_j t) + \beta_j \sin(\omega_j t)\right] + \varepsilon_t$$

$$Y_t = \sum_j \left[A_j \cos(\omega_j t - \varphi_j)\right] + \varepsilon_t$$

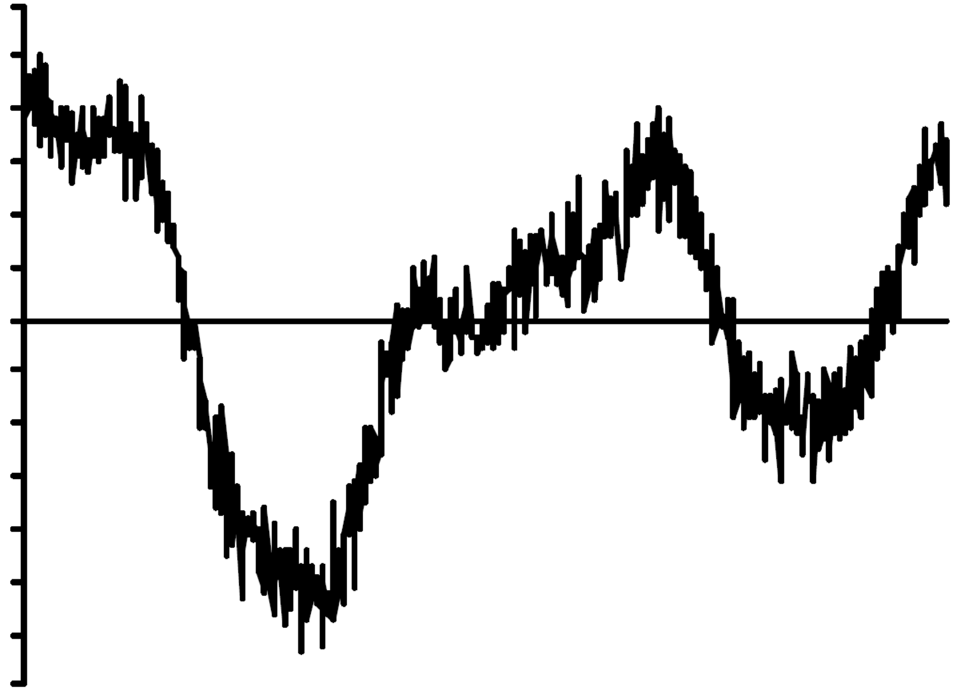
avec $A_j = \sqrt{\alpha_j^2 + \beta_j^2}$ Poids = L'amplitude de la j$^{\text{ième}}$ composante périodique (harmonique)

ω_j Fréquence (nb de répétitions par unité de temps), avec période (durée entre 2 répétitions) $T = \frac{2\pi}{\omega}$

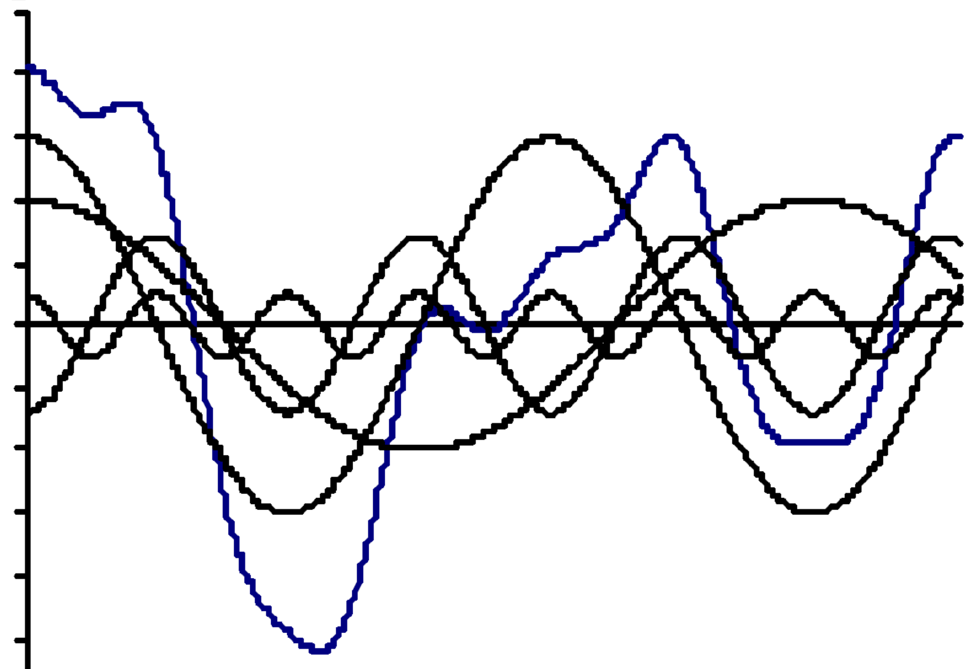
$\varphi_j = \tan^{-1} \frac{\beta_j}{\alpha_j}$ Phase (décalage) ($\alpha_j \neq 0$)

ε_t Suite de V.A. i.i.d de moyenne 0 et de variance σ^2_ε= Bruit Blanc

Série temporelle observée



4 fonctions sinusoïdales d’amplitudes A différentes
La courbe bleue est la somme des 4 harmoniques



En astronomie: observation de cycles parfaitement périodiques
Autres domaines (économie, sciences de la vie…) : ce n'est pas le cas

Yule (1927): Modèle auto-Regressif AR

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \varepsilon_t$$

Généralisable AR(p)

$$Y_t = \alpha_1 Y_{t-1} + \ldots + \alpha_p Y_{t-p} + \varepsilon_t$$

Slutsky(1927): Modèle Moyennes Mobiles MA

$$Y_t = \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \varepsilon_t$$

Généralisable MA(q)

$$Y_t = \beta_1 \varepsilon_{t-1} + \ldots + \beta_q \varepsilon_{t-q} + \varepsilon_t$$

ε_t V.A. i.i.d de moyenne 0 et de variance σ^2_ε= Bruit Blanc

Développement:

Un processus Y_t est un processus ARMA(p,q) s'il existe ε_t (bruit blanc) tel que

$$\forall t,$$

$$Y_t = \underbrace{\alpha_1 Y_{t-1} + \ldots + \alpha_p Y_{t-p} + \varepsilon_t}_{\text{AR(p)}} + \underbrace{\beta_1 \varepsilon_{t-1} + \ldots + \beta_q \varepsilon_{t-q}}_{\text{MA(q)}}$$

Poursuite de l'histoire

Prévisions structurelles

1936, Keynes: $\mathrm{Conso}_t = \alpha \,\mathrm{Rev}_t + \beta$

1952, Brown: $\mathrm{Conso}_t = \alpha \,\mathrm{Rev}_t + \beta + \gamma \mathrm{Conso}_{t-1}$

1970, Box-Jenkins (prévisions non structurelles): déterminer les différentes composantes d'une série temporelle, et faire de la prévision à partir d'une série stationnaire.

Variable mesurée séquentiellement dans le temps sur un intervalle fixé (intervalle d’échantillonnage) => les données observées forment une série temporelle historique.
Ces observations sont des réalisations d’une séquence de variables aléatoires.
Une séquence de variables aléatoires (à intervalles d’échantillonnage fixes) est parfois appelée processus stochastique en temps discret, ou modèle de série temporelle.

Les principales caractéristiques recherchées sont une tendance ou une variation saisonnière, qui peuvent être modélisées de façon déterministe par des fonctions du temps.
Un aspect important des séries temporelles est que les observations proches en temps tendent à être corrélées (dépendance temporelle). La plupart des méthodes ont pour but d’expliquer cette dépendance et les caractéristiques des observations. Une fois la modélisation faite, le modèle peut être utilisé pour prévoir le futur.

On ne peut jamais rien prévoir …

- 1873, « *L'abdomen, la poitrine et le cerveau sont à jamais interdits à l'intrusion de la connaissance et de la chirurgie humaine* » John Eric Ericksen, médecin personnel de la Reine Victoria.
- 1876, « *Le téléphone a bien trop de défauts et de lacunes pour que nous le considérions sérieusement comme un moyen de communication. Cet appareil n'a aucune valeur à nos yeux* » Note de service, Western Union.
- 1895, *« Il est impossible d'imaginer des machines volantes plus lourdes que l'air* » William T. Kelvin, président de la Société Royale des Sciences.
- 1929, juste avant la crise, « *Le marché de la Bourse semble avoir atteint un haut plateau permanent* » Irving Fisher, Professeur d'économie à l'Université de Yale.
- 1943, « *Je crois que le marché mondial pourrait peut-être accueillir cinq ordinateurs* » Thomas Watson, président d'IBM.
- 1977, « *Il n'y a aucune raison de vouloir posséder un ordinateur à la maison* » Ken Olson, PDG-fondateur de la société Digital Equipment.

II. Graphiques, tendances et saison

Exemple 1: nombre d'enregistrements sur la Pan Am (USA) ,1949-1960

```
data(AirPassengers)
AP<-AirPassengers
AP
```

?De quelle classe est l'objet « AP » ?

```
class(AP)

[1] "ts"
```

La classe `ts` (time series) => spécifique des séries temporelle.

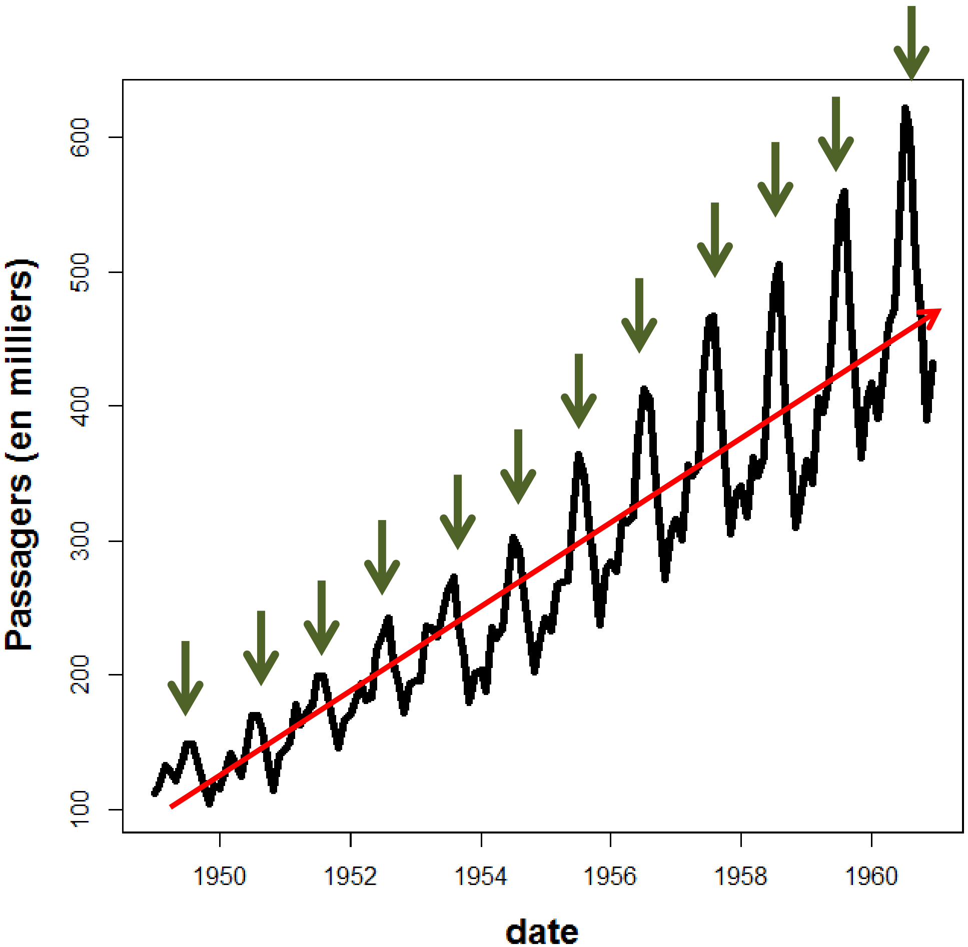
début de la série, 1iere année	`start(AP)`	`[1] 1949    1`
Fin de la série, 12ieme année	`end(AP)`	`[1] 1960   12`
Fréquence des observations, 12/an	`frequency(AP)`	`[1] 12`

pour toute série d'observation,=>spécifier objet de classe `ts`

```
ts(Y, start=xx, end=yy, frequency=zz)
```

Visualiser la série temporelle

```
plot(AP,ylab="Passagers (en milliers)", xlab="date")
```



• un changement systématique non périodique => tendance

•un changement répété => saisonnalité, (même pour période non annuelle) :

Recherche de causes explicatives, pour améliorer la modélisation:

• Tendance:

$\Rightarrow$augmentation du niveau de vie

$\Rightarrow$disponibilité et sécurité des avions

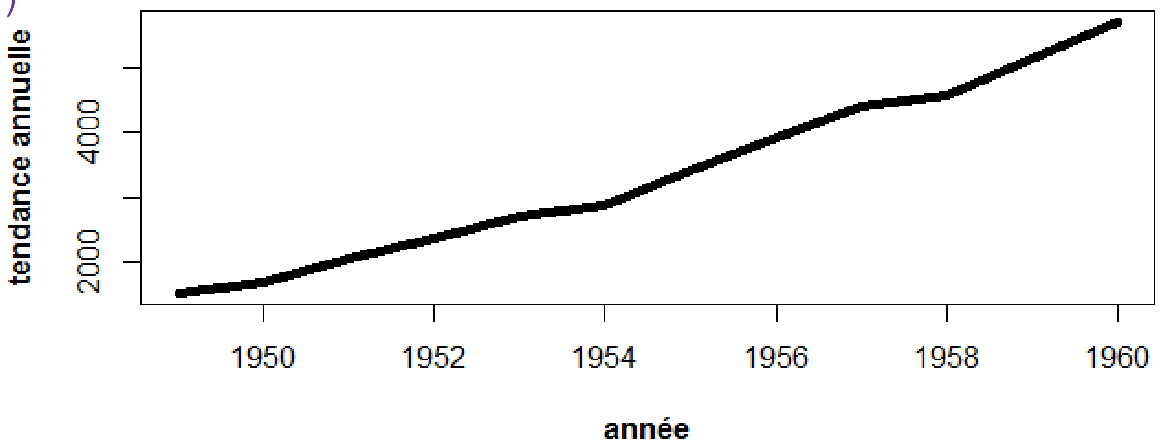
• Saison:

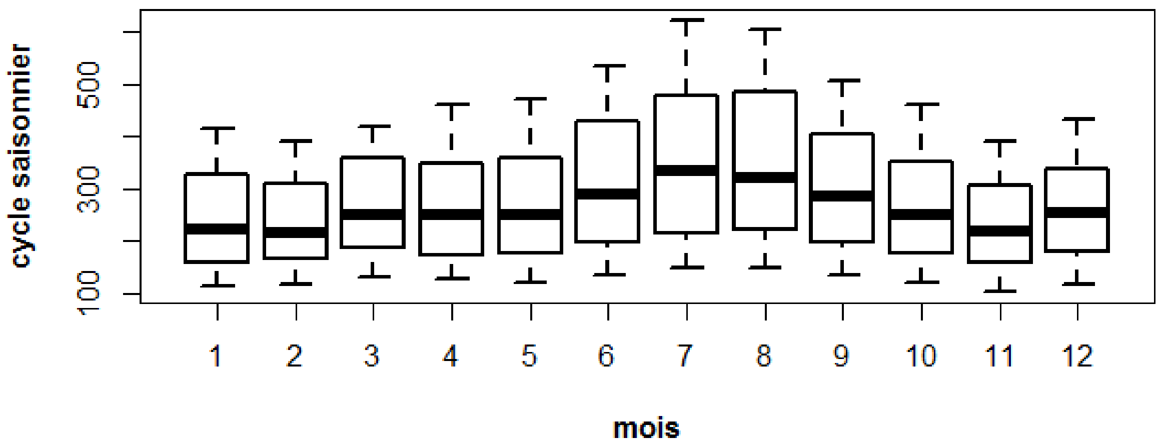
$\Rightarrow$période haute, été (juin-aout)

$\Rightarrow$période basse, automne/hivers (nov-fev)

Visualiser la tendance et la saisonnalité

```
layout(1:2)
plot(aggregate(AP))
boxplot(AP~cycle(AP))
```





Exemple 2: taux de chômage 01-1996 à 08-2006, Maine (USA)

```
www <-
"http://www.mif.vu.lt/~rlapinskas/2012-2013/Erasmus/Data/Cowpertwait,Metcalfe/Maine.dat"

Maine.month<-read.table(www,header=TRUE)
```

OU, à partir du fichier téléchargé depuis AMeTICE

```
Maine.month<-read.table("C:/Mon_dossier/Maine.dat",header=TRUE)

attach(Maine.month)
class(Maine.month)        [1] "data.frame"
```

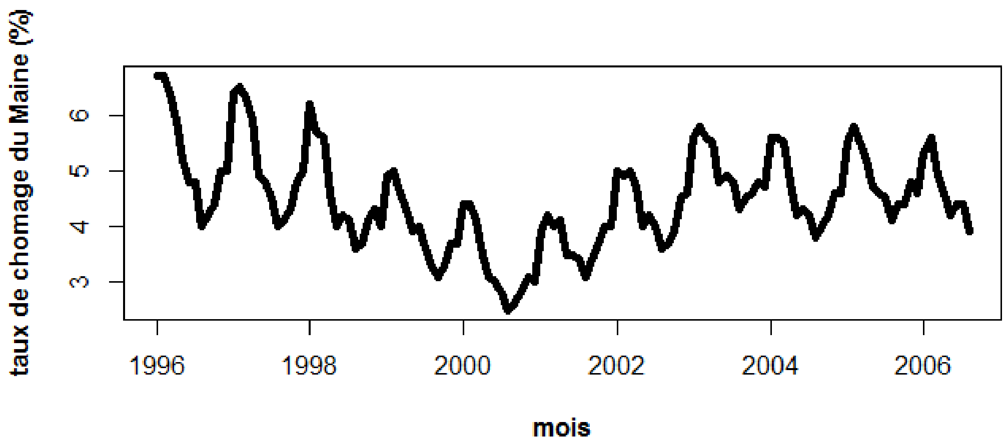
C'est un dataframe, pas une classe ts => il faut donc convertir

```
Maine.m.ts<-ts(unemploy,start=c(1996,1),freq=12)
```

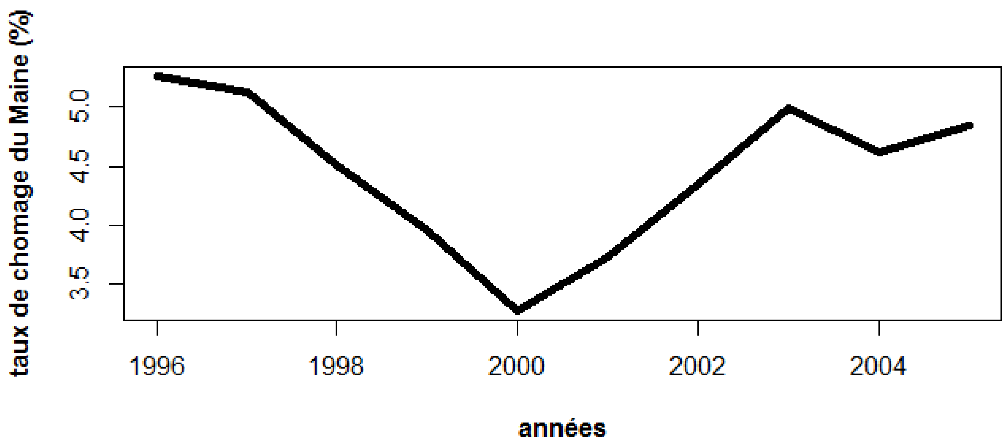
Pour avoir les taux annuels:

```
Maine.an.ts<-aggregate(Maine.m.ts)/12
```

```
layout(1:2)
plot(Maine.m.ts,ylab="taux de chomage,Maine(%)",xlab="mois")
plot(Maine.an.ts,ylab="taux de chomage,Maine(%)",xlab="années")
```



variations mensuelles =>
février >> moyennes
Aout << moyennes



variations inter-annuelles=>
2000 << moyennes

Expliquer les variations mensuelles:
Extraire des fenêtres de temps: février et Aout

```
Maine.Feb<-window(Maine.m.ts, start=c(1996,2), freq=TRUE)
Maine.Aug<-window(Maine.m.ts, start=c(1996,8), freq=TRUE)
```

Calculer le ratio de chaque fenêtre par rapport à la moyenne

```
Feb.ratio<-mean(Maine.Feb)/mean(Maine.m.ts)
Aug.ratio<-mean(Maine.Aug)/mean(Maine.m.ts)
Feb.ratio
[1] 1.222529
Aug.ratio
[1] 0.8163732
```

? Expliquer cette variation intra-annuelle ?

Exemple 3: taux de chômage 01-1996 à 08-2006, USA

```
www <-
"http://www.mif.vu.lt/~rlapinskas/2012-2013/Erasmus/Data/Cowpertwait,Metcalfe/USunemp.dat"

USA.month<-read.table(www,header=TRUE)
```

OU, à partir du fichier téléchargé depuis AMeTICE

```
USA.month<-read.table("C:/Mon_dossier/USunemp.dat",header=TRUE)

attach(USA.month)
class(USA.month)
```

```
[1] "data.frame"
```

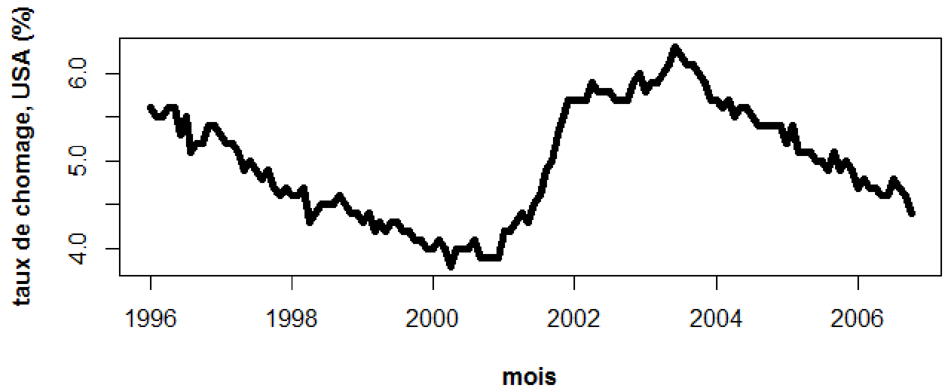
C'est un dataframe, pas une classe ts => il faut donc convertir

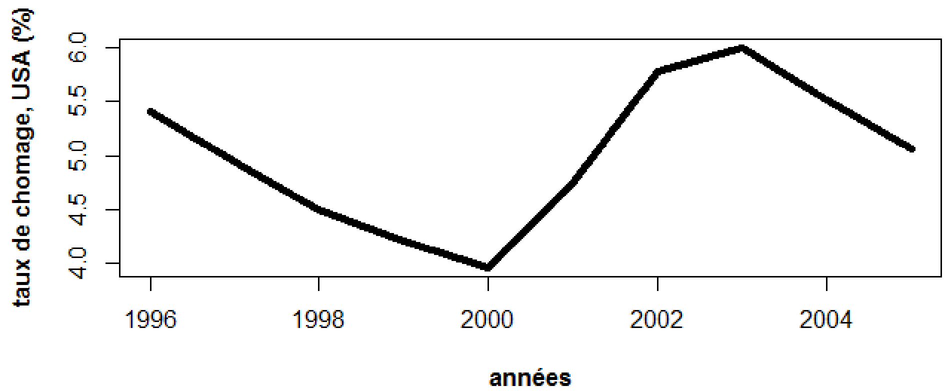
```
USA.m.ts<-ts(USun,start=c(1996,1),freq=12)
```

Pour avoir les taux annuels:

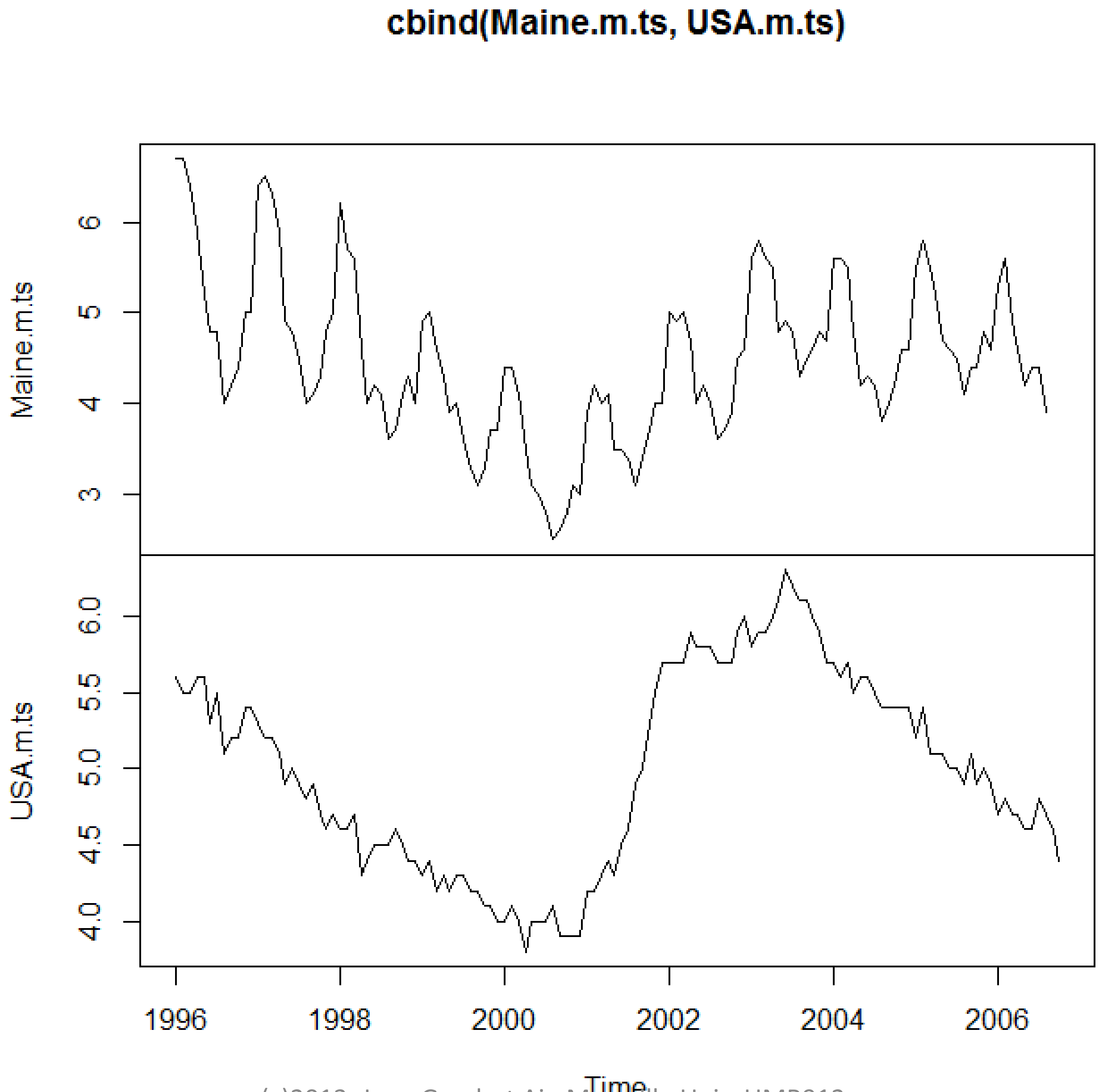
```
USA.an.ts<-aggregate(USA.m.ts)/12
```

```
layout(1:2)
plot(USA.m.ts, ylab="taux de chomage, USA (%)", xlab="mois")
plot(USA.an.ts, ylab="taux de chomage, USA (%)", xlab="années")
```





```
plot(cbind(Maine.m.ts,USA.m.ts))
```



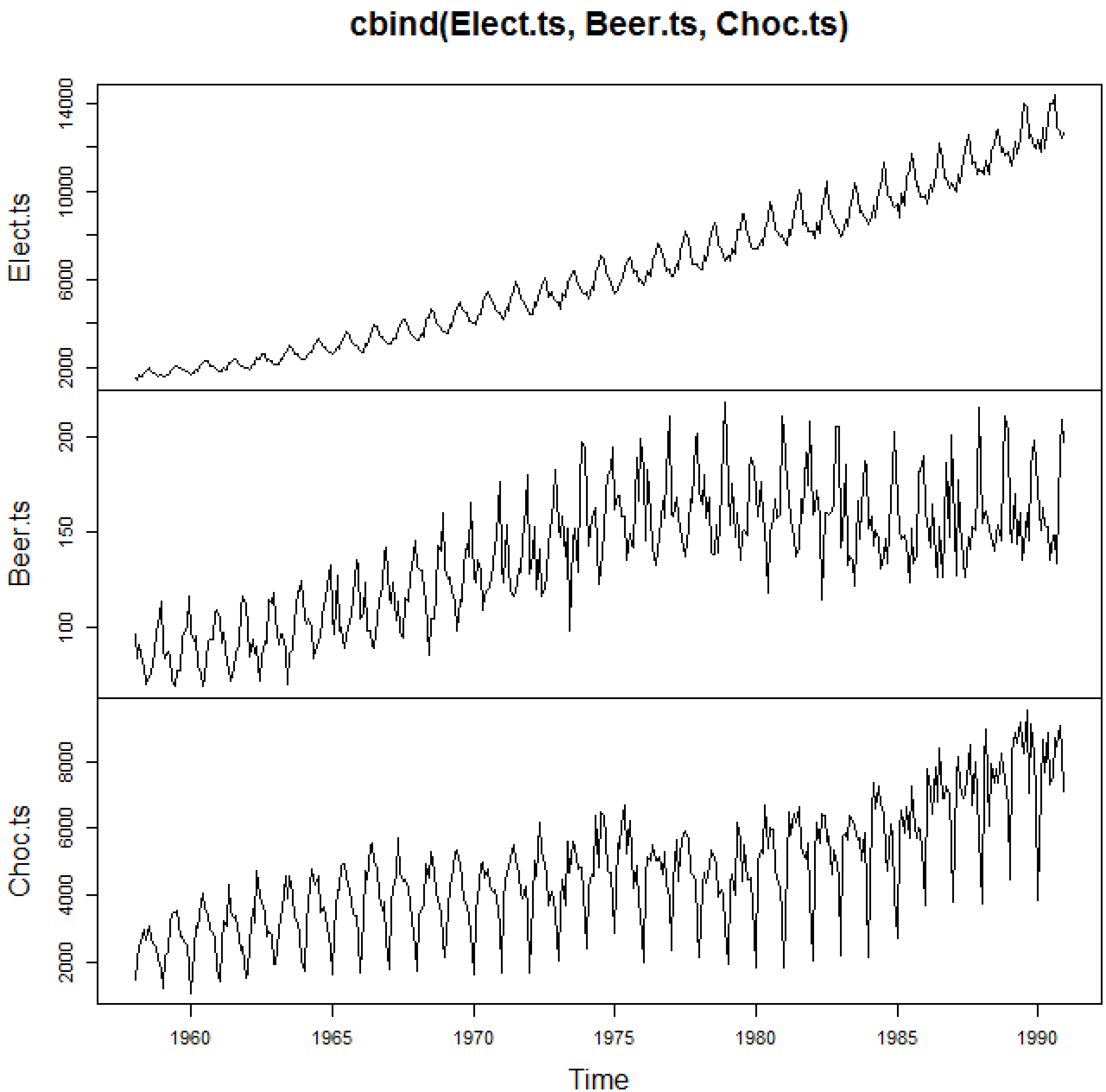
Exemple 4: consommation en Australie de bière, chocolat et électricité

```
www <-
"http://www.mif.vu.lt/~rlapinskas/2012-2013/Erasmus/Data/Cowpertwait,Metcalfe/cbe.dat"

CBE<-read.table(www,header=TRUE)
```

OU, à partir du fichier téléchargé depuis AMeTICE

```
CBE<-read.table("C:/Mon_dossier/cbe.dat",header=TRUE)
```

```
Elect.ts<-ts(CBE[,3],start=1958, freq=12)
Beer.ts<-ts(CBE[,2],start=1958, freq=12)
Choc.ts<-ts(CBE[,1],start=1958, freq=12)

plot(cbind(Elect.ts, Beer.ts, Choc.ts))
```



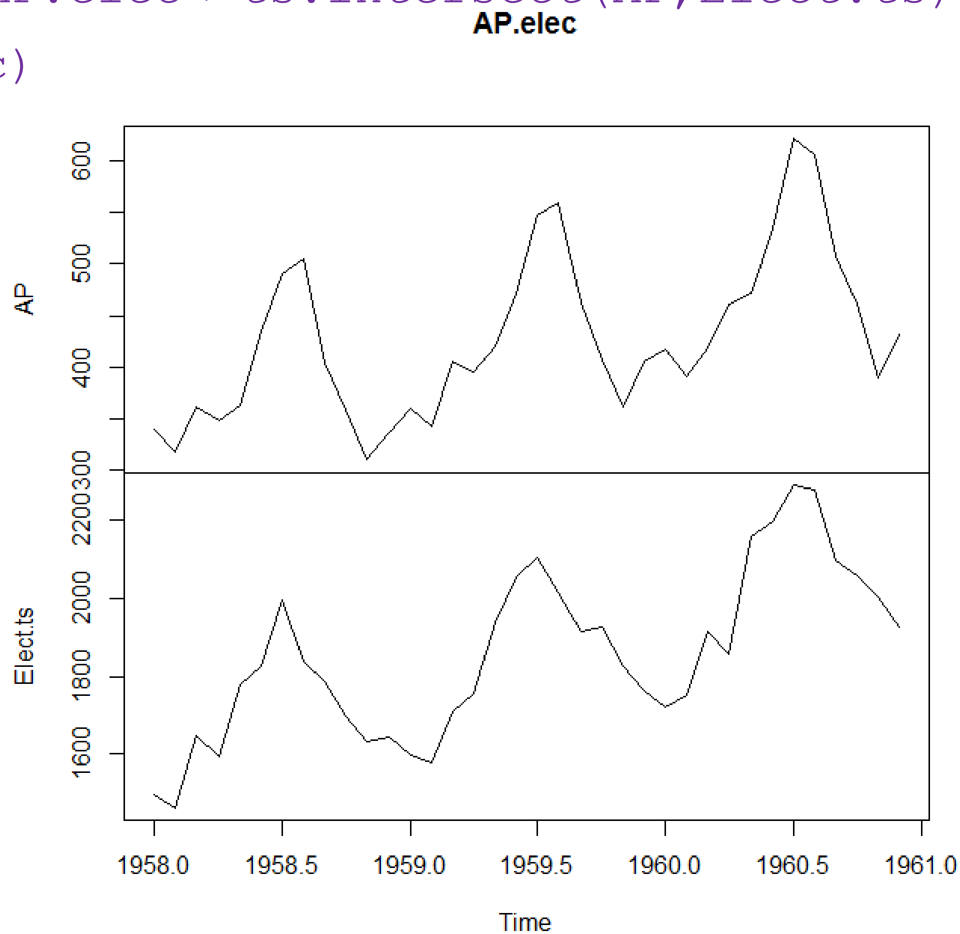
Pendant cette période:
- Pop. 10 -> 18 millions
- Conso. Electrique x7
- Conso. Chocolat x4
- Conso. Bière x1,5

ATTENTION: concomitances fortuites

Consommation Electricité Australie VS passagers PanAM USA

La fonction `ts.intersect` permet d'extraire les périodes communes

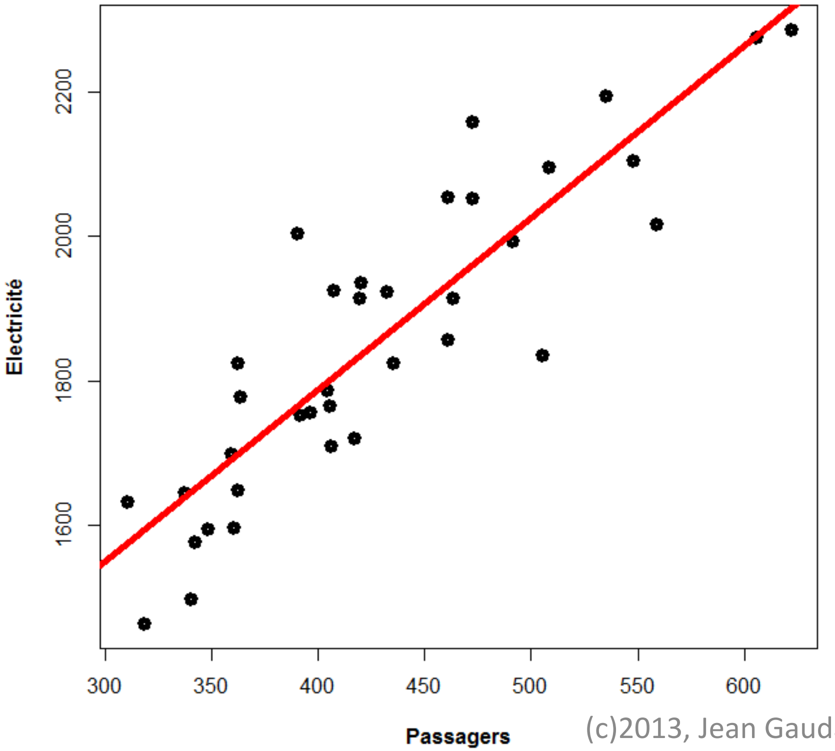
```
AP.elec<-ts.intersect(AP,Elect.ts)
plot(AP.elec)
```



```
plot(as.vector(AP.elec[,1]),as.vector(AP.elec[,2]))
abline(reg=lm(as.vector(AP.elec[,2])~as.vector(AP.elec[,1])))
```

```
cor.test(as.vector(AP.elec[,2]),as.vector(AP.elec[,1]))
```

```
  Pearson's product-moment correlation
 data:  as.vector(AP.elec[, 2]) and as.vector(AP.elec[, 1])
```




```
t=11.0356,df=34, p-value=8.817e-13
alternative hypothesis: true
correlation is not equal to 0
95 percent confidence interval:
 0.7831087 0.9397295
sample estimates:
      cor
0.8841668
```

?Expliquer comment les passagers de la PanAm (USA) influent sur la consommation d'électricité en Australie ?

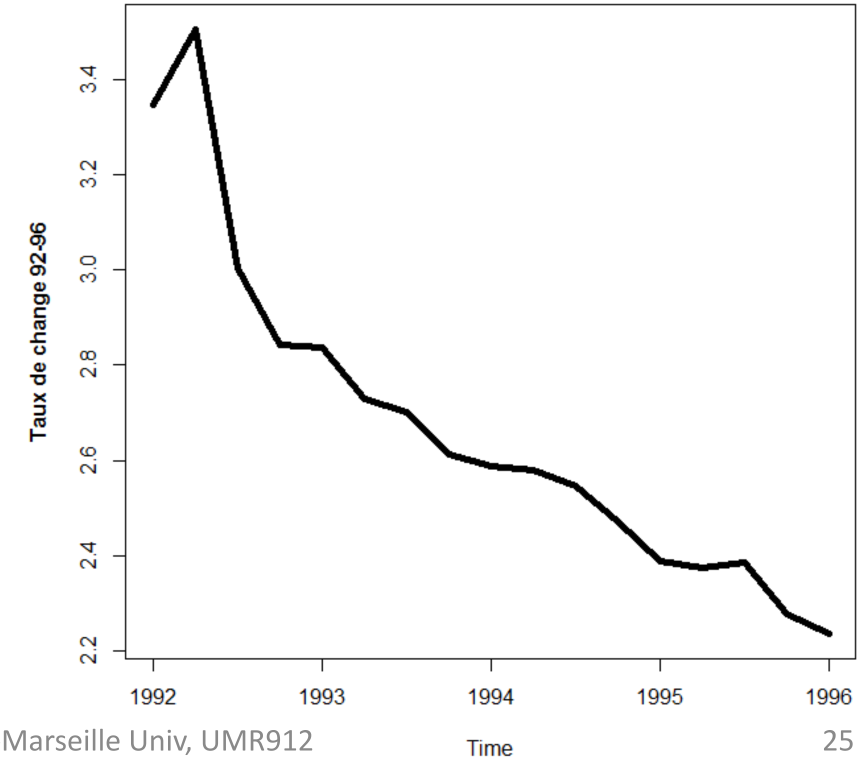
Exemple 5: Taux de change trimestriel de la livre sterling en dollar NZ

```
Z<- read.table("C:/Mon_dossier/pounds_nz.dat",header=TRUE)
Z.ts<-ts(Z, start=1991,freq=4)
```

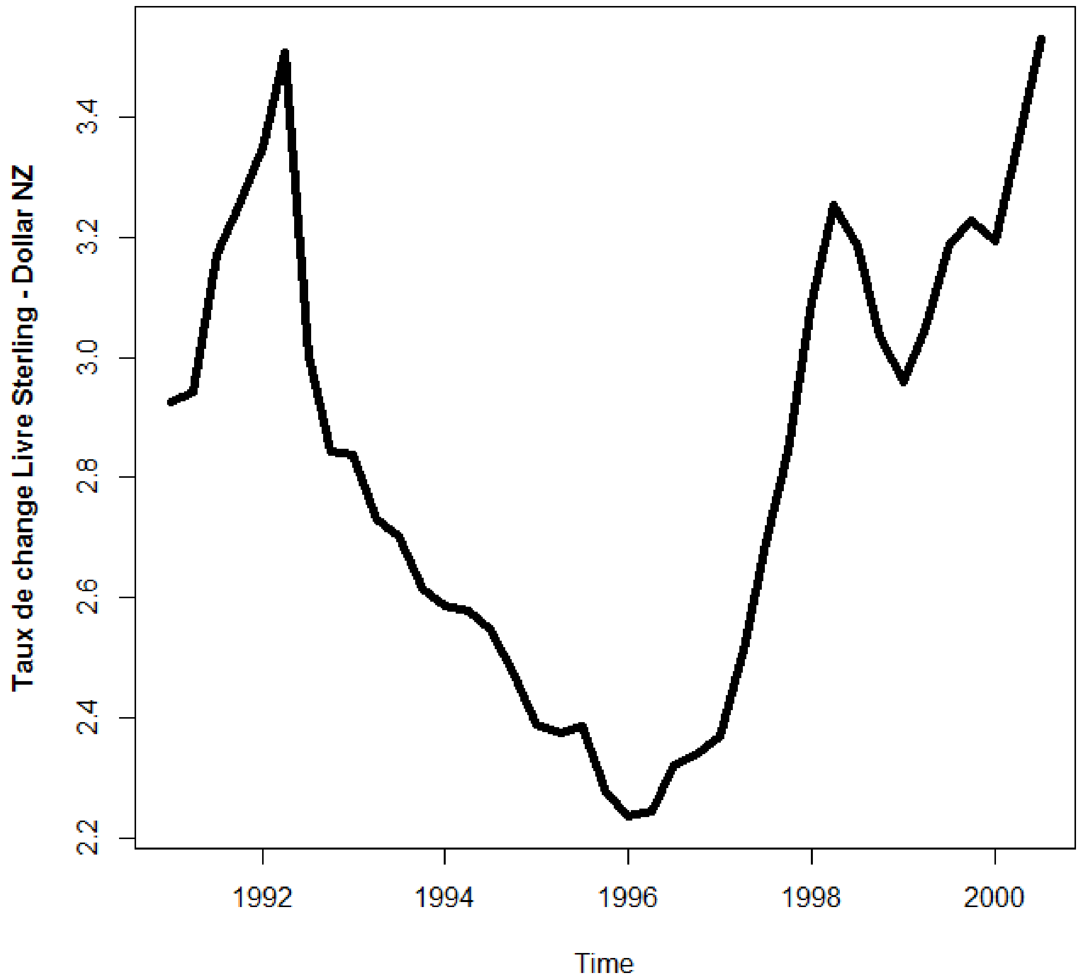
Série de 01/1992 à 01/1996

```
Z.92.96<-window(Z.ts, start=c(1992,1),end=c(1996,1))
plot(Z.92.96)
```

?Que peut-on prévoir pour la suite?



```
Z.96.98<-window(Z.ts, start=c(1996,1),end=c(1998,1))
plot(Z.ts, ylab="Taux de change Livre Sterling - Dollar NZ")
```



Réaliser des prévisions = extrapoler, => 2 hypothèses:

1/ le modèle est adéquat (+observations non erronées et non partielles, compréhension du processus et modélisation correcte)

2/ Poursuite du processus à l'identique.

NB: modèle polynomial d'ordre élevé => très adéquat pour les observations, mais interdit la prévision

III. Décomposition

Décomposition additive

tendance saison résidu

$$y_t = m_t + s_t + \varepsilon_t$$

Décomposition multiplicative

$$y_t = m_t \times s_t + \varepsilon_t$$

Estimation de la tendance $\hat{m}_t$: par exemple moyenne mobile centrée sur y_t

Estimation de la saison: $\hat{s}_t = y_t - \hat{m}_t$

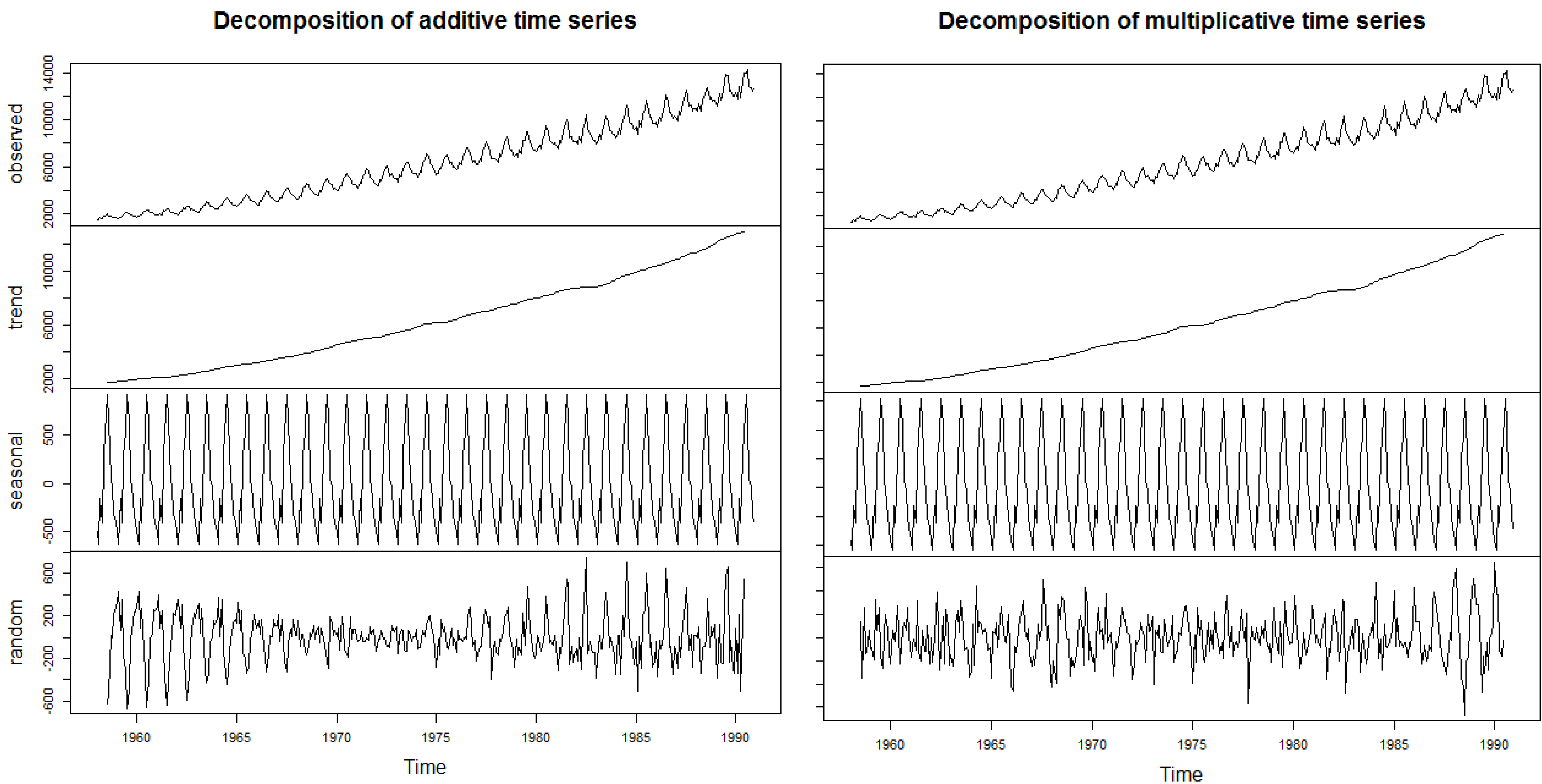
ou $\hat{s}_t = y_t / \hat{m}_t$

La moyenne mobile centrée est un exemple de fonction de lissage (ou filtre),
$\Rightarrow$Série temporelle historique => identification d'une tendance
(fonction `decompose`)

NB: Autre moyen de lissage:
- Fonction linéaire ou polynomiale du temps $m_t = \alpha + \beta t$ ou $m_t = \alpha + \beta t + \delta t^2$
- algorithme *loess* (fonction `stl`) => régression locale pondérée

```
Elec.decom.add<-decompose(Elect.ts,type="add")

Elec.decom.mult<-decompose(Elect.ts,type="mult")
```

```
plot(Elec.decom.add)
plot(Elec.decom.mult)
```



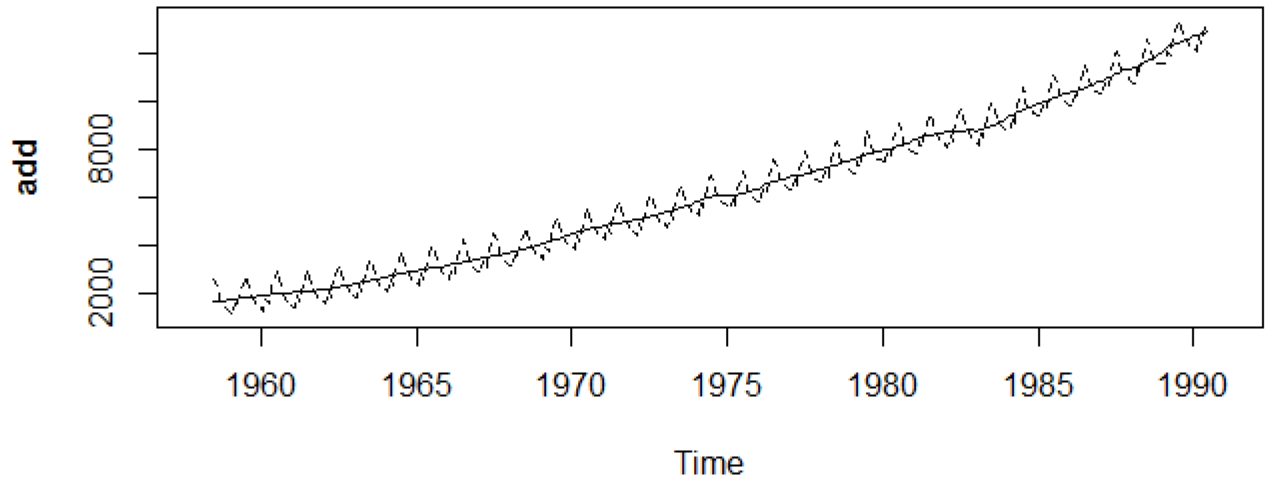
?que remarquez vous?

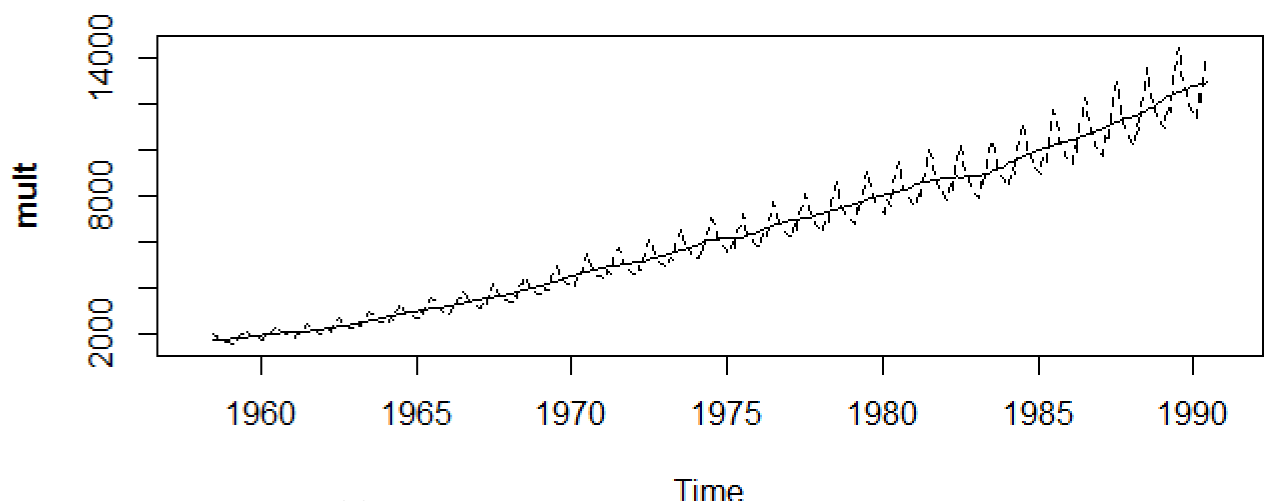
Extraction des tendances et des saison des modèles additif et multiplicatif

```
Trend.add<-Elec.decom.add$trend
Sais.add<-Elec.decom.add$seasonal
Trend.mult<-Elec.decom.mult$trend
Sais.mult<-Elec.decom.mult$seasonal
```

```
layout(1:2)
ts.plot(cbind(Trend.add,Trend.add+Sais.add),lty=1:2)
ts.plot(cbind(Trend.mult,Trend.mult*Sais.mult),lty=1:2)
```

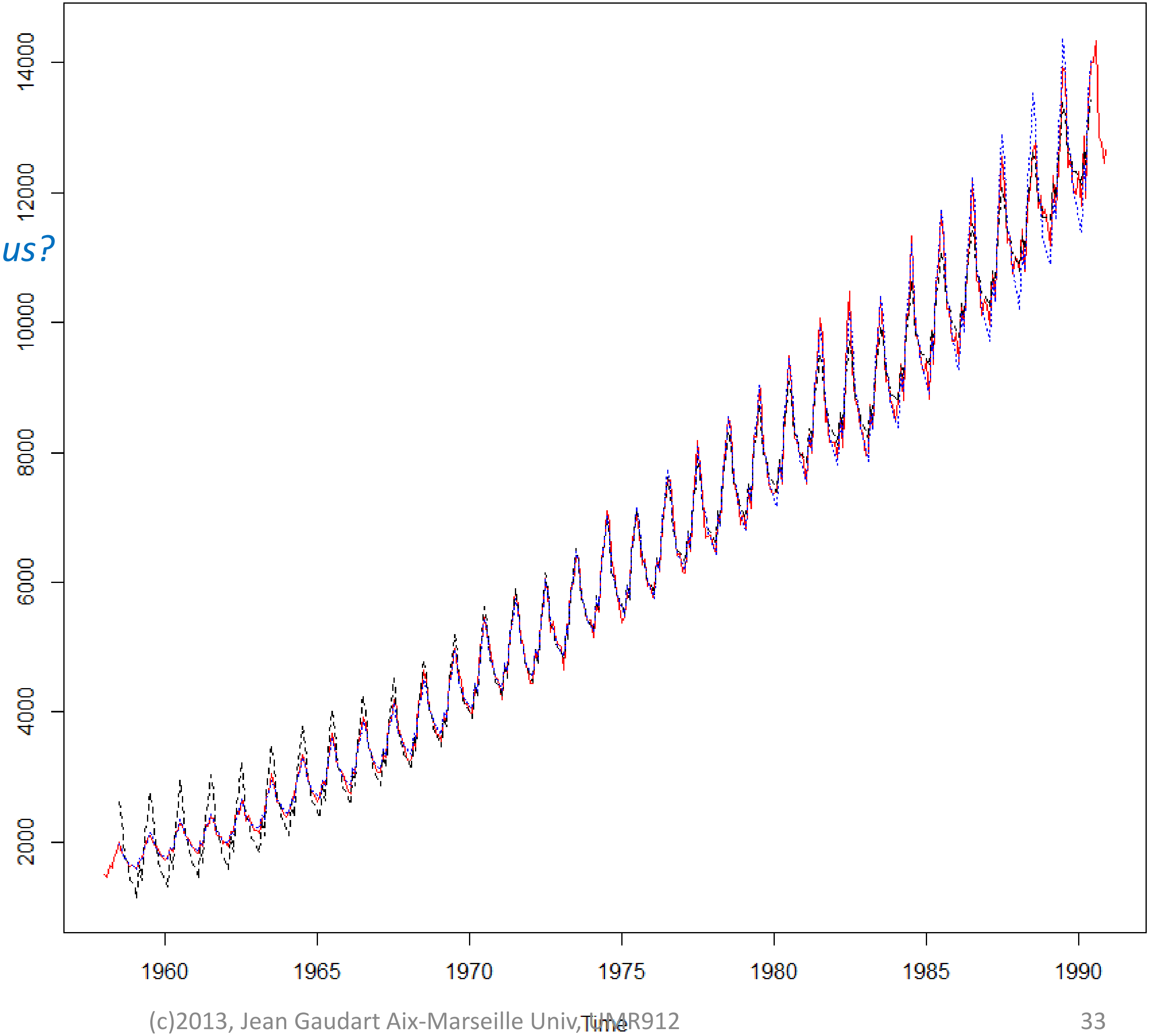
?que remarquez vous?





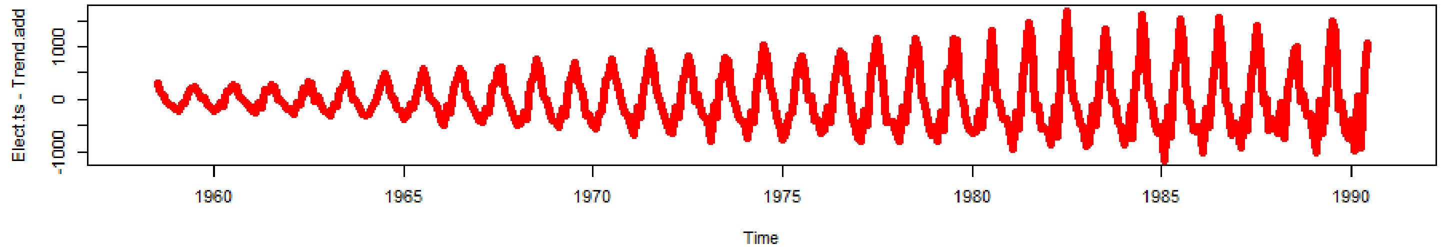
```
ts.plot(cbind(Elect.ts,Trend.add+Sais.add,Trend.mult*Sais.mult)
,col=c("red","black","blue"),lty=1:3)
```

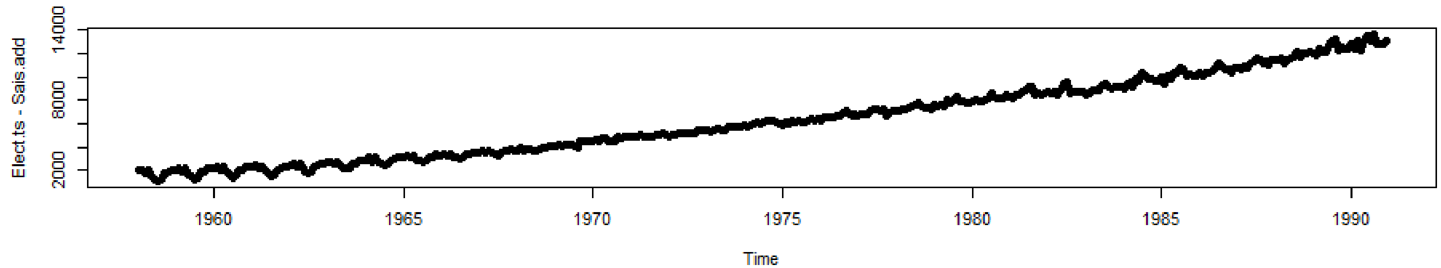
?que remarquez vous?

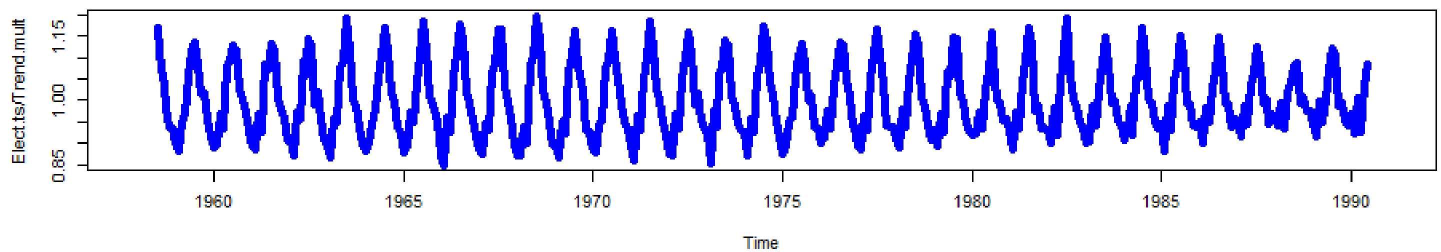


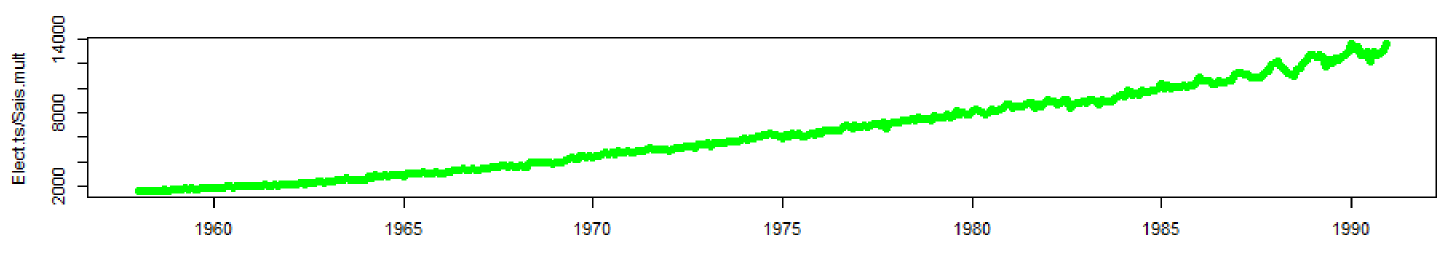
```
layout(1:4)
ts.plot(Elect.ts-Trend.add,col="red")
ts.plot(Elect.ts-Sais.add,col="black")
ts.plot(Elect.ts/Trend.mult,col="blue")
ts.plot(Elect.ts/Sais.mult,col="green")
```

?que remarquez vous?









IV. Corrélation

IV.1 Rappels:

Espérance d'une VA X $E(X) = \mu$ et variance $E\left[(X - E(X))^2\right] = \sigma^2$

Co-variance de 2 VA X et Y (mesure d'association linéaire)

$$\gamma(X, Y) = E\left[(x - \mu_x)(y - \mu_y)\right]$$

Corrélation de 2 VA X et Y

$$\rho(X, Y) = \frac{E\left[(x - \mu_x)(y - \mu_y)\right]}{\sigma_x \sigma_y} = \frac{\gamma(X, Y)}{\sigma_x \sigma_y}$$

IV.2 Ensembles, stationnarité, ergodicité

Espérance d'une série temporelle $E(Y_t) = \mu(t)$

$\Rightarrow$Moyenne, fonction d'un temps, d'un ensemble de séries temporelles (population)
$\Rightarrow$une seule réalisation observée
$\Rightarrow$On estime l'ensemble en faisant l'hypothèse d'une structure particulière (tendance, saison)

l'espérance peut-être indépendante du temps, estimée par

$$\hat{\mu} = \sum_{t=1}^{n} \frac{y_t}{n}$$

C'est pourquoi on décompose une série

Série suffisamment longue pour « capturer » la structure sous-jacente:

$$\lim_{n\to\infty} \sum_{t=1}^{n} \frac{y_t}{n} = \mu$$ Série ergodique en moyenne

Variance d’une série temporelle $E\left[\left(y_t - \mu\right)^2\right] = \sigma^2(t)$

La variance peut être indépendante du temps estimée par

$$\hat{\sigma}^2 = \frac{\sum\left(y_t - \hat{\mu}\right)^2}{n-1}$$

Si les observations sont corrélées => les observations proches sont similaires
=> sous-estimation de la variance sur une courte série

Auto-covariance et Auto-corrélation

Pour un « lag » k

$$\gamma_k = E\left[(y_t - \mu)(y_{t+k} - \mu)\right]$$

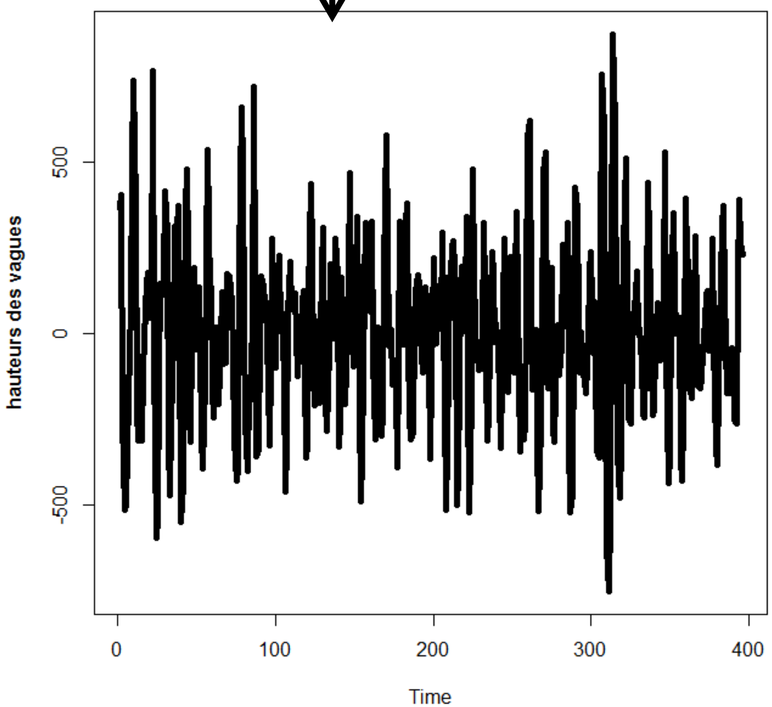
$$\rho_k = \frac{\gamma_k}{\sigma^2}$$

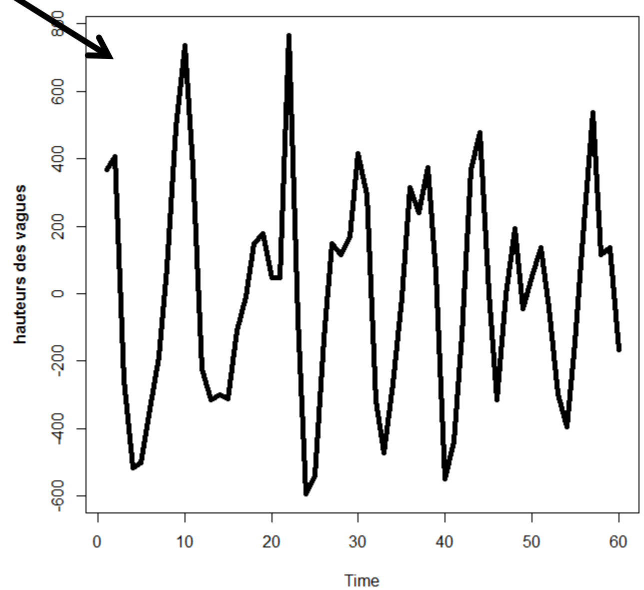
Une série temporelle est dite stationnaire si

$$\begin{cases} \hat{\mu} = \sum_{t=1}^{n} \frac{y_t}{n} & \forall t \\ \operatorname{cov}(Y_t, Y_{t+k}) = \gamma_k & \forall t \end{cases}$$

Exemple 6: hauteur de l’eau au centre d’un réservoir (vagues)

```
wave<- read.table("C:/mon_dossier/wave.dat",header=TRUE)
attach(wave)
plot(ts(waveht)); plot(ts(waveht[1:60]))
```






```
decompose(waveht, type="add")
```

? Que remarquez vous?

autocorrelation d'ordre 1: ie entre yt et yt+1

```
acf(waveht)$acf[2]            [1] 0.4702564
```

autocovariance d’ordre 1

```
acf(waveht, type=c("covariance"))$acf[2]    [1] 33328.39
```

?Que donne le script suivant? Pourquoi?

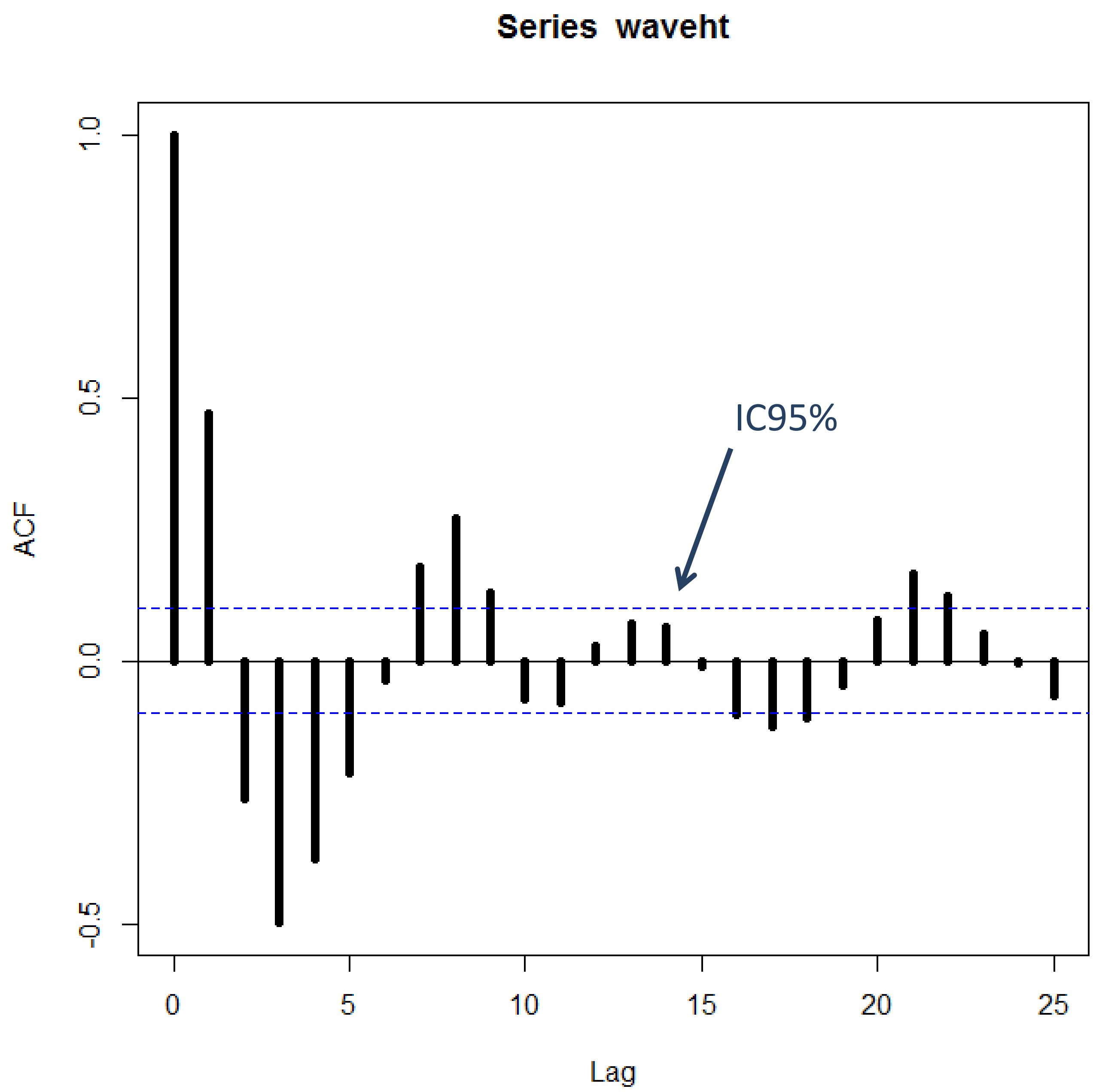
```
acf(waveht)$acf[1]
```

Remarque: si $\rho_k = 0$ alors $r_k \sim> N\left(-1/n;1/n\right)$

=> IC95% de r_k $-\frac{1}{n} \pm 1{,}96 \frac{1}{\sqrt{n}}$

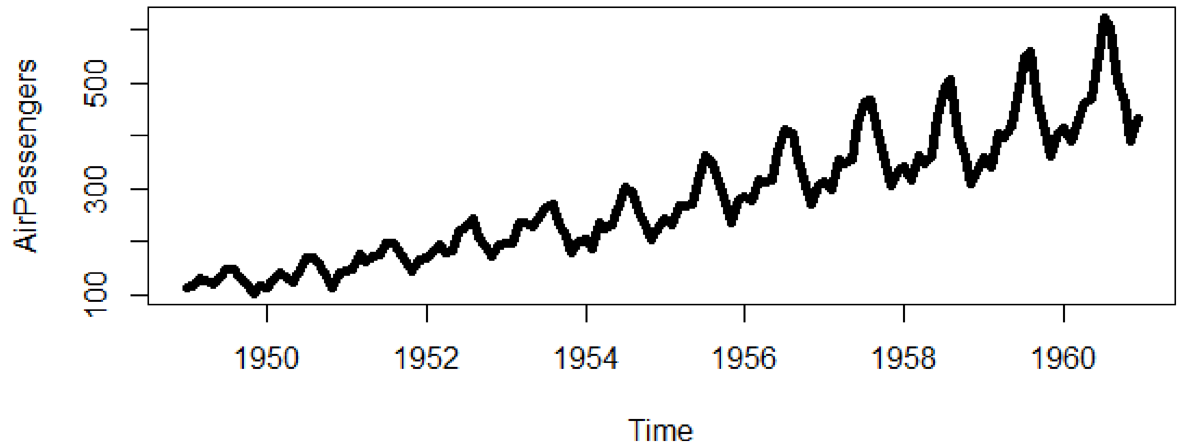
Corrélogramme

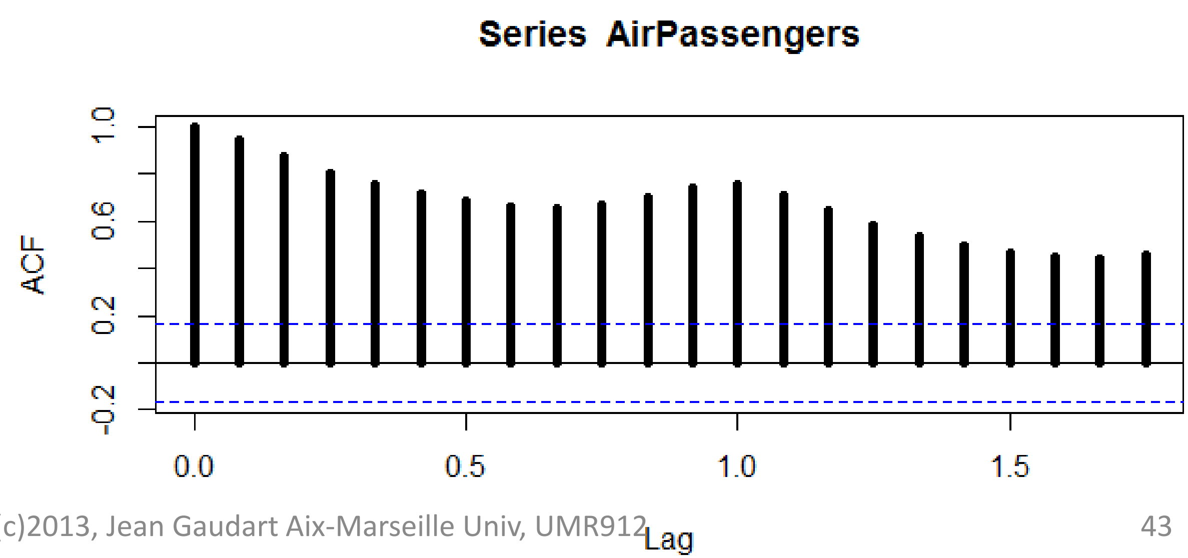
```
acf(waveht)
```



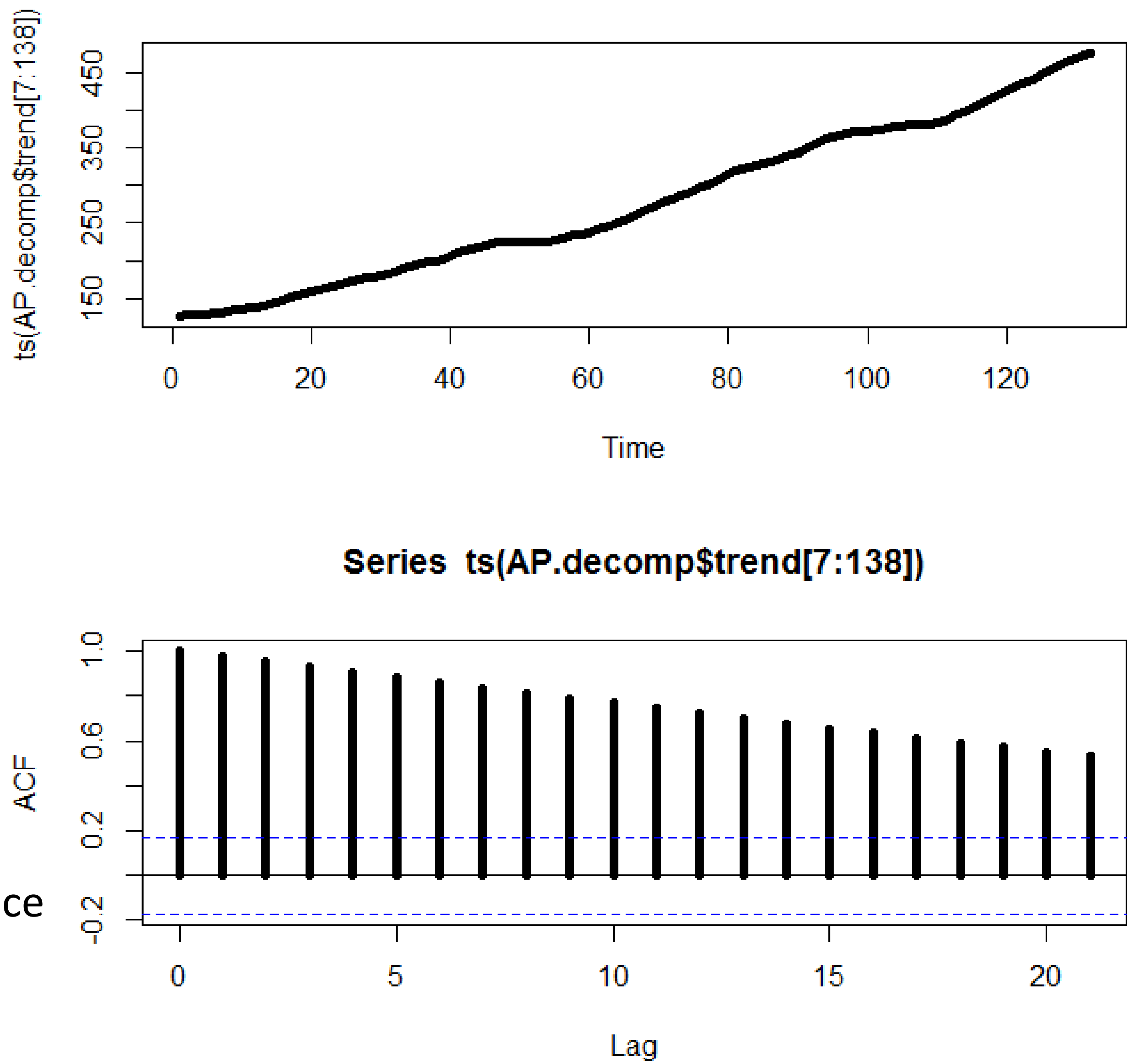
Reprise exemple 1: PanAm

```
layout(1:2)
plot(AirPassengers)
acf(AirPassengers)
AP.decomp<-decompose(AirPassengers, type="mult")
```



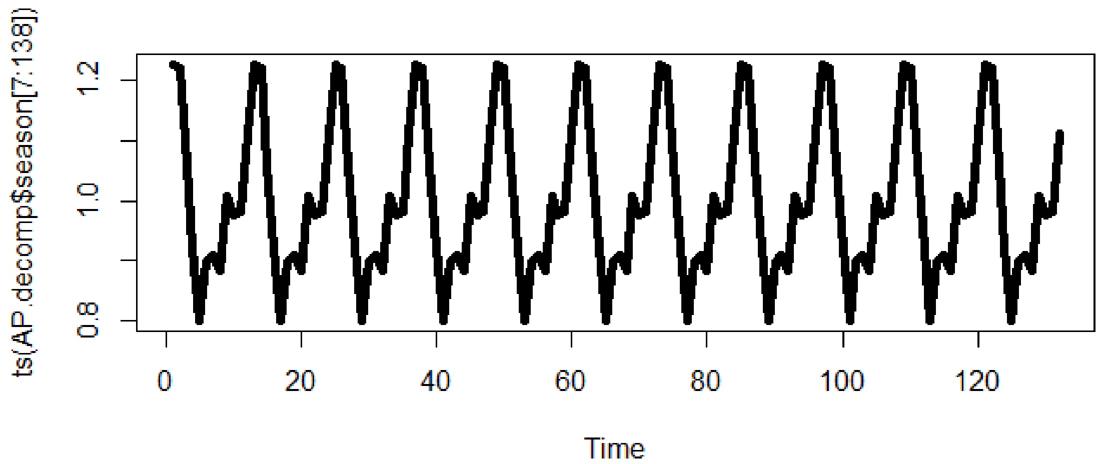


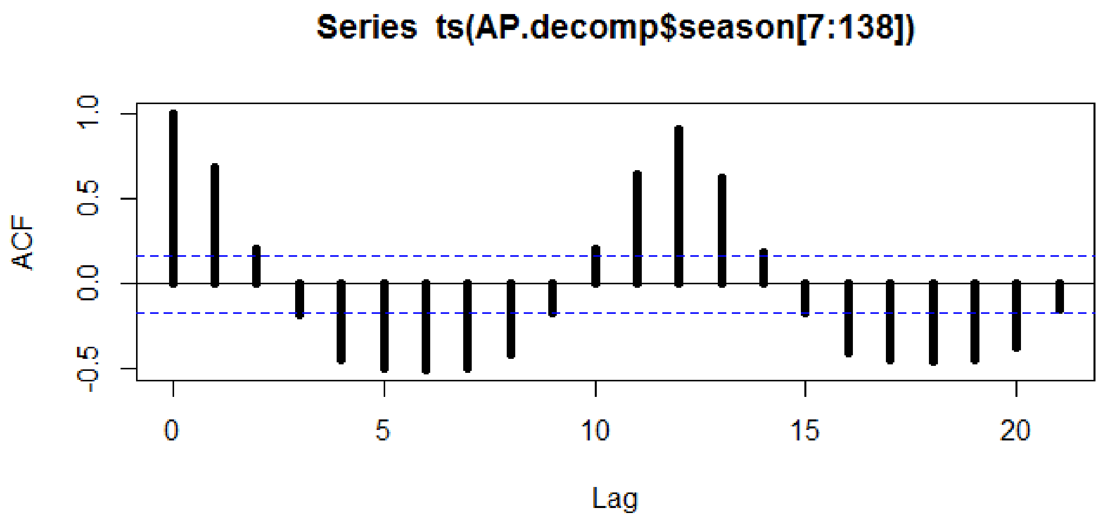
```
layout(1:2)
plot(ts(AP.decomp$trend[7:138]))
acf(ts(AP.decomp$trend[7:138]))
```



Aspect typique de tendance

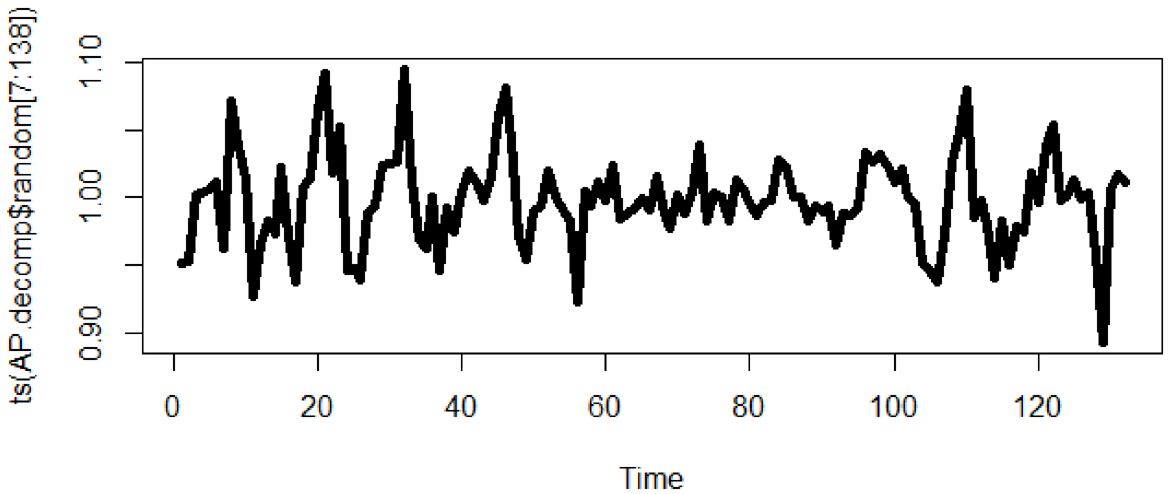
```
layout(1:2)
plot(ts(AP.decomp$season[7:138]))
acf(ts(AP.decomp$season[7:138]))
```



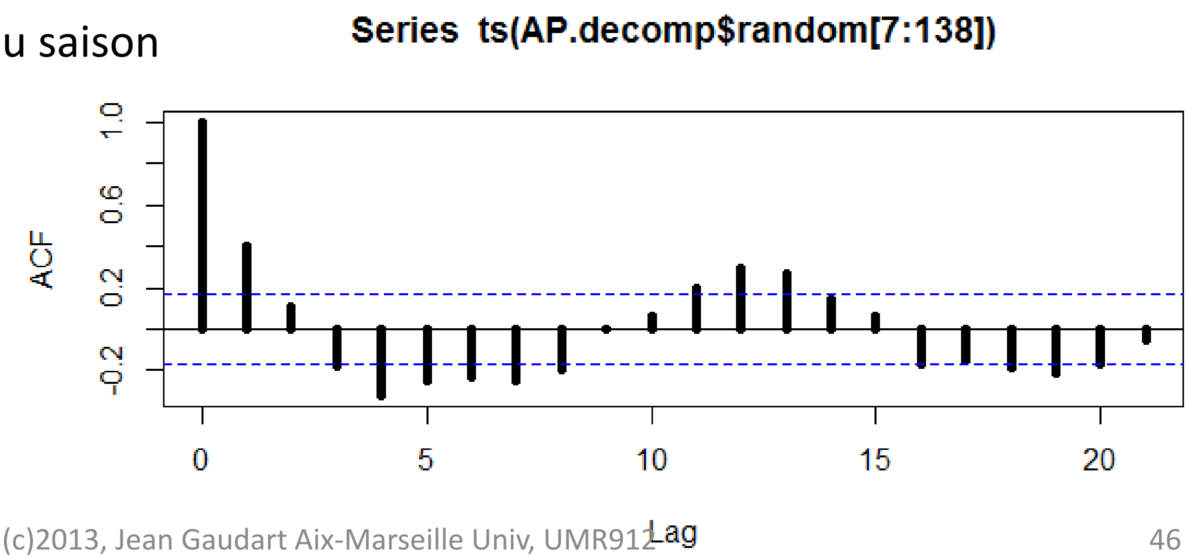


Aspect typique saison

```
layout(1:2)
plot(ts(AP.decomp$random[7:138]))
acf(ts(AP.decomp$random[7:138]))
```



sinusoïde amortie =>AR(2) ou saison



?La dé-saisonnalité est-elle suffisante?
?Comment faire?

?La dé-saisonnalité est-elle suffisante?
?Comment faire?

```
sd(AP[7:138])                        [1] 109.4187

sd(AP.decomp$random[7:138])          [1] 0.0333884
```

?Qu'en pensez vous?

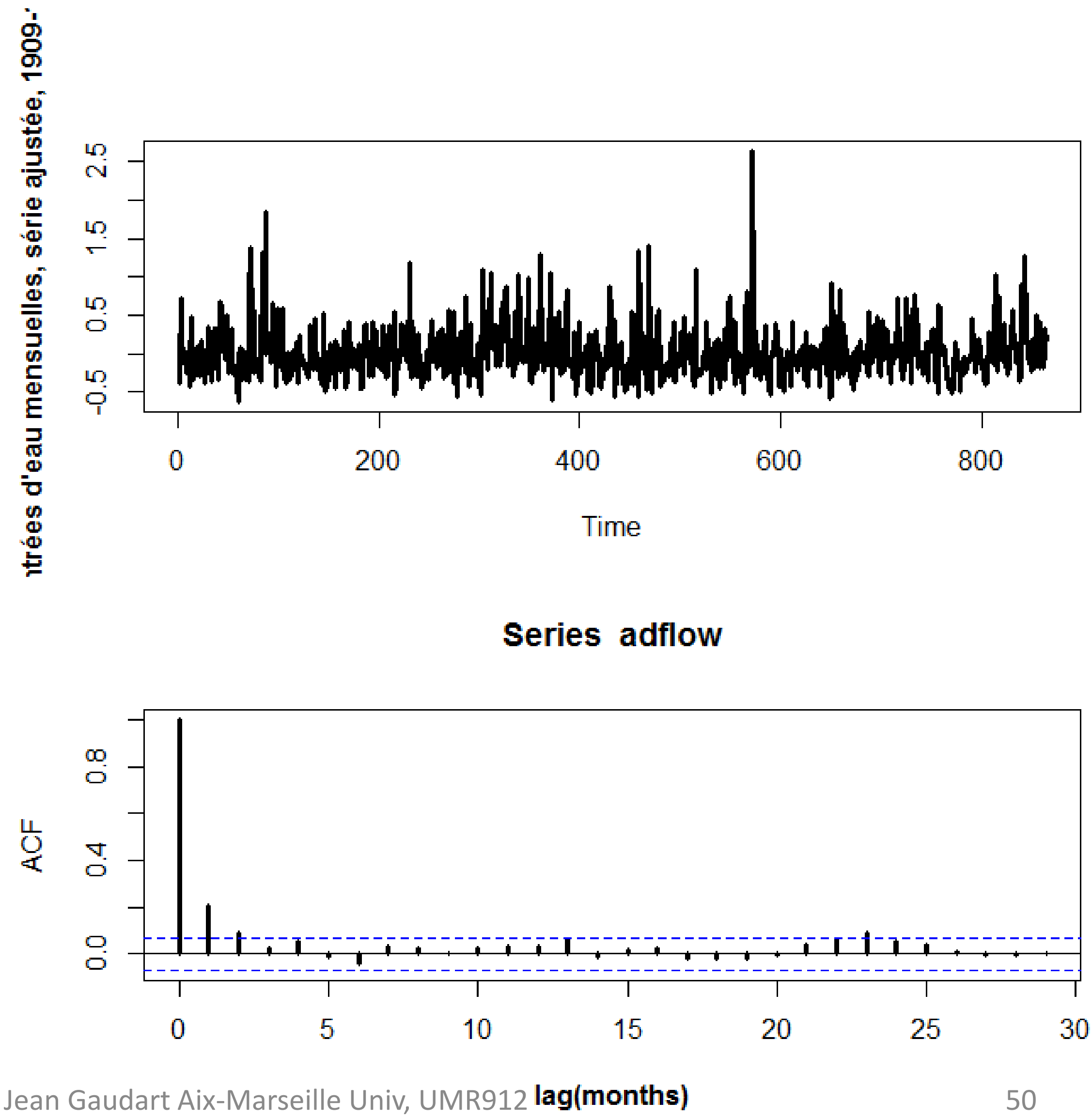
Exemple 7: entrées dans un réservoir d'eau, Ecosse

```
Res<- read.table("C:/mon_dossier/Fontdsdt.dat",header=TRUE)
attach(Res)
summary(Res)
```

```
    adflow
 Min.   :-0.6213
 1st Qu.:-0.2151
 Median :-0.0744
 Mean   : 0.0000
 3rd Qu.: 0.1427
 Max.   : 2.6388
```

```
layout(1:2)
plot(ts(adflow), ylab="entrées/mois, ajustées, 1909-1980")
acf(adflow, xlab="lag(mois)")
```

?interprétez?



décroissance exponentielle
=> AR(1)

V. Co-facteur

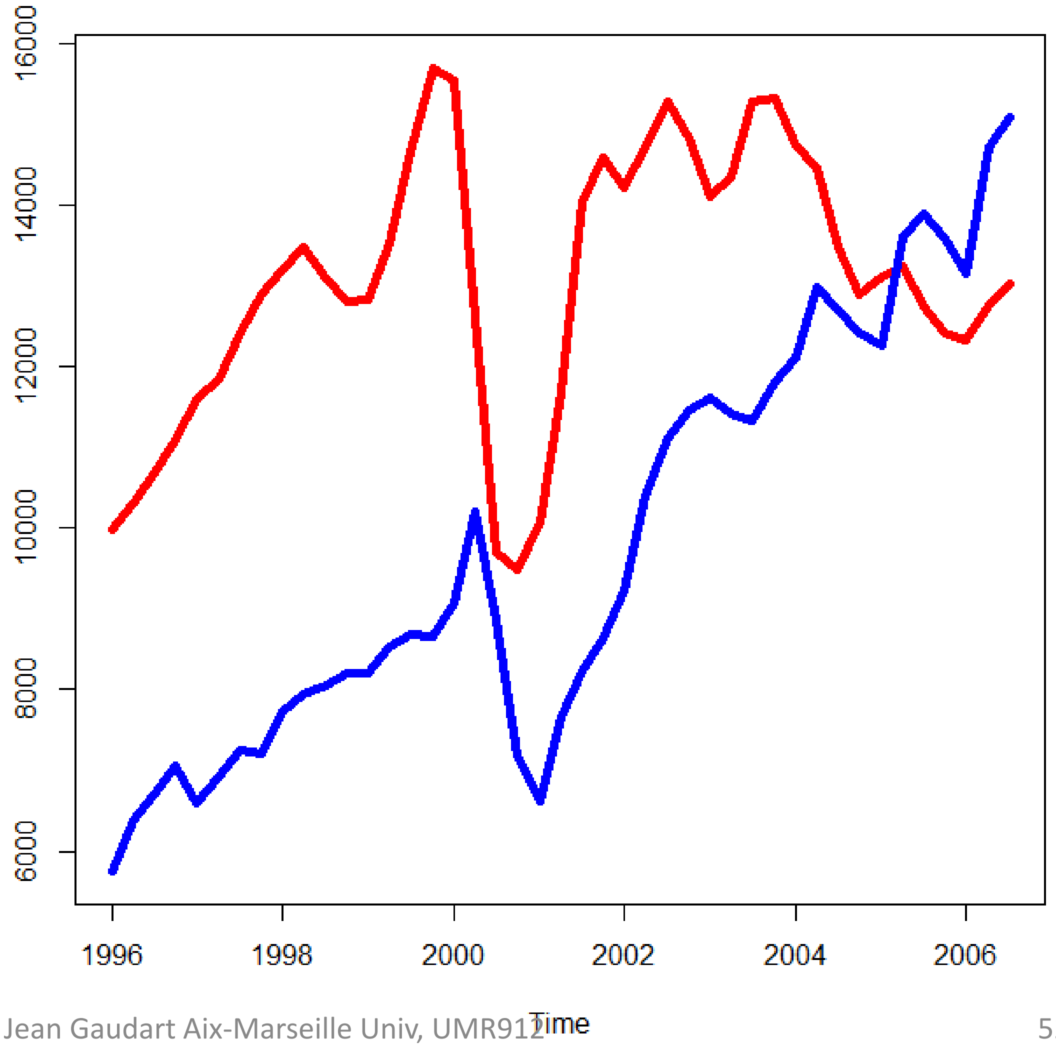
Chercher si un co-facteur qui « conduit » la série

Exemple 8: Permis de construire et constructions, Australie

```
Construct<-
read.table("C:/mon_dossier/ApprovActiv.dat",header=TRUE)
attach(Construct)
Perm.ts<-ts(Approvals, start=c(1996,1), freq=4)
Act.ts<-ts(Activity, start=c(1996,1), freq=4)


ts.plot(Perm.ts, Act.ts, col=c("red","blue"))
```

```
ts.plot(Perm.ts, Act.ts, col=c("red","blue"))
```

?Que remarquez vous?



Corrélation croisée: cross-corrélation et cross-covariance

Assumons 2 séries temporelles X_t et Y_t stationnaires

Cross-covariance:

$$\gamma_k(x,y) = E\left[\left(x_{t+k} - \mu_x\right)\left(y_t - \mu_y\right)\right]$$

Attention: la relation n'est pas symétrique: décalage entre y et x d'un lag k
=> x est en avance sur y d'un délai k (x est VA explicative, y VA réponse)

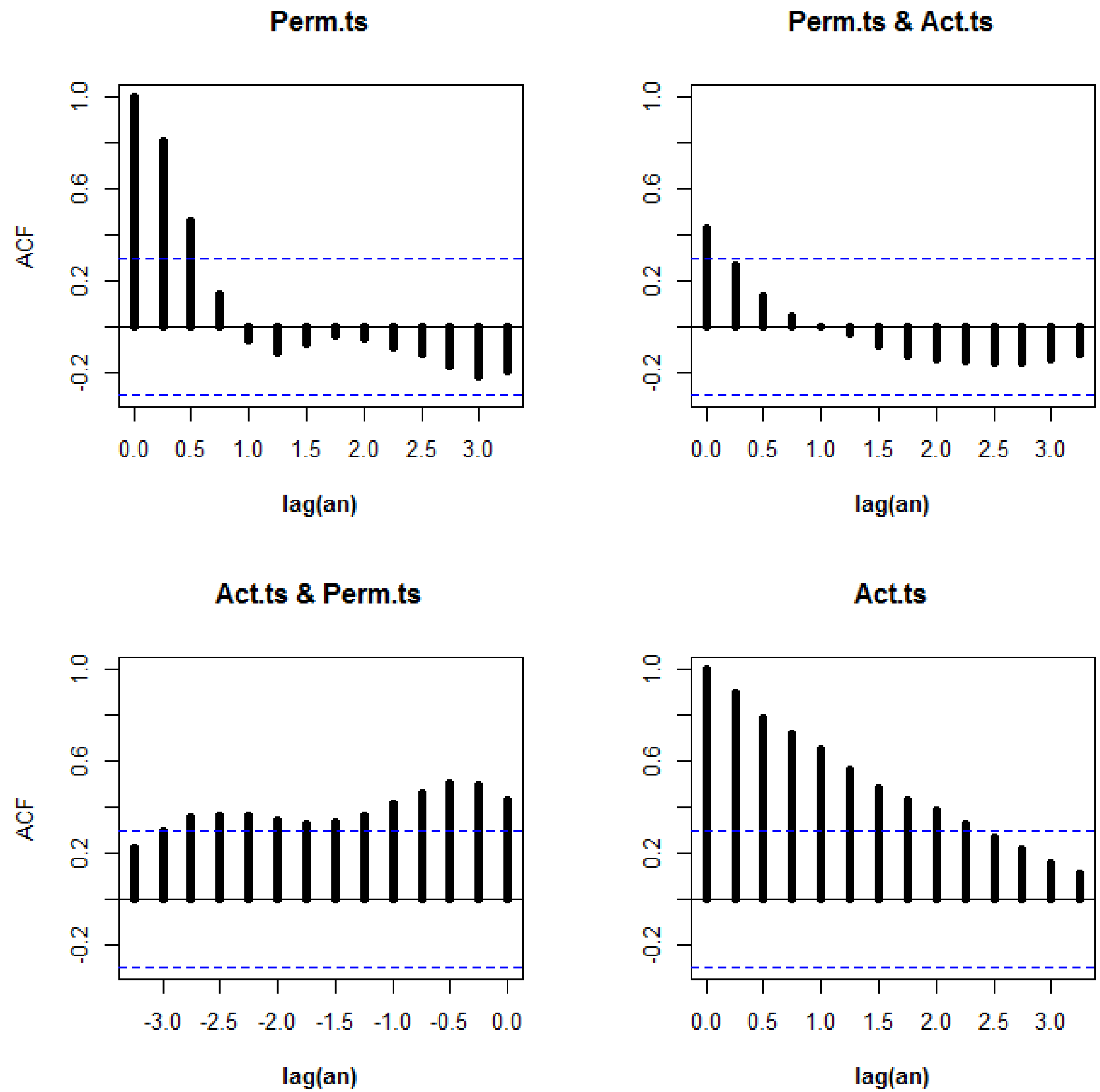
Cross-corrélation:

$$\rho_k(x,y) = \frac{\gamma_k(x,y)}{\sigma_x \sigma_y}$$

Combiner 2 séries temporelle de fréquence similaire: fonction `ts.union()`

```
acf(ts.union(Perm.ts,Act.ts))
```

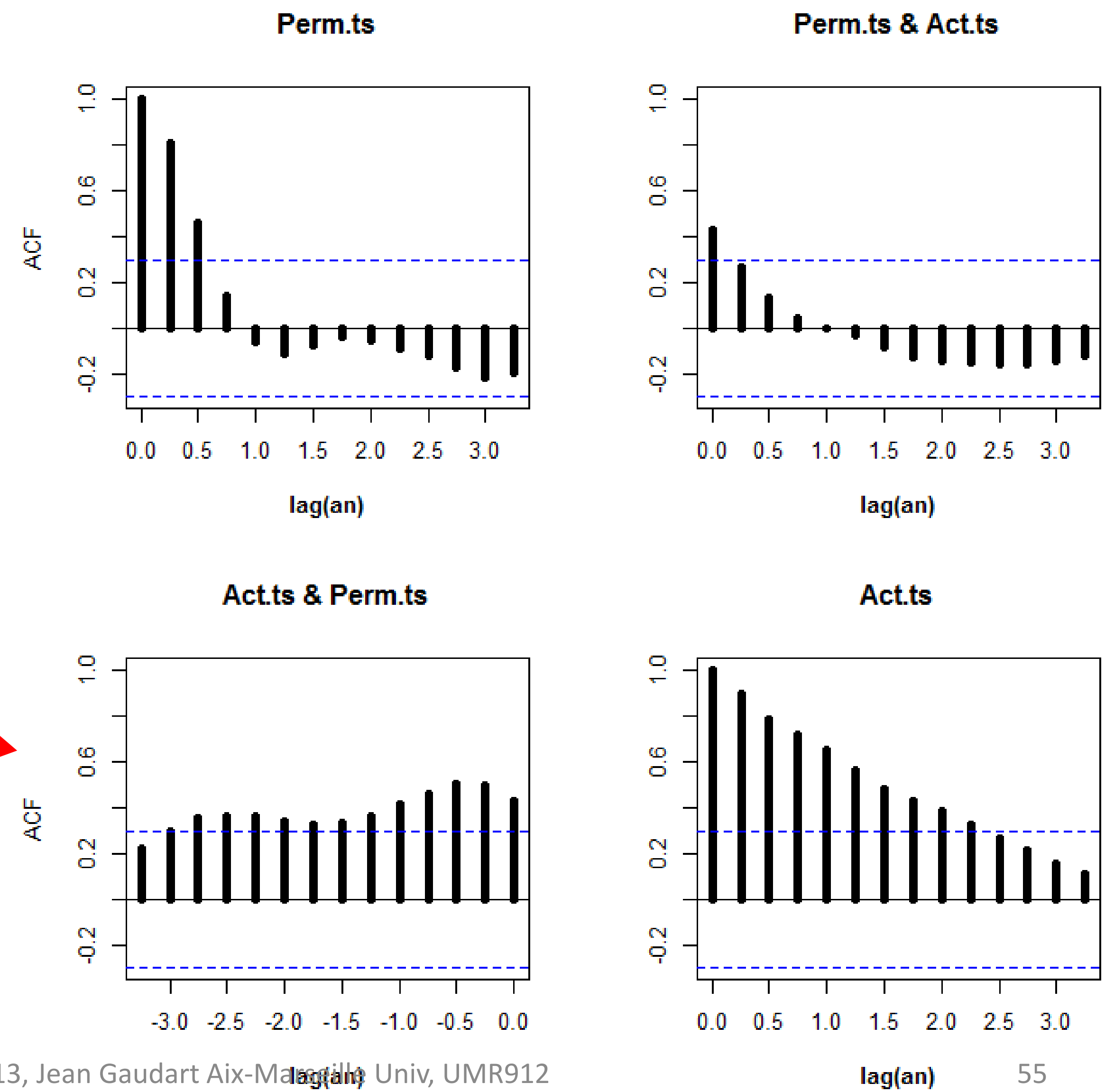
?Que remarquez vous?



Combiner 2 séries temporelle de fréquence similaire: fonction `ts.union()`

```
acf(ts.union(Perm.ts,Act.ts))
```

Cross-corrélation négative significative:
Perm(t+k) <-> Act(t)



ATTENTION: si les 2 séries ont une même tendance ou saison, cross-corrélation +++ => stationnariser les séries

?Comment faire pour enlever tendance et saison?

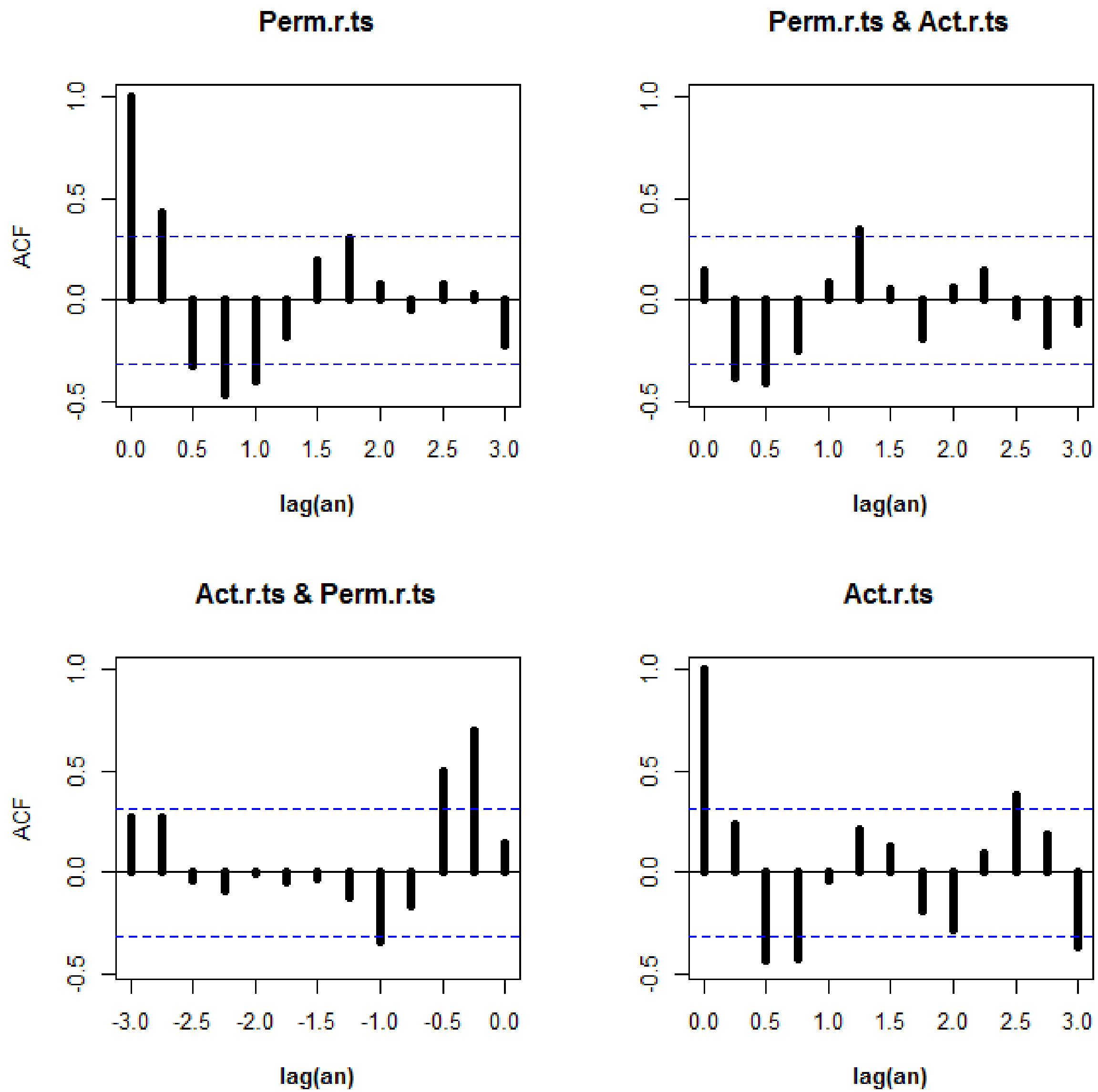
?Comment faire pour enlever tendance et saison?

```
Perm.r<-decompose(Perm.ts)$random
Perm.r.ts<-window(Perm.r,start=c(1996,3),end=c(2006,1))
Act.r<-decompose(Act.ts)$random
Act.r.ts<-window(Act.r,start=c(1996,3),end=c(2006,1))
```

```
acf(ts.union(Perm.r.ts,Act.r.ts))
```

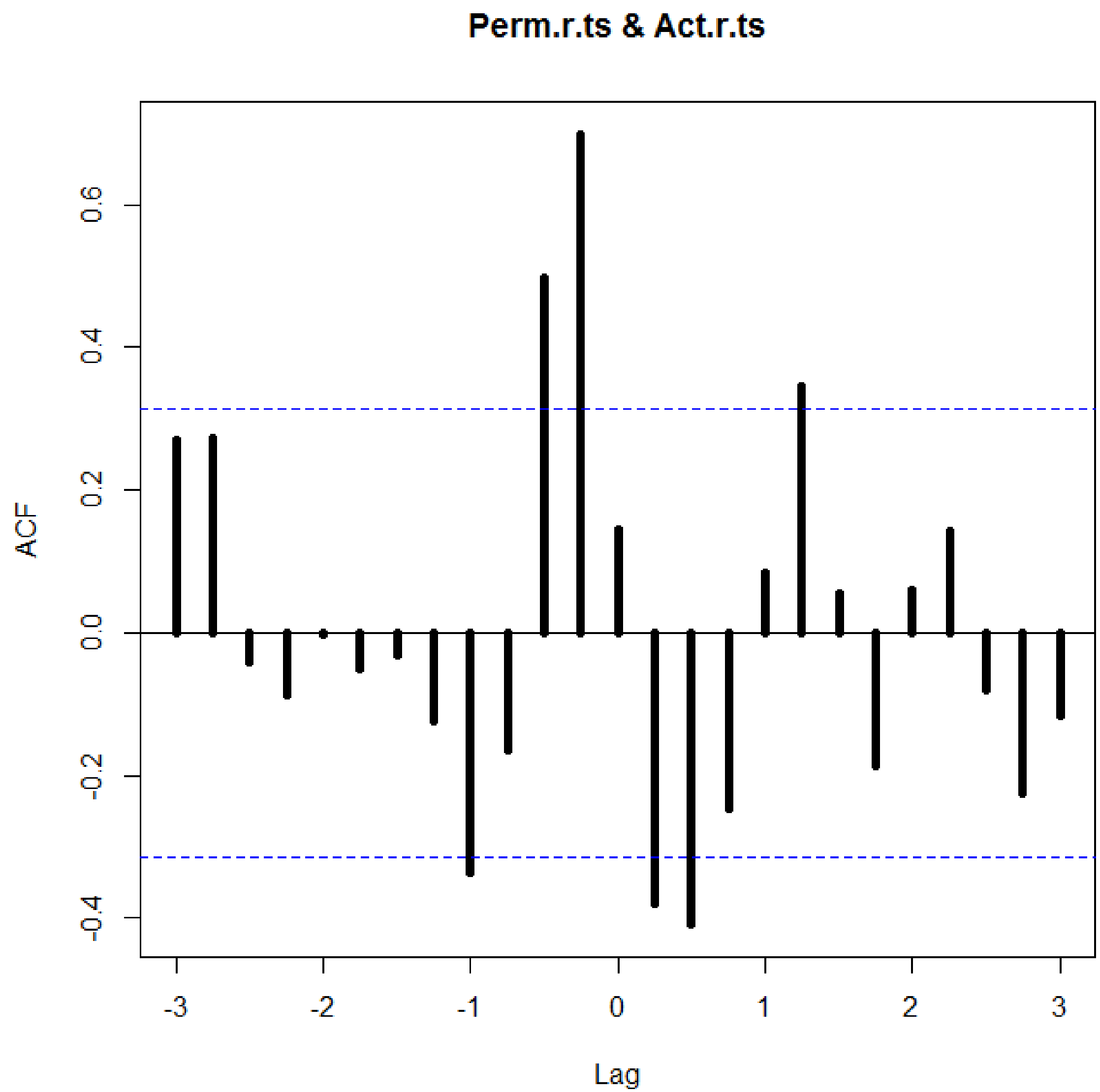
```
acf(ts.union(Perm.r.ts,Act.r.ts))
```

?interpréter?



```
ccf(Perm.r.ts,Act.r.ts)
```

Cross-correlation entre Perm(t+k) et Act(t)



VI. Lissage exponentiel

VI.1 Lissage Exponentiel

Objectif: prévoir le futur y_{n+k} en fonction du passé ($y_1 \ldots y_n$)
Assumer: pas de tendance ni d'effet saison sur le processus (ou ajuster la série)
=> La moyenne du processus change de t = t+1, mais pas d'info.

$$y_t = \mu_t + \varepsilon_t$$

Un bonne estimation de μ_t est donné par

$$\mu_t = \alpha y_t + (1-\alpha)\mu_{t-1} \qquad 0 < \alpha < 1$$

=EWMA: exponentially weighted moving average

Si $\alpha \cong 1$, peu de lissage et $\mu_t = y_t$ utile si le changement attendu est grand / σ
Si $\alpha \cong 0$, lissage ++, $\mu_t = \mu_{t-1}$, aucun impact du temps, seulement σ

En pratique, souvent α=0,2 ; mais peut être estimé

Prévision: $$\mu_t = \alpha y_t + (1-\alpha)\mu_{t-1}$$

$$\Leftrightarrow \mu_t = \alpha y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \ldots$$

=> Le futur est une combinaison linéaire du présent et du passé, avec une pondération décroissante avec le lag.

=>suite géométrique

Erreur du modèle (one-step-ahead prediction error):

$$\varepsilon_t = y_t - \hat{y}_{t/t-1} = y_t - \mu_{t-1}$$

(Cf erreur de prédiction en RL)

Estimation du paramètre α en minimisant la somme des carrés des erreurs (SS1PE sum of squared one-step-ahead predictor errors)

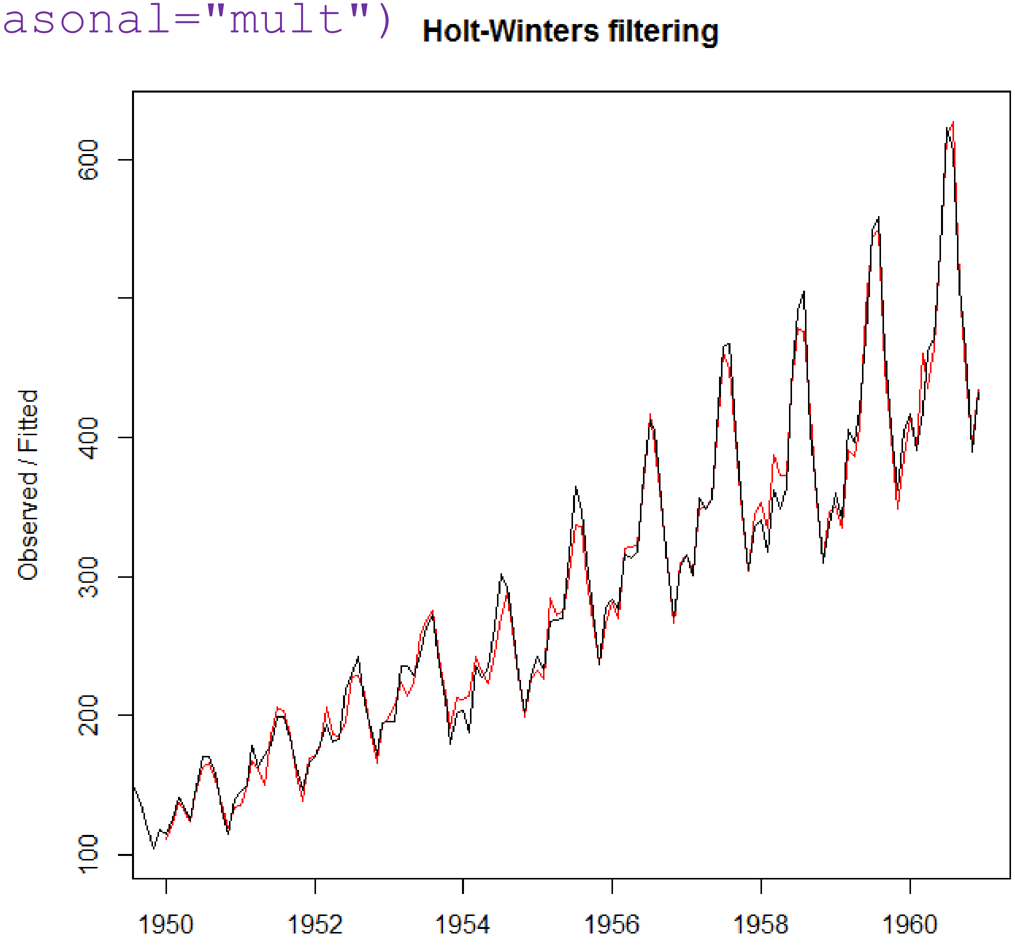
$$\mathrm{SS1PE} = \sum_{t=2}^{n} \varepsilon_t^2$$

VI.2 Algorithme Holt-Winters

Lissage exponentiel : cas particulier de HW.
HW analyse: tendance, Saison, et Changement (level=moyenne ajustée)

Reprise Exemple 1: voyageurs de la PanAm

```
AP.liss<-HoltWinters(AP, seasonal="mult")
plot(AP.liss)
```

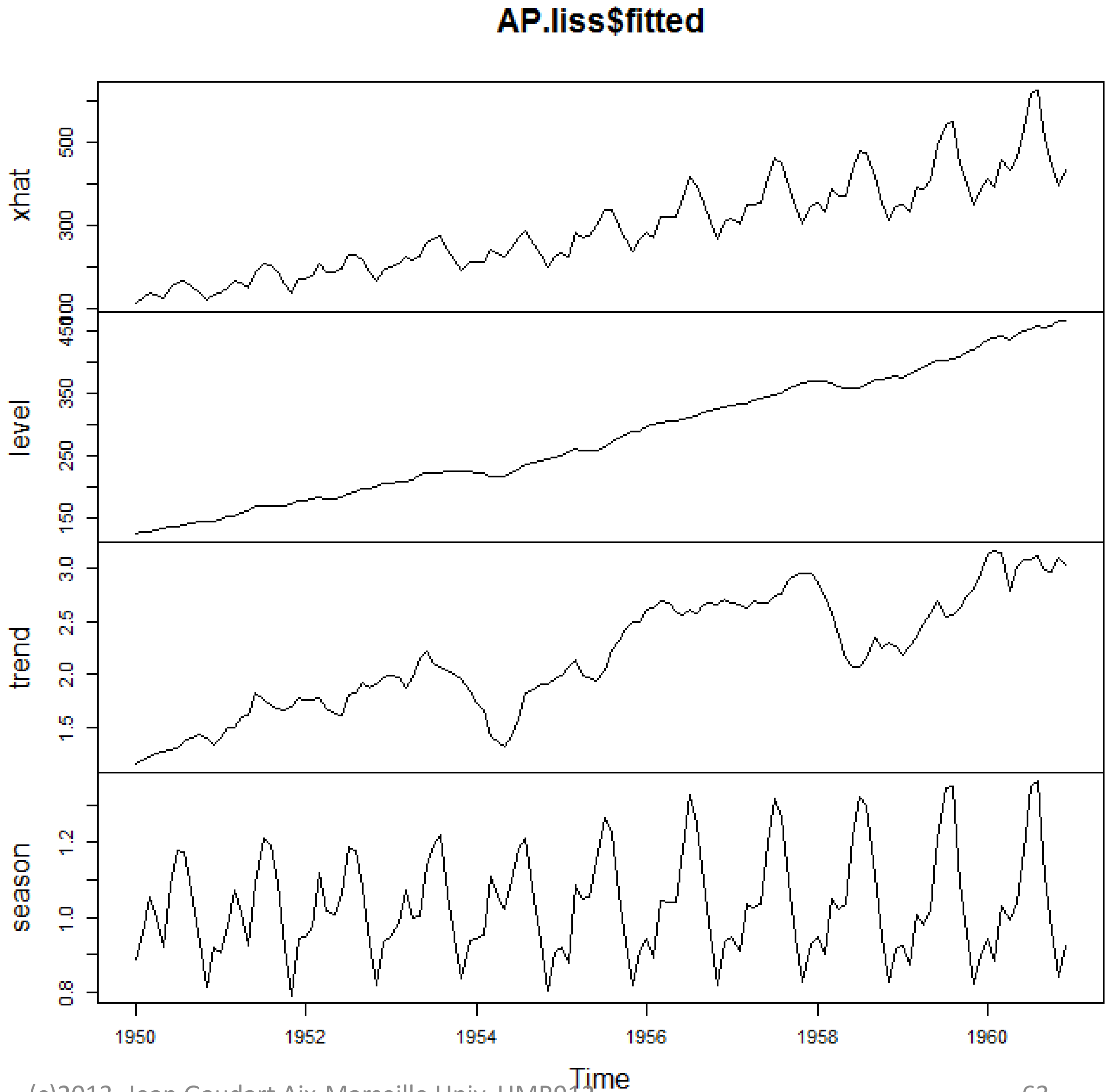



```
sd(AP)
sqrt(AP.liss$SSE/length(AP))
```

```
[1] 119.9663
[1] 10.72729
```

?Que remarquez vous?

```
plot(AP.liss$fitted)
```



Prévision sur 4 ans

```
AP.prev<-predict(AP.liss,n.ahead=4*12)
ts.plot(AP,AP.prev, col=c("black","red"))
```

ATTENTION:
Modèle adéquate?
Processus inchangé?