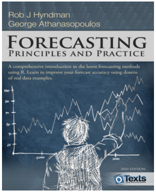
Chapitre 1

Désaisonnalisation et Décomposition des Séries Temporelles

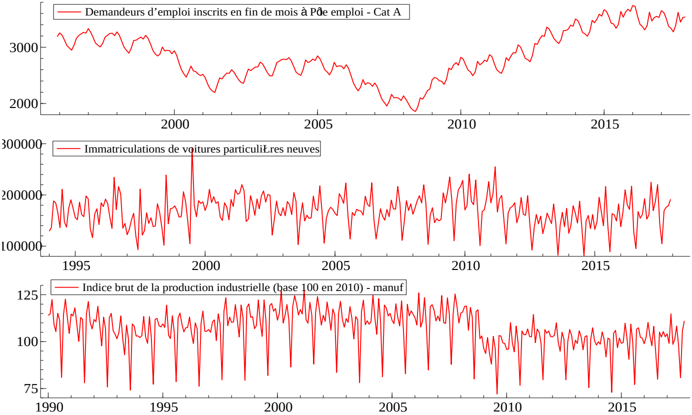
Techniques de Prévision et Conjoncture (M1 EKAP)
2020-2021

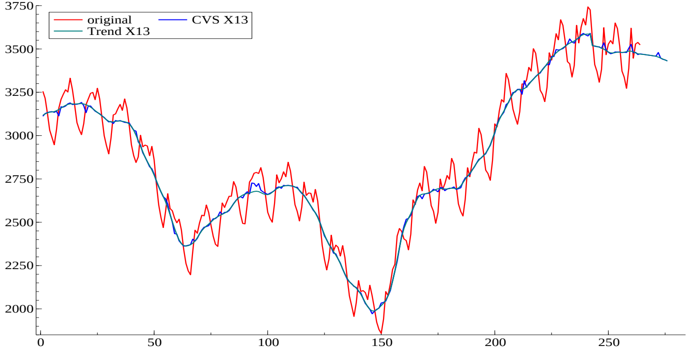
Olivier DARNÉ

Références Bibliographiques



- Hyndman et Athanasopoulos (2018), *Forecasting: Principles and Practice*, OTexts. **[online]**
- Eurostat (2018), *Handbook of Seasonal Adjustment*, European Commission.
- Aragon (2016), *Série temporelles avec R*, Springer.


Demandeurs d'emploi inscrits en fin de mois à Pôle emploi - Cat A
3000
2000
2000
2005
2010
2015
Immatriculations de voitures particulières neuves
300000
200000
100000
1995
2000
2005
2010
2015
Indice brut de la production industrielle (base 100 en 2010) - manuf
125
100
75
1990
1995
2000
2005
2010
2015


original
CVS X13
Trend X13
3750
3500
3250
3000
2750
2500
2250
2000
0
50
100
150
200
250

Les composantes inobservables

W.M. Persons (1919) a donné une définition et une formalisation standards des composantes inobservables (*unobserved components*, UC) d'une série temporelle X_t observée, largement popularisées par les succès du baromètre de Harvard (analyse et prédiction du cycle) fondé sur cette méthodologie.

Il distingue quatre types d'éléments :

- une tendance de long terme ou tendance séculaire : T_t
- un mouvement cyclique superposé sur la tendance : C_t
- un mouvement saisonnier, cad un mouvement cyclique relativement régulier de période intra-annuelle : S_t
- une variation résiduelle : I_t
  Cette composante rassemble tout ce que les autres composantes n'ont pu expliquer du phénomène observé (fluctutations accidentelles, catastrophes naturelles, guerres ...)

Ces composantes inobservables peuvent être combinées selon les schémas de décomposition additif ou multiplicatif (idée ancienne datant des travaux de Buys Ballot, 1847)

$$\begin{aligned} \text{Schéma additif :} \quad & X_t = TC_t + S_t + I_t \\ \text{Schéma multiplicatif :} \quad & X_t = TC_t \times S_t \times I_t \\ \Leftrightarrow \text{Schéma log-multiplicatif :} \quad & \log(X_t) = \log(TC_t) + \log(S_t) + \log(I_t) \end{aligned}$$

avec TC_t : la composante de tendance-cycle ou tendance de court terme.

La composante cyclique est un phénomène se répétant mais contrairement à la saisonnalité sur des durées qui ne sont pas fixes et généralement plus longues. Sans informations spécifiques, il est généralement très difficile de dissocier tendance et cycle, notamment les séries courtes (15 à 20 années de données)

Par conséquent, on regroupe ces deux composantes pour former celle de tendance-cycle, aussi appelé extra-saisonnier

De nos jours, les schémas de décomposition des différents éléments tiennent compte des composantes déterministes qui affectent un certain nombre de chroniques socio-économiques (notamment celles de flux), telles que les variations de jours ouvrables (CJO) ou les effets de fêtes mobiles

Ces schémas s'écrivent alors :

$$\begin{aligned} \text{Schéma additif :} \quad & X_t = TC_t + S_t + WD_t + EE_t + I_t \\ \text{Schéma multiplicatif :} \quad & X_t = TC_t \times S_t \times TD_t \times HE_t \times I_t \\ \Leftrightarrow \text{Schéma log-multiplicatif :} \quad & \log(X_t) = \log(TC_t) + \log(S_t) + \log(WD_t) \\ & + \log(EE_t) + \log(I_t) \end{aligned}$$

- WD_t : composante de jours ouvrables (*Working Days*) mesurant l'impact sur la série de la composition journalière du mois
- HE_t : composante mesurant l'effet jours de vacances (*Holiday Effect*).
  On parle plus généralement des effets de fêtes mobiles (*moving holiday effects*) pour caractériser les effets provoqués par les fêtes qui tombent dans un mois ou dans un autre, selon le calendrier. (fête du travail, fête de Pâques, Thanksgiving ...)

Détection de la saisonnalité

Un certain nombre de tests de saisonnalité sont disponibles afin de savoir s'il est nécessaire ou pas de désaisonnaliser la série : étape préliminaire indispensable

Chaque test est basée sur l'hypothèse nulle H_0 d'absence de saisonnalité, en supposant que

- la série est $\{z_t\}$ est stationnaire (ou rendue stationnaire)
- T le nombre d'observations
- s observations par année

Les 5 tests sont les suivants :

- le test QS modifié
- le test de Friedman
- le test de Kruskal-Wallis
- le test du périodogramme
- le F-test sur des variables dichotomiques saisonnières (*seasonal dummies*)

Le test QS modifié (QS) : il contrôle si la série présente des autocorrélation positives significatives aux retards saisonniers :

$$QS = T(T+2)\left(\frac{\widehat{\rho}^2(s)}{T-s} + \frac{[\max\{0, \widehat{\rho}(2s)\}]^2}{T-2s}\right) \sim \chi^2(2)$$

avec $\rho(h) = \gamma(h)/\gamma(0)$ l'autocorrélation d'ordre h, où $\gamma(h) = E[z_{t+h}z_t] - [E(z_t)]^2$ est l'autocovariance à l'ordre h. Si $\widehat{\rho}(s) < 0$ alors $QS = 0$.

Le test de Friedman (FT) : il contrôle les différences significatives entre les rangs moyens à des périodes spécifiques pour les observations :

$$FT = \frac{s-1}{s}\sum_{i=1}^{s}\frac{n[\bar{r}_i - (s+1)/2]^2}{(s^2-1)/12} \sim \chi^2(s-1)$$

en supposant que chaque période $i = 1, \ldots, s$ possède n observations, et avec $\bar{r}_i$ le rang moyen des observation dans la période i, où les rangs sont assignés séparément pour chaque année.

Le test de Kruskal-Wallis (KW) : il suit la même idée que le test de Friedman mais se base sur le nombre d'observations spécifique à chaque période n_i et assigne les rangs sur la période d'observation entière :

$$KW = \frac{T-1}{T}\sum_{i=1}^{s}\frac{n_i[\bar{r}_i-(T+1)/2]^2}{(T^2-1)/12} \sim \chi^2(s-1)$$

Les tests FT et KW peuvent être interprétés comme des tests ANOVA

Le test du périodogramme (PD) : il contrôle si la somme pondérée de l'estimateur du périodogramme aux fréquences saisonnières est significativement différentes de zéro. L'estimateur est donné par $(2\pi)^{-1}I(\omega_j)$ avec

$$I(\omega_j) = \begin{cases} \sum_{|h|\leq T}\widehat{\gamma}(h)e^{-ih\omega_j} & \text{si } \omega_j \neq 0 \\ T|\bar{z}|^2 & \text{si } \omega_j = 0 \end{cases}$$

où $\omega_j = 2\pi j/T$ est la j-ième fréquence de Fourier pour $j = -\lfloor(T-1)/2\rfloor, \ldots, /2\rfloor$, avec $\lfloor x\rfloor$ le plus grand entier n'excédant pas x.

La somme pondérée de l'estimateur du périodogramme aux fréquences saisonnières suit une distribution de Fisher $F(s-1;T-s)$

Le F-test sur des variables dichotomiques saisonnières (SD) : il contrôle si les effets des $(s-1)$ variables dichotomiques dans le modèle RegARIMA sont tous nuls de manière simultanée.

Le modèle RegARIMA inclue un modèle SARIMA non saisonnier (ou ARIMA) + variable dichotomiques saisonnières :

$$ARIMA(p,d,q)+\mu+\sum_{i=1}^{s-1}\beta_i D_{i,t}$$

Une version pose le modèle ARIMA$(p,d,q)=$ARIMA$(0,1,1)$, et une seconde détermine les ordres p,d,q via l'algorithme de Hannan-Rissanen. Le test est

$$SD=\frac{\widehat{\beta}'Cov(\widehat{\beta})^{-1}\widehat{\beta}}{s-1}\times\frac{T-d-p-q-s-1}{T-d-p-q}\sim F(s-1;T-d-p-q-s-1)$$

avec $\widehat{\beta}$ le vecteur des coefficients estimés des variables dichotomiques saisonnières.

Le R package seastest de Ollech (2019) **[pdf]** permet d'utiliser ces 5 tests également disponibles dans JDemetra et RegARIMA.

Basé sur les travaux de Webel et Ollech (2019) le package propose un nouveau test combinant les tests QS et KW avec la fonction `wo`

Pour le test SD, la sélection automatique des ordres du modèle ARIMA est réalisée à partir de l'algorithme de Hyndman et Khandakar (2008) **[pdf]**, également utilisé dans la fonction `auto.arima` du package forecast :

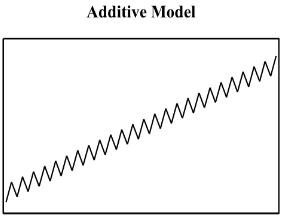
- sélection de l'ordre de différenciation d par le test de racine unitaire KPSS
- sélection des polynômes non saisonniers p et q en minimisant le critère AICc, avec $\max(p) = \max(q) \leq 3$
- la recherche se fait pas à pas (*stepwize*)

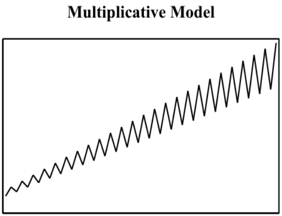
```
library(readxl)
library(seastests)
y <- read.table("c://R//data//demandbrut.txt")
yy <- ts(data = y, start=c(1996,01),frequency=12)

# Friedman test
ft <- fried(yy)
show(ft)
# Testing the seasonality of series
# a boolean value is returned : TRUE or FALSE
is <- isSeasonal(yy, test="wo")
show(is)
# Kruskal-Wallis test
kwt <- kw(yy)
show(kwt)
# QS test
qst <- qs(yy)
show(qst)
# Seasonal dummies (SD)
sd <- seasdum(yy)
show(sd)
# Welch test
w <- welch(yy)
show(w)
# Webel-Ollech test
wot <- wo(yy)
show(wot)
```

Friedman test
```
Test used:  Friedman rank
Test statistic:  205.41
P-value:  0
```
Testing the seasonality of series
```
[1] TRUE
```
Kruskal-Wallis test
```
Test used:  Kruskall Wallis
Test statistic:  223.92
P-value:  0
```
QS test
```
Test used:  QS
Test statistic:  398.29
P-value:  0
```
Seasonal dummies (SD)
```
Test used:  SeasonalDummies
Test statistic:  139.15
P-value:  0
```
Welch test
```
Test used:  Kruskall Wallis
Test statistic:  139.8
P-value:  5.772899e-55
```
Webel-Ollech test
```
Test used:  WO
Test statistic:  1
P-value:  0 0 0
```

Choix du modèle (additif ou multiplicatif)

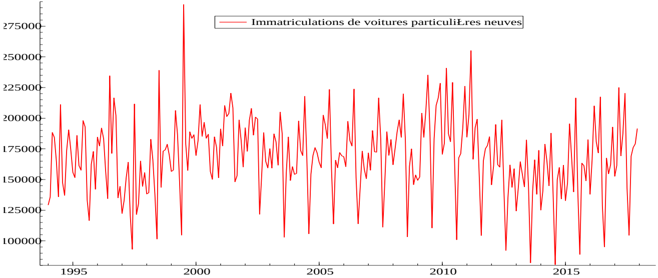
Avant toute modélisation et étude aprofondie du modèle, on tente d'abord de déterminer si on est en présence d'une spécification additive ou multiplicative





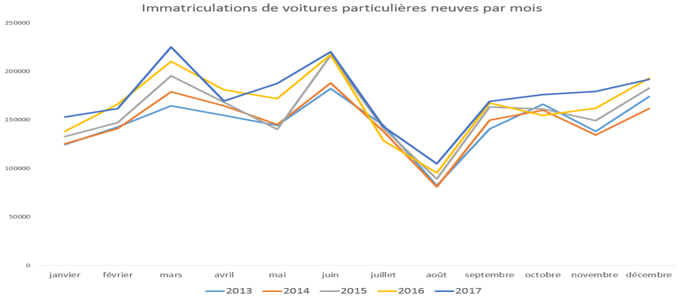
Il existe 4 approches pour déterminer le type de schéma :

- Méthode du profil (graphique)
- Méthode de la bande
- Méthode analytique ou test de Buys-Ballot
- Test log-level



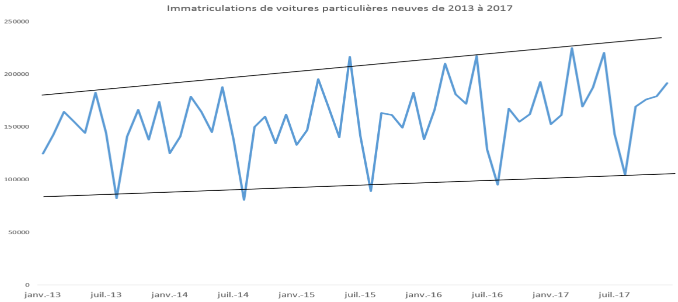
Méthode du profil (graphique) : on superpose les saisons représentées par des courbes de profil sur un même graphique.

Si ces courbes sont parallèles, le modèle est additif, autrement le modèle est multiplicatif



Méthode de la bande : on fait un graphique représentant la série chronologique, puis on trace une droite passant respectivement par les minima et par les maxima de chaque saison.

Si ces deux droites sont parallèles, nous sommes en présence d'un modèle additif. Dans le cas contraire, c'est un modèle multiplicatif.
(attention à prendre une période sans changement de tendance)



Méthode analytique ou test de Buys-Ballot : on test si les moyennes et les écarts-types sont indépendantes.
On calcule les moyennes et les écarts-types pour chacune des périodes considérées puis la droite des moindres carrées

$$\sigma = a\overline{x} + b$$

- Si $a = 0$ significativement $\Rightarrow \sigma$ et $\overline{x}$ indépendants $\Rightarrow$ modèle additif
- Si $a \neq 0$ significativement $\Rightarrow \sigma$ et $\overline{x}$ dépendants $\Rightarrow$ modèle multiplicatif

date	2013	2014	2015	2016	2017
janvier	124384	125125	132815	138050	152708
février	142556	140717	147060	166230	161374
mars	164448	178758	195362	209981	224936
avril	154407	164400	168144	180746	169342
mai	144078	144988	139957	171731	187481
juin	182245	187756	216465	217289	220196
juillet	144303	138048	141489	128370	143062
août	82157	80669	88984	95114	104626
septembre	140652	149766	163328	167466	169077
octobre	165992	159789	161254	154689	176078
novembre	137993	134223	148972	162082	179075
décembre	173736	161616	182403	192725	191556
moyenne	146412.6	147154.6	157186.1	165372.8	173292.6
écart-type	26191.5	27834.2	32669.6	34234.8	32573.2
a =	0.25				
b =	-9135.47				

Test log-level : les méthodes X13-ARIMA, TRAMO-SEATS et X13-SEATS-ARIMA proposent une procédure de pré-test pour tester la spécification log-level.

Ils comparent le critère AICc de l'estimation du maximum de vraisemblance d'une transformation Box-Cox avec

- $\lambda = 0$: transformation en logs $\Rightarrow$ décomposition multiplicative
- $\lambda = 1$: transformation en niveau (*level*) $\Rightarrow$ décomposition additive

$$B(X_t, \lambda) = \begin{cases} \frac{X_t^\lambda - 1}{\lambda} & \text{si } \lambda \neq 0 \\ \log(X_t) & \text{si } \lambda = 0 \\ X_t & \text{si } \lambda = 1 \end{cases}$$

L'utilisateur peut choisir dans le package seasonal entre

- pas de transformation : `transform.function = "none"`
- transformation logarithmique : `transform.function = "log"`
- test la spécification log-level (option de défaut)

La fonction `out` du package seasonal affiche le contenu de la sortie principale du X-13 dans le navigateur

```
library(seasonal)
seasX <- seas(yy)
out(seasX)
```

Reading input spec file from *C:\Users\darne-o\AppData\Local\Temp\Rtmpy2RVq5\x13out8c8076f328eb/iofile.spc*

Reading data from *C:\Users\darne-o\AppData\Local\Temp\Rtmpy2RVq5\x13out8c8076f328eb/data.dta*

Previous Table | Index | Next Table

U. S. Department of Commerce, U. S. Census Bureau

X-13ARIMA-SEATS monthly seasonal adjustment Method, Release Version 1.1 Build 39

This software application provides an enhanced version of Statistics Canada's X-11-ARIMA extension (Dagum, 1980) of the X-11 variant of the Census Method II of Shiskin, Young and Musgrave (1967).

It also provides an ARIMA model-based method following Hillmer and Tiao (1982) and Burman (1980) that is very similar to the update of the method of SEATS (Gómez and Maravall, 1996) produced at the Bank of Spain by G. Caporello and A. Maravall for TSW (Caporello and Maravall, 2004). The present application includes additional enhancements.

X-13ARIMA-SEATS includes an automatic ARIMA model selection procedure based largely on the procedure of Gómez and Maravall (1998) as implemented in TRAMO (1996) and subsequent revisions.

Primary Programmers: Brian Monsell, Mark Otto and, for the ARIMA model-based signal extraction, Gianluca Caporello and Víctor Gómez

Index for x13out8c8076f328eb/iofile.ht

Likelihood statistics for model fit to untransformed series.

Likelihood Statistics	
Number of observations (nobs)	264
Effective number of observations (nefobs)	251
Number of parameters estimated (np)	3
Log likelihood (L)	-1192.3287
AIC	2390.6574
AICC (F-corrected-AIC)	2390.7546
Hannan Quinn	2394.9136
BIC	2401.2338

Likelihood statistics for model fit to log transformed series.

Likelihood Statistics	
Number of observations (nobs)	264
Effective number of observations (nefobs)	251
Number of parameters estimated (np)	3
Log likelihood	793.2709
Transformation Adjustment	-1991.9574
Adjusted Log likelihood (L)	-1198.6865
AIC	2403.3730
AICC (F-corrected-AIC)	2403.4702
Hannan Quinn	2407.6292
BIC	2413.9494

***** AICC (with aicdiff=-2.00) prefers **no transformation** *****

La désaisonnalisation

La désaisonnalisation ou correction des variations saisonnières (CVS) ou ajustement saisonnier (seasonal adjustment, SA) a pour objectif d'éliminer les composantes saisonnières des séries temporelles économiques afin de mettre en évidence les autres éléments (tendance et cycle) qui jouent un rôle important dans l'analyse économique.

La série désaisonnalisée ou corrigées des variations saisonnières (CVS) est définie par

$$SA_t = TC_t + I_t$$

La série CVS et corrigée des jours ouvrables (CVS-CJO) est définie par

$$SA_t = TC_t + I_t - WD_t$$

L'idée d'identifier et de supprimer la saisonnalité dans les chroniques est très ancienne, et remonte au XIXème siècle. En effet, Jevons (1862) préconise

> *Chaque type de fluctuations périodiques, qu'elles soient journalières, hebdomadaires, trimestrielles ou annuelles, doit être détecté. [. . .] En effet, nous devons établir et éliminer de telles variations périodiques avant de pouvoir correctement montrer celles qui sont irrégulières et non périodiques, et qui probablement présentent plus d'intérêt et d'importance.*

William Stanley Jevons
économiste anglais
(1835-1882)

D'après Hylleberg (1992), le problème de la désaisonnalisation se produit dans au moins deux contextes économiques différents :

- les études historiques des cycles d'affaires (Burns et Mitchell, 1946 ; Kaiser et Maravall, 2000)
- l'évaluation des conditions économiques actuelles ou conjoncturelles : analyse des phases récentes de récession et d'expansion de l'activité économique (Moore, 1961 ; Dagum, 2001)

Hylleberg souligne également que l'utilisation de l'analyse économétrique sur des données désaisonnalisées augmente le risque d'une mauvaise spécification du modèle, entraînant ainsi des relations dynamiques fallacieuses et de faibles prévisions (tests de racines unitaires et de cointégration saisonnières)

Néanmoins, la plupart des modèles macro-économétriques emploient des données corrigées des variations saisonnières (CVS) produites par les agences de statistique.

Un certain nombre de méthodes de désaisonnalisation sont apparues dans les années 20 et 30, mais c'est principalement depuis l'essor de l'informatique après la seconde guerre mondiale qu'elles se sont développées.

De nombreuses recherches se sont intéressées à la correction des variations saisonnières présentes dans les données sans toutefois en modifier les caractéristiques.

Cette opération est délicate, ce qui explique le développement de diverses méthodes dont le but est d'améliorer l'ajustement saisonnier.

Approches temporelle et fréquentielle

Une série temporelle peut être considérée de deux points de vue : celui du temps et celui des fréquences

- Dans le **domaine temporel**, on regarde la série $\{X_t\}$ comme une succession de T valeurs observées aux instants t, avec $t = 1, \ldots, T$
- Dans le **domaine des fréquences**, on part de l'expression de la série $\{X_t\}$ comme une somme de fonctions sinusoïdales. On mesure alors pour chaque fréquence, l'importance qu'elle a dans la composition de la série

Ces deux approches s'avèrent souvent complémentaires

Approche temporelle : il est facile de représenter graphiquement son évolution au cours du temps. On note que cette série est caractérisée par une forte saisonnalité traduisant la chute de l'activité industrielle au mois d'août

Les modélisations de la série, ou de ses composantes, mettant en relation la valeur à l'instant t et celles des instants passés sont facile à formaliser.

C'est le cas par exemple de la modélisation de la série par un modèle ARIMA saisonnier (SARIMA), de l'expression d'une tendance linéaire, exponentielle ou encore localement polynomiale, ou de la modélisation de la composante irrégulière par un bruit blanc.

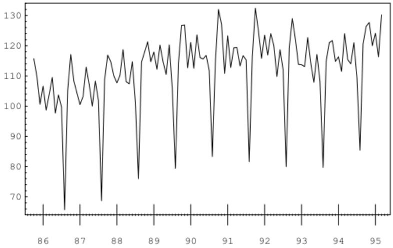


FIG. 3.1 – Indice mensuel de la production industrielle en France, d'octobre 1985 à mars 1995.

Approche fréquentielle : le spectre de la série associe à chaque fréquence son importance dans la série. Le spectre de l'indice de la production industrielle laisse apparaître une forte contribution (pic spectral) de la fréquence $\pi/6 = 30^\circ$, et de ses multiples $2\pi/6 = 60^\circ$, $3\pi/6 = 90^\circ, \ldots, 6\pi/6 = 180^\circ$.
La période associée à cette fréquence est $\omega = 2\pi/f = 2\pi/(\pi/6) = 12$, et nous retrouvons la saisonnalité mensuelle observée sur la graphique précédent.

Les basses fréquences correspondent par nature à des composantes évoluant lentement, tendance et cycle par exemple, et les hautes fréquences à des composantes évoluant plus vite comme la composante irrégulière.

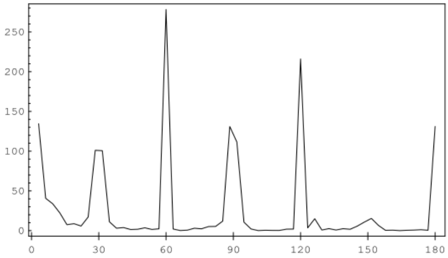
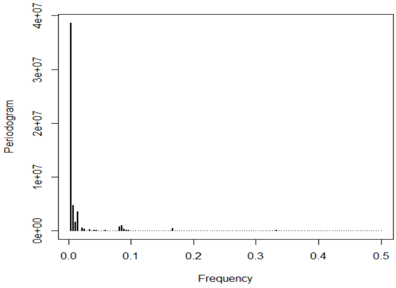


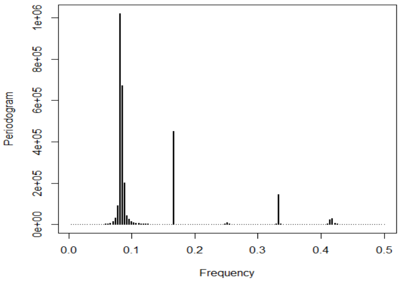
FIG. 3.2 – Spectre de l'indice de la production industrielle française.

Le **périodogramme** permet de représenter l'évolution des séries dans le domaine des fréquences : package TSA

```
library(TSA)
y <- read.table("c://R//data//demandbrut.txt")
yy <- ts(data = y, start=c(1996,01),frequency=12)
periodogram(yy)
```



série originale



série en différence 1ère

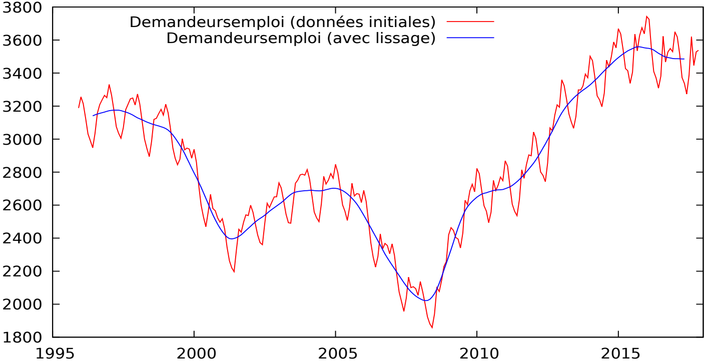
Les procédures sont généralement classées en trois catégories :

- les méthodes non paramétriques (ou modèles implicites) fondées sur des filtres de lissage linéaire où les variables inobservables sont modélisées de manière implicite
- les méthodes paramétriques (ou modèles explicites) pour lesquelles les composantes inobservables sont modélisées et estimées explicitement
  - les méthodes paramétriques déterministes où chaque composante a un comportement « déterministe » dans le sens de la décomposition de Wold (1938) représentée par des modèles de régression
  - les méthodes paramétriques stochastiques basée sur des comportements « stochastiques » ou « purement linéaires » en utilisant des modèles univariés de type ARIMA (*Autoregressive Integrated Moving Average*)
- les méthodes semi paramétriques (ou modèles mixtes) basées sur des modèles hybrides (combinant des modèles implicites et explicites)

1. Les méthodes non paramétriques (modèles implicites)

En matière de désaisonnalisation des séries économiques, les filtres de lissage linéaire ou moyennes mobiles sont déjà connus au début des années 20, mais sont rarement utilisés.

L'instrument le plus utilisé à l'époque pour éliminer la composante saisonnière mensuelle était la moyenne mobile centrée d'ordre 12


Demandeursemploi (données initiales)
Demandeursemploi (avec lissage)
3800
3600
3400
3200
3000
2800
2600
2400
2200
2000
1800
1995
2000
2005
2010
2015

La moyenne mobile centrée d'ordre 12 s'avère un piètre estimateur de la tendance-cycle :

- cette moyenne ne peut pas suivre avec précision les sommets et les creux des cycles conjoncturels à court terme (d'une périodicité de 5 ans ou plus)
- à moins que les fluctuations accidentelles aient une faible amplitude, elle ne comprend pas suffisamment de termes pour assurer un lissage adéquat des données
- les moyennes mobiles sont très sensibles aux cas erratiques (points atypiques, *outliers*) et nécessitent de ce fait un traitement *a priori* des valeurs extrêmes

Vers la fin des années 20, l'élaboration de nouveaux filtres de lissage et de techniques d'application différentes permet de diffuser cette approche non paramétrique pour la désaisonnalisation des séries temporelles économiques (Macaulay, 1931)

Progressivement, les gouvernements et les bureaux de statistique commencent à appliquer les procédés de lissage pour désaisonnaliser leurs séries, mais la méthode est coûteuse, longue et subjective, parce que les ajustements sont réalisés, pour la plupart, à la main.

Le développement de l'informatique, après la seconde guerre mondiale, contribue à la propagation et à l'amélioration des méthodes non paramétriques de désaisonnalisation.

La méthode la plus connue est sûrement la méthode X-11 (Shiskin, Young et Musgrave, 1967).

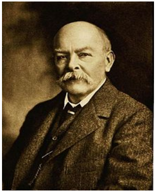
Par la suite, de nombreuses méthodes dont le principe de construction est identique à celui de X-11 ont été développées et sont appelées « style X-11 »

Les moyennes mobiles

Ces outils font partie des premières méthodes pour l'analyse des séries chronologiques. Il semble que le physicien John Henry Poynting soit le premier, en 1884, à avoir utilisé les moyennes mobiles pour éliminer les variations accidentelles ou périodiques d'une série.

En 1904, Spencer introduit une moyenne mobile (symétrique d'ordre 15) permettant de conserver les polynômes de degré 3.

Puis à partir de 1914, les grands personnages de la statistique tels que Karl Pearson et George Yule par exemple, s'intéressent à ce genre de problèmes.



John Henry Poynting
physicien anglais

Karl Pearson
mathématicien britannique
(1857 -1936)

Karl Pearson
mathématicien britannique

George Yule
mathématicien britannique
(1871 –1951)

George Yule
mathématicien britannique

Définition

Une **moyenne mobile** (*moving average*) **de coefficients** θ_i, noté $M(\theta_i)$ ou M, est un opérateur linéaire du type

$$M = \sum_{i=-p}^{f} \theta_i L^{-i} = \sum_{i=-p}^{f} \theta_i F^{i}$$

- p et f sont 2 entiers positifs
- $\theta_{-p}, \dots, \theta_{-1}, \theta_0, \theta_1, \dots, \theta_f$: **coefficients de** M (nombres réels)
- L : **opérateur de retard** (*backward*) $L^p X_t = X_{t-p}$, avec $L^0 = 1$
- F : **opérateur d'avance** (*forward*) $F = L^{-1}$
- ordre de le moyenne mobile M : $p + f + 1$ (nombre de coeffcients)
- une moyenne mobile M est **centrée** si $f = p$
- une moyenne mobile M est **symétrique** si M est centrée et si $\theta_i = \theta_{-i}$
- moyenne mobile centrée est automatiquement d'ordre impair : $p + f + 1 = 2p + 1$

Pour une série chronologique X_t une moyenne mobile en t est une combinaison linéaire finie des valeurs de la série correspondant à des dates entourant t, elle réalise donc un lissage de la série, une moyennisation

$$\begin{aligned} M(X_t) &= \sum_{i=-p}^{f} \theta_i X_{t+i} = \theta_{-p} X_{t-p} + \cdots + \theta_{-1} X_{t-1} + \theta_0 X_t + \theta_1 X_{t+1} + \cdots + \theta_f X_{t+f} \\ &= \sum_{i=-p}^{f} \theta_i L^i X_t = \theta(L) X_t \end{aligned}$$

- pour une série temporelle on doit avoir : $p+1 \leq t \leq T-f$

Propriétés

- M **conserve** X_t ou X_t est invariante par M si $M(X) = X$
- X_t est **absorbée** par M si $M(X) = 0$

Propriétés : préservation de la tendance

- une moyenne mobile M centrée **conserve** les suites constantes ssi

$$\sum_{i=-p}^{p} \theta_i = 1$$

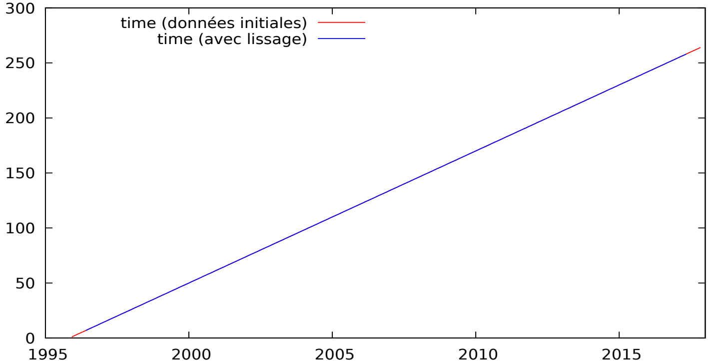
- si M est symétrique et conserve les constantes, alors M **conserve** les polynômes de degré 1 (i.e $X_t = at + b$) ssi

$$\sum_{i=-p}^{p} \theta_i = 1 \qquad \text{et} \qquad \sum_{i=-p}^{p} i\theta_i = 0$$

- une moyenne mobile M centrée conserve tout polynôme de degré d ssi ses coefficients vérifient

$$\sum_{i=-p}^{p} \theta_i = 1 \qquad \text{et} \qquad \sum_{i=-p}^{p} i^h\theta_i = 0 \qquad\qquad 1 \leq h \leq d$$

Figure: Moyenne mobile centrée à 12 termes sur une tendance linéaire



Les moyennes mobiles arithmétiques

Pour éliminer la saisonnalité, les **moyennes mobiles d'ordre** k permettent d'éliminer la **saisonnalité de période** k.

Définition

Une **moyenne mobile arithmétique** M **d'ordre impair** $k = 2p+1$ est définie par $\theta_i = \frac{1}{(2p+1)}, \forall i = -p, \ldots, p$:

$$M(X_t) = \frac{1}{2p+1} \sum_{i=-p}^{p} X_{t+i}$$

Propriété

Une **moyenne mobile arithmétique d'ordre impair** $2p+1$ absorbe les saisonnalités de périodes $k = 2p+1$ qui sont nulles en moyenne, et elle préserve les polynômes de degré 1

Toutefois, en pratique, il est souvent nécessaire de considérer des saisonnalités de période paire (données trimestrielles ou mensuelles).

Pour cela, si la période est $k = 2p$, on utilise la moyenne mobile $M = \frac{1}{2}M_1 + \frac{1}{2}M_2$

$$
\begin{aligned}
M_1 &= \frac{1}{2p}\sum_{i=-p}^{p-1} X_{t+i} = \frac{1}{2p}(X_{t-p} + \cdots + X_{t+p-1}) \\
M_2 &= \frac{1}{2p}\sum_{i=-p+1}^{p} X_{t+i} = \frac{1}{2p}(X_{t-p+1} + \cdots + X_{t+p})
\end{aligned}
$$

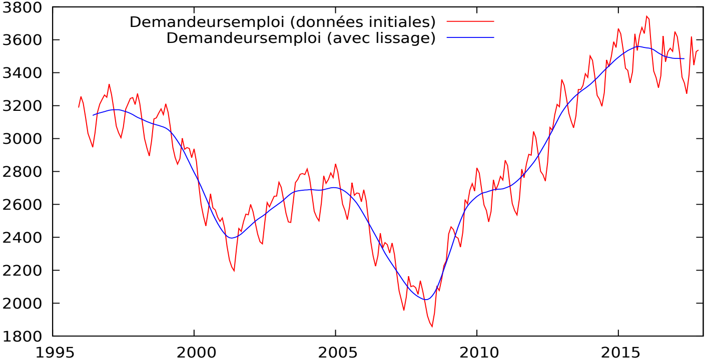
Les moyennes mobiles M_1 et M_2 absorbent les saisonnalités de période $2p$ et de moyenne nulle mais elles ne sont pas symétriques

Définition

La moyenne mobile arithmétique M **d'ordre pair** $k = 2p$ est **symétrique**, **annule** les saisonnalités de période $2p$ et de moyenne nulle, et **préserve** les polynômes de degré 1

$$\begin{aligned} M(X_t) &= \frac{1}{2p}\left(\frac{X_{t-p}}{2} + \sum_{i=-p+1}^{p-1}(X_{t+i}) + \frac{X_{t+p}}{2}\right) \\ &= \frac{1}{2p}\left(\frac{X_{t-p}}{2} + X_{t-p+1} + \cdots + X_{t+p-1} + \frac{X_{t+p}}{2}\right) \end{aligned}$$

Figure: Moyenne mobile centrée à 12 termes



Les moyennes mobiles simples composées

Une moyenne mobile dite $P \times Q$, notée $M_{P \times Q}$ s'obtient en composant

- une moyenne mobile simple d'ordre P, de coefficients tous égaux à $1/P$
- une moyenne mobile simple d'ordre Q, de coefficients tous égaux à $1/Q$
- cela revient à appliquer successivement à la série les deux moyennes mobiles simples

Une moyenne mobile $M_{P \times Q}$ est une **moyenne mobile symétrique d'ordre** $P+Q+1$

Quelques exemples de moyennes mobiles $M_{P \times Q}$

- Estimation de la tendance-cycle TC_t : $M_{2 \times 12}$
- Estimation de la saisonnalité S_t : $M_{3 \times 3}$, $M_{3 \times 5}$ ou $M_{3 \times 9}$

NB : voir en annexe l'expression sous forme de polynome de la moyenne mobile centrée à 12 termes $M_{2 \times 12}$

Effet d'une moyenne mobile sur les fluctuations irrégulières

Par construction, une moyenne mobile consiste à un lissage de la série.

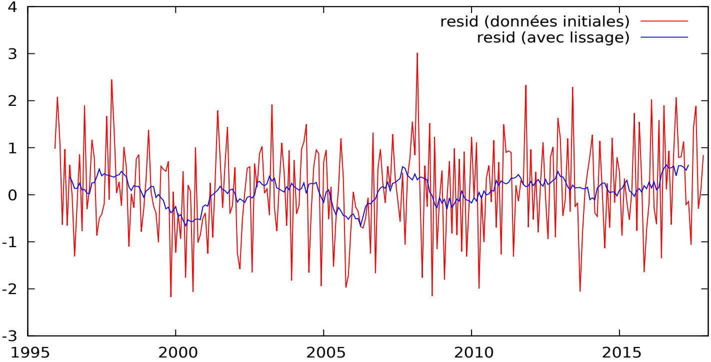
L'effet de la composante irrégulière est d'autant plus atténué que l'ordre de la moyenne mobile est grand

Les moyennes mobiles arithmétiques d'ordre $k = 2p+1$ sont les moyennes mobiles **minimisant la variance** d'un bruit blanc parmi les moyennes mobiles centrées telles que

$$\sum_{i=-p}^{p} \theta_i = 1$$

Ainsi les moyennes arithmétiques d'ordre $k = 2p+1$ transforment un bruit blanc en un processus **centré** (inchangé) et de variance $\sigma^2/(2p+1)$ **réduite**

Figure: Moyenne mobile arithmétique à 13 termes sur la composante irrégulière



	Moyenne	Médiane	S.D.	Min	Max
resid	0.1124	0.09426	0.9876	-2.173	3.017
ma_resid	0.09331	0.1289	0.2747	-0.6597	0.6358

La méthode X-11

En 1954, Julius Shiskin met au point une méthode (*Method I*) au US Bureau of the Census .

Cette technique de désaisonnalisation est suivie par 11 versions eXpérimentales d'une *Method II* (X0, X1, etc) pour finalement aboutir au logiciel X-11 en 1965 (Shiskin, Young et Musgrave, 1967).

Inspirées directement des lissages par moyennes mobiles et des travaux de Macaulay (1931), ces diverses versions constituent les premières méthodes automatiques de désaisonnalisation et X-11 devient rapidement un standard utilisé dans le monde entier.

La méthode X-11 est basée sur l'utilisation de différentes sortes de moyennes mobiles symétriques et asymétriques pour décomposer une série temporelle en ses éléments de tendance, saisonnier et irrégulier, sans modèle explicite sous-jacent.

Les estimations des composantes saisonnière et de tendance résultent d'un filtrage en cascade, cad de l'application successive de divers filtres linéaires individuels, appliqués de manière séquentielle. Chaque composante s'obtient après plusieurs itérations

Pour des données mensuelles, ce filtrage est composé de :

- une moyenne mobile centrée à 12 termes
- deux moyennes mobiles saisonnières d'ordre $3 \times (2p+1)$, et
- une moyenne mobile de Henderson : estimation lisse (*smooth*) de la tendance-cycle (voir annexe)
- pour les observations en début et en fin de série, X-11 emploie des moyennes mobiles asymétriques dérivées de la méthode de Musgrave (1964) afin de prolonger la série (voir annexe)

L'algorithme de la méthode X-11 est présenté en annexe

Avantage

- La force principale du logiciel X-11 vient de son application massive dans les organismes et les instituts de statistique nationaux et internationaux durant des années, ce qui a permis de l'enrichir avec des mises à jour régulières.

Inconvénients

- Son inconvénient majeur a été une estimation peu fiable des valeurs désaisonnalisées en fin de série (application des filtres asymétriques $\Rightarrow$ révisions importantes lorsque de nouvelles observations sont disponibles).
- Ces valeurs en fin de série sont importantes pour évaluer la direction de la tendance de court terme et identifier un point de retournement cyclique dans l'économie.
- En outre, les tests statistiques étaient peu nombreux et les indicateurs pour évaluer la qualité de la désaisonnalisation étaient absents.

Solution : méthodes semi-paramétriques X11-ARIMA

Par la suite, de nombreuses méthodes dont le principe de construction est identique à celui de X-11 ont été développées et sont appelées « style X-11 »

- La méthode SABL (*Seasonal Adjustment at Bell Laboratories*), développée par Cleveland, Dunn et Terpenning (1978) à Bell Laboratories
- La méthode SEASABS (*SEASonal Analysis at Australian Bureau of Statistics*), développée par l'Australian Bureau of Statistics (1987)
- Le logiciel GLAS (*General Linear Abstraction of Seasonality*), développé par Young (1992) à la Banque d'Angleterre
- Le logiciel STL (*Seasonal-Trend decomposition based on Loess*, proposé par Cleveland *et alii* (1990) à Bell Laboratories
  Le package stats avec la fonction stl ou le package stlplus décompose la série à partir de l'**approche STL**

La méthode STL

Dans la méthode STL (*Seasonal-Trend decomposition based on Loess*), proposée par Cleveland *et alii* (1990), chaque composante est déterminée par une régression locale (Cleveland, 1979, 1981 ; Cleveland et Devlin, 1988), aussi appelée :

- lissage LOESS (*LOcally Estimated Scatterplot Smoothing*) ou
- lissage LOWESS (*LOcally WEighted Scatterplot Smoothing*)

C'est une méthode non paramétrique, appelée aussi régression polynomiale locale, avec pondération locale, basée sur la méthode des k plus proches voisins (k-NN, k*-nearest neighbors*, méthode d'apprentissage supervisé).

La procédure est la suivante :

- Pour chaque valeur x d'une série X, on va considérer les k plus proches voisins à gauche et à droite notées x_i (k-NN)
- La distance locale maximale est calculée parmi les k-NN :

$$\lambda(x) = \max_i |x_i - x|$$

- Des poids sont attribués à chaque point du voisinage k-NN x_i :

$$\omega_i(x) = W\left(\frac{|x_i - x|}{\lambda(x)}\right)$$

  avec $W(.)$ une fonction de poids cubique :

$$W(u) = \begin{cases} (1-u^3)^3 & \text{pour } 0 \leq u < 1 \\ 0 & \text{pour } u \geq 1 \end{cases}$$

  Les x_i proches de x auront les poids les plus importants

  Les poids décroissent lorsque x_i s'éloigne de x et deviennent nuls au k-ième point le plus éloigné

- Ensuite, un polynôme de degré d est ajusté aux données avec les poids $\omega_i(x)$
  - soit un polynôme de degré $d = 1$ (localement linéaire)
  - soit un polynôme de degré $d = 2$ (localement quadratique)
  - le degré $d = 0$ revient à calculer des moyennes mobiles pondérées

2. Les méthodes semi-paramétriques (modèles hybrides)

La vulgarisation des modèles ARIMA (*AutoRegressive Integrated Moving Average*) à partir de l'ouvrage de Box et Jenkins (1970) a permis de faire progresser les outils de désaisonnalisation.

⇒ évolution de la méthode X-11 vers la famille des méthodes de type X11-ARIMA, en améliorant notamment l'estimation des observations récentes (défaut majeur de X-11) sous forme de logiciel :

- X11-ARIMA/80 par Dagum (1980) à Statistique Canada
- X12-ARIMA par Findley et alii (1988, 1998) au US Bureau of the Census
- X13-ARIMA-SEATS par Monsell et alii (2013) au US Bureau of the Census

La série initiale X_t est modélisée par un processus ARIMA saisonnier (SARIMA) puis prolongée en début et en fin de série, limitant ainsi les révisions des estimations lorsque l'on dispose d'une observation supplémentaire.

Cette approche permet ainsi de prévoir jusqu'à trois années de données supplémentaires

Néanmoins, il faut noter que les modèles ARIMA ne peuvent pas être appliqués directement sur les séries économiques sans une analyse préalable de celles-ci, notamment à cause

- du problème de stationnarité
- des effets déterministes (effets de calendriers, outliers . . .)

Un module de pré-traitement des données, appelé RegARIMA (*Regression ARIMA*) a été développé, permettant une estimation simultanée des points aberrants (*outliers*), des variations de jours ouvrables et des effets de calendrier, avec un modèle SARIMA.

La procédure de sélection du modèle SARIMA dans la version de la méthode X-13ARIMA-SEATS, à travers le module RegARIMA, utilise l'algorithme de Hannan et Rissanen (1982) dont l'avantage est sa rapidité de calcul.

Le module RegARIMA (idem TRAMO) impose des contraintes sur les polynômes des modèles SARIMA$(p, d, q)(P, D, Q)_S$ avec

- les polynômes non saisonniers : $p, q \leq 3$
- les polynômes saisonniers : $P, Q \leq 1$
- les ordres de différentiation non saisonnier et saisonnier : $d \leq 3$ et $D \leq 2$

L'algorithme de Hannan-Rissanen est utilisée pour sélectionner les ordres des polynômes non saisonniers (p, q) les plus adaptés, ayant pour objectif la sélection d'un modèle parcimonieux et équilibré :

- parcimonieux : un petit nombre de paramètres avec un bon pouvoir explicatif
- équilibré : un modèle pour lequel l'ordre des polynômes AR et de différentiation est égal à celui du polynôme MA

L'algorithme de Hannan-Rissanen pour un modèle ARMA(p,q) suit les étapes suivantes :

- la procédure de Hannan-Rissanen opère sur une transformation stationnaire de la série originale $\Rightarrow$ identification uniquement des paramètres AR et MA
- ajustement d'un modèle $AR(m)$ où $m > \max(p,q)$ sur X_t. Dans le module RegARIMA $p_{\max} = q_{\max} = 3$ (idem module TRAMO)
- les résidus obtenus $\widehat{a}_k$ sont utiliser pour estimer les innovations du modèle ARMA a_t

$$a_t = X_t - \sum_{k=1}^{m} \widehat{a}_k X_{t-k}$$

- les paramètres p et q du modèle ARMA sont estimés à partir d'une régression linéaire des moindres carrés de X_t sur $(X_{t-1}, \ldots, X_{t-p}, a_{t-1}, \ldots, a_{t-q})$
- une paire (p,q) est sélectionnée à partir du critère BIC$_{p,q}$ (*Bayesian information criterion*) :

$$BIC_{p,q} = \log(\sigma^2_{p,q}) + \frac{(p + q\log(n-d))}{n-d}$$

avec $\sigma^2_{p,q}$ l'estimateur du maximum de vraisemblance de σ^2, n le nombre d'observations, et d l'ordre de différentiation

Automatic ARIMA Model Selection

Procedure based closely on TRAMO ,method of Gomez and Maravall (2000)
"Automatic Modeling Methods for Univariate Series",
A Course in Time Series (Edited by D. Pena, G. C. Tiao, R. S. Tsay),
New York : J. Wiley and Sons

Maximum order for regular ARMA parameters : 2

Maximum order for seasonal ARMA parameters : 1

Maximum order for regular differencing : 2

Maximum order for seasonal differencing : 1

Results of Unit Root Test for identifying orders of differencing:

Regular difference order : 2 Seasonal difference order : 1

Mean is not significant.

Best Five ARIMA Models

Model # 1 : (0 2 1)(0 1 1) (BIC2 = 9.313)
Model # 2 : (1 2 1)(0 1 1) (BIC2 = 9.325)
Model # 3 : (0 2 2)(0 1 1) (BIC2 = 9.326)
Model # 4 : (2 2 0)(0 1 1) (BIC2 = 9.346)
Model # 5 : (2 2 1)(0 1 1) (BIC2 = 9.347)

Preliminary model choice : (0 2 1)(0 1 1)

ARIMA Model

ARIMA Model: (0 2 1)(0 1 1)

Nonseasonal differences: 2
Seasonal differences: 1

ARIMA Model

	Estimate	Standard Error
Nonseasonal MA		
Lag 1	0.69654	0.04526
Seasonal MA		
Lag 12	0.43662	0.05987

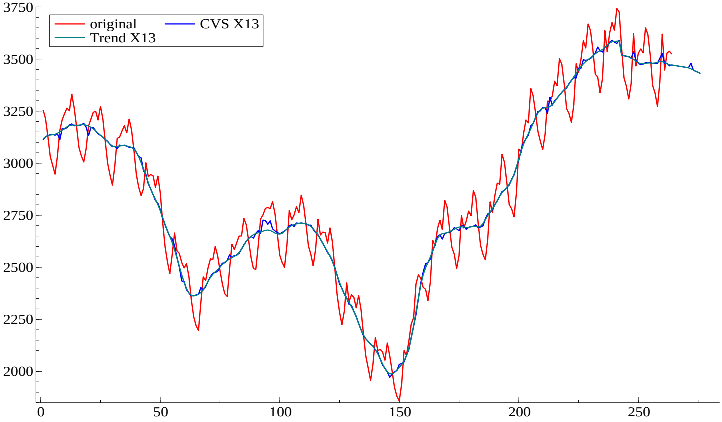
Model Innovation Variance

Variance	0.59935E+03
Standard Error of Variance	0.53607E+02

Previous Table | Index | Next Table

Likelihood Statistics

Number of observations (nobs)	264
Effective number of observations (nefobs)	250
Number of parameters estimated (np)	4
Log likelihood (L)	-1155.8212
AIC	2319.6425
AICC (F-corrected-AIC)	2319.8057
Hannan Quinn	2325.3116
BIC	2333.7283


original
CVS X13
Trend X13
3750
3500
3250
3000
2750
2500
2250
2000
0
50
100
150
200
250

3. Les méthodes paramétriques (modèles explicites)

Certains auteurs ont critiqué les approches non paramétriques de désaisonnalisation basées sur les filtres linéaires ou moyennes mobiles, cad des procédures « empiriques » *ad hoc* alors qu'il existe des outils mathématiques adéquats.

L'insatisfaction de ces méthodes conduit à l'utilisation de modèles explicites de séries temporelles pour la désaisonnalisation des séries

Deux types de méthodes se sont développées :

- les méthodes fondées sur des modèles déterministes
- les méthodes basées sur des modèles stochastiques

Les méthodes déterministes

Ces méthodes déterministes supposent que chaque composante a un comportement « déterministe » dans le sens de la décomposition de Wold (1938), représentée par des modèles de régression

Définition : Théorème de Wold

Il est toujours possible de décomposer un processus **stationnaire** X_t en

- une **composante déterministe** d_t (parfaitement prévisible) et
- une **composante stochastique** u_t

telles que

$$X_t = d_t + u_t \qquad \text{avec } u_t = \sum_{i=0}^{\infty} \alpha_i \varepsilon_{t-i}$$

où ε_t est un bruit blanc, cad moyenne nulle, variance constante et non autocorrélé.

Les méthodes de régression fournissent les premières approches de désaisonnalisation basées sur le modèle (AMB, *Model-Based Approaches*)

Ces méthodes s'appuient sur

- une modélisation de la série originale (X_t) et de chacune des composantes (C_t, T_t, S_t) par des fonctions paramétriques simples
- une estimation des paramètres par des méthodes de type MCO (OLS, *Ordinary Least Squares*)

A l'heure actuelle, cette approche ne suscite plus un grand intérêt car elle ne permet pas de prendre en compte les propriétés stochastiques des chroniques économiques.

Les méthodes stochastiques

Les méthodes stochastiques sont fondées sur

- la spécification de modèles ARIMA pour modéliser les composantes inobservables (UCARIMA, *Unobserved Component ARIMA*)
- l'utilisation de la théorie de l'extraction de signal

Le problème de l'extraction de signal est d'estimer le signal S_t dans les observations

$$X_t = S_t + N_t$$

L'extraction de signal est utilisée en désaisonnalisation en identifiant :

- S_t : élément saisonnier et
- N_t : élément non saisonnier (le « bruit »)

et en modélisant convenablement X_t, S_t et N_t

On distingue deux approches :

- l'approche des séries temporelles structurelles (STS, *Structural Time Series*)
- l'approche basée sur le modèle ARIMA (AMB, *ARIMA Model-Based*).

L'approche basée sur le modèle ARIMA (AMB)

L'approche AMB a pour objet de modéliser la série observée à partir d'un modèle ARIMA saisonnier (SARIMA) de manière à déduire les éléments à partir de la structure du modèle en utilisant les estimations spectrales

Puisque les composantes sont inobservables et afin d'obtenir une décomposition unique à partir du modèle ARIMA général ajusté à la série d'origine, Hillmer et Tiao (1982) proposent ce qu'ils appellent la décomposition canonique.
Elle a les propriétés, entre autres,

- de maximiser la variance de la composante irrégulière
- de minimiser celle de la composante saisonnière

Les modèles ARIMA sont très sensibles aux points atypiques et ne peuvent pas estimer correctement les éléments déterministes.

Par conséquent, d'autres développements sont réalisés en combinant des modèles de régression avec des variables dichotomiques et des erreurs ARIMA.

C'est dans cette optique que le logiciel TRAMO-SEATS a été proposé.

Le logiciel TRAMO-SEATS

Gómez et Maravall (1997) ont développé TRAMO-SEATS à la Banque d'Espagne

- TRAMO : *Time Series Regression with ARIMA Noise, Missing Observations, and Outliers*
- SEATS : *Signal Extraction in Arima Time Series*

L'application de SEATS suppose que la série soit stationnaire ou qu'elle puisse être stationnarisée afin d'être modélisée en utilisant des modèles SARIMA globaux.

Cette hypothèse de stationnarité est rarement justifiée pour les séries temporelles économiques.

Par conséquent, de la même manière que RegARIMA, le pré-programme TRAMO est utilisé pour

- rendre la série stationnaire
- supprimer les effets déterministes (effets de calendrier, outliers, . . .)
- permettre l'estimation, la prévision et l'interpolation de modèles de régression avec des observations manquantes et des erreurs ARIMA, en présence de plusieurs types possibles de points atypiques

A partir de la série linéarisée par TRAMO, SEATS commence par ajuster un modèle SARIMA.

Ce modèle est déterminé par une procédure automatique d'identification de modèles, identique à RegARIMA, fondée sur

- des contraintes concernant les ordres des polynômes (saisonniers et non saisonniers)
- l'algorithme de Hannan-Rissanen pour les polynômes non saisonniers
- le critère d'information Bayésien (*Bayesian Information Criteria*, BIC)

Ensuite, SEATS utilise la méthode AMB pour décomposer la série en composantes de tendance-cycle, saisonnier et irrégulier.

Définition

Un **modèle SARIMA**$(p,d,q)(P,D,Q)_s$ est définit de la manière suivante (Box et Jenkins, 1978)

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D y_t = \theta_q(B)\Theta_Q(B^s)a_t$$

- s : périodicité de la composante saisonnière
- B : opérateur de retard tel que $By_t = y_{t-1}$
- $\phi_p(B) = (1-\phi_1 B - \cdots - \phi_p B^p)$: polynôme non saisonnier autorégressif (AR) d'ordre p
- $\Phi_P(B^s) = (1-\Phi_1 B^s - \cdots - \Phi_p B^{Ps})$: polynôme saisonnier autorégressif (SAR) d'ordre P
- $\theta_q(B) = (1-\theta_1 B - \cdots - \theta_q B^q)$: polynôme non saisonnier de moyenne mobile (MA) d'ordre q
- $\Theta_Q(B^s) = (1-\Theta_1 B^s - \cdots - \Theta_Q B^{Qs})$: polynôme saisonnier de moyenne mobile (SMA) d'ordre Q
- $\Delta^d = (1-B)^d$: opérateur de différence (non saisonnière) d'ordre d-ième
- $\Delta_s^D = (1-B^s)^D$: opérateur de différenciation saisonnière d'ordre D-ième
- $\delta(B) = \Delta^d \Delta_s^D$: opérateur de différenciation (filtres)
- a_t : un **processus de bruit blanc**, normalement distribué de moyenne zéro et de variance σ_a^2 : $a_t \sim N(0,\sigma_a^2)$

Le modèle SARIMA complet peut se réécrire de la manière suivante :

$$\Psi(B)y_t = \pi(B)a_t + c$$

- $\Psi(B) = \phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D$
- $\pi(B) = \theta_q(B)\Theta_Q(B^s)$
- $\Psi(B)$ et $\pi(B)$: polynômes respectivement autorégressif et de moyenne mobile satisfaisant les conditions de stationnarité et d'inversibilité
- $c = \Psi(B)\overline{y}$ avec $\overline{y}$ la moyenne de y_t

Le modèle pour la série différenciée z_t peut être exprimé de la manière suivante

$$\phi_p(B)\Phi_P(B^s)(z_t - \bar{z}) = \theta(B)a_t$$

- $z_t = \delta(B)y_t$: processus stationnaire
- $\bar{z}$: moyenne de z_t

Le polynôme $\Psi(B)$ est factorisé de la manière suivante :

$$\Psi(B) = \Psi_T(B)\Psi_S(B)\Psi_C(B)$$

où $\Psi_T(B)$, $\Psi_S(B)$ et $\Psi_C(B)$ sont des polynômes autorégressifs qui contiennent respectivement les racines des composantes de tendance, saisonnière et cyclique.

On obtient ainsi :

$$y_t = \frac{\pi(B)}{\Psi(B)} a_t = \frac{\pi_T(B)}{\Psi_T(B)} a_{Tt} + \frac{\pi_S(B)}{\Psi_S(B)} a_{St} + \frac{\pi_C(B)}{\Psi_C(B)} a_{Ct} + u_t$$

- u_t : un bruit blanc Gaussien
- les bruits a_{Tt}, a_{St} et a_{Ct} sont non corrélés deux à deux.

Soit $g_j(\omega)$ le spectre de $[\pi(B)/\Psi(B)]a_t$.

Si les spectres de toutes les composantes sont non négatifs, alors la décomposition est dite admissible

Les modèles pour les composantes de tendance, saisonnière et cyclique sont définies respectivement par :

$$\Psi_T(B)T_t = \pi_T(B)a_{Tt}$$

$$\Psi_S(B)S_t = \pi_S(B)a_{St}$$

$$\Psi_C(B)C_t = \pi_C(B)a_{Ct}$$

Il existe une infinité de décomposition admissible à partir d'un modèle SARIMA pour des composantes inobservables

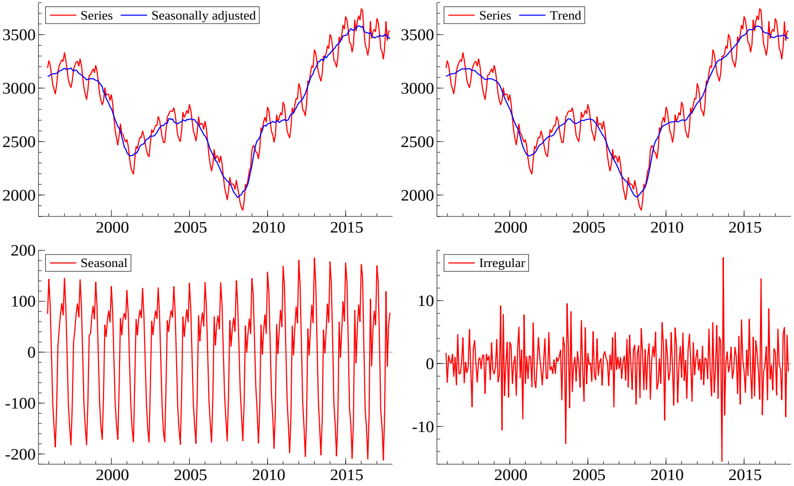
Hillmer et Tiao (1982) proposent ce qu'ils appellent la décomposition canonique :

- mettre tout le bruit blanc dans la composante irrégulière
- maximiser la variance de l'irrégulier
- minimiser la variance de la composante saisonnière

L'application de la décomposition canonique permet d'obtenir une décomposition unique du modèle

Les paramètres des composantes de tendance-cycle (TC_t) et saisonnière (S_t) sont ensuite estimés par les estimateurs des erreurs moyennes quadratiques minimums (MMSE, *Minimum Mean Squared Error*), en utilisant le filtre de Weiner-Kolmogorov.

Figure: Décomposition de la série par TRAMO-SEATS via JDemetra+



L'approche des séries temporelles structurelles (STS)

L'approche STS consiste à spécifier directement des modèles ARIMA pour chaque variable inobservable.

Cette approche s'inscrit dans la continuité des méthodes de régression outre le fait qu'au lieu d'utiliser des modèles déterministes pour estimer chaque composante, elle emploie des modèles stochastiques simples appartenant à la classe ARIMA avec prédominance IMA (*Integrated Moving Average*).

La méthode de décomposition du modèle structurel est en 2 étapes (description est donnée en annexe) :

1. une équation d'observation (quelquefois appelée équation de mesure) composée d'éléments inobservables, cad la tendance-cycle, la saisonnalité et les irréguliers.
   Pour chaque variable inobservable, on suppose un modèle ARIMA très simple qu'on explicite avec des équations dites d'état.
2. Le modèle structurel est mis sous forme d'espace d'état et est souvent estimé par le filtre de Kalman

L'approche STS est implémentée dans des logiciels :

- Le module STAMP disponible dans OxMetrics
- Le package R stats avec la fonction `StrucTS`

Le module STAMP

Koopman et alii (1995, 2010) ont développé le logiciel STAMP (*Structural Time series Analyser, Modeller and Predictor*) à la London School of Economics and Political Science.

STAMP, dans son option par défaut, suggère d'utiliser le modèle structurel fondamental (BSM, *Basic Structural Model*) proposé par Harvey (1981), cad un modèle sans élément exogène ni cycle, pour modéliser les composantes inobservables :

$$y_t = T_t + S_t + I_t$$

La tendance est modélisée comme une marche aléatoire avec dérive, où la dérive (la pente de la tendance) suit également une marche aléatoire.

$$\begin{aligned} T_t &= T_{t-1} + \beta_{t-1} + \eta_t \\ \beta_t &= \beta_{t-1} + \zeta_t \end{aligned}$$

- $\eta_t \sim N(0, \sigma^2_\eta)$ et $\zeta_t \sim N(0, \sigma^2_\zeta)$ sont deux processus de bruit blanc mutuellement non corrélés
- si $\sigma^2_\zeta = 0 \Rightarrow$ processus $I(1)$: $T_t = T_{t-1} + \beta_1 + \eta_t$
- si $\sigma^2_\zeta = 0$ et $\sigma^2_\eta = 0 \Rightarrow$ tendance linéaire : $T_t = T_1 + \beta_1 t$

La composante irrégulière est généralement supposée être un bruit blanc Gaussien de moyenne zéro et de variance σ^2_ε :

$$I_t = \varepsilon_t \sim N(0, \sigma^2_\varepsilon)$$

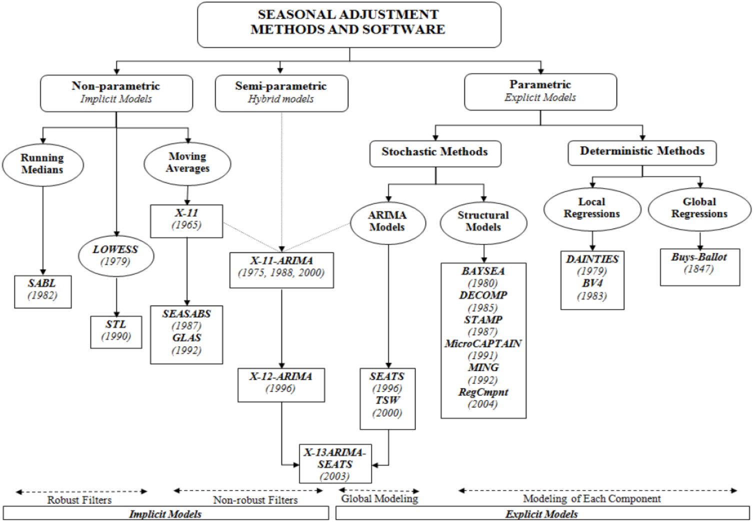
Dans le BSM, la composante saisonnière est un modèle trigonométrique stochastique (Harvey, 1989) :

$$S_t = \sum_{j=1}^{[s/2]} \gamma_{j,t}$$

où chaque $\gamma_{j,t}$ est générée par :

$$\begin{bmatrix} \gamma_{j,t} \\ \gamma^*_{j,t} \end{bmatrix} = \begin{bmatrix} \cos\lambda_j & \sin\lambda_j \\ -\sin\lambda_j & \cos\lambda_j \end{bmatrix} \begin{bmatrix} \gamma_{j,t-1} \\ \gamma^*_{j,t-1} \end{bmatrix} + \begin{bmatrix} \omega_{j,t} \\ \omega^*_{j,t} \end{bmatrix}$$

- $\lambda_j = 2\pi j/s$, $j = 1, \ldots, [s/2]$, s : le nombre d'observations par année, et $t = 1, \ldots, T$.
- les innovations saisonnières $\omega_{j,t}$ et $\omega^*_{j,t}$ sont mutuellement non corrélées, de moyenne zéro et de variance commune σ^2_ω


SEASONAL ADJUSTMENT METHODS AND SOFTWARE
Non-parametric Implicit Models
Semi-parametric Hybrid models
Parametric Explicit Models
Running Medians
Moving Averages
Stochastic Methods
Deterministic Methods
X-11 (1965)
ARIMA Models
Structural Models
Local Regressions
Global Regressions
LOWESS (1979)
X-11-ARIMA (1975, 1988, 2000)
BAYSEA (1980) DECOMP (1985) STAMP (1987) MicroCAPTAIN (1991) MING (1992) RegCmpnt (2004)
DAINTIES (1979) BV4 (1983)
Buys-Ballot (1847)
SABL (1982)
STL (1990)
SEASABS (1987) GLAS (1992)
X-12-ARIMA (1996)
SEATS (1996) TSW (2000)
X-13ARIMA-SEATS (2003)
Robust Filters
Non-robust Filters
Global Modeling
Modeling of Each Component
Implicit Models
Explicit Models

La détection des points atypiques

La présence de point atypiques (*outliers*) de nature transitoire ou permanent pose de nombreux problèmes d'estimation pour les méthodes de désaisonnalisation mais également pour les méthodes d'estimation (MCO), l'estimation des paramètres des modèles ARIMA . . .

Les méthodes de désaisonnalisation actuelle proposent un pré-traitement de ces valeurs avant d'estimer les différentes composantes : TRAMO et RegARIMA.

Ces méthodes sont fondées sur les travaux de Box et Tiao (1975) sur l'analyse d'intervention.

Considérons une série temporelle univariée X_t décrite par un modèle ARIMA (p, d, q)

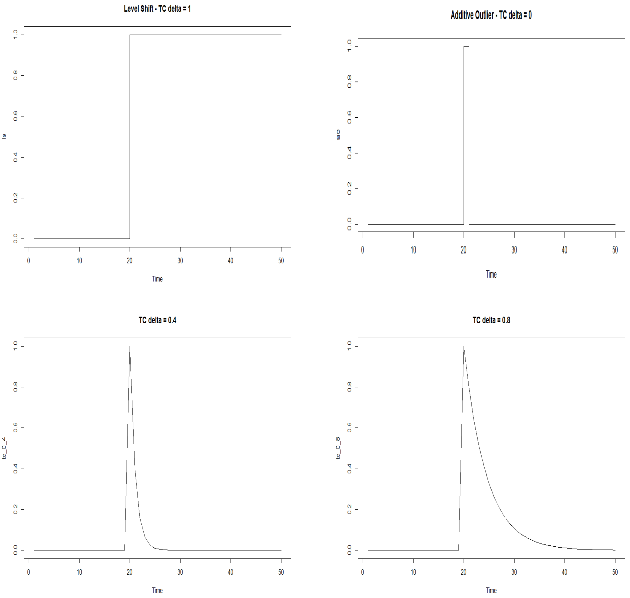
$$
\begin{aligned}
X_t &= X_t^* + f(t) && (1)\\
\alpha(L)\phi(L)X_t^* &= \theta(L)a_t \qquad t = 1, \ldots, T && (2)\\
f_t &= \sum_I \omega_i \xi_i(L) I_t(\tau) \qquad i = 1, \ldots, 4 && (3)
\end{aligned}
$$

- X_t^* : modèle ARIMA(p, d, q)
- $f(t)$: processus contenant les points atypiques
- $L = B$: opérateur retard
- a_t : processus de bruit blanc $a_t \sim N(0, \sigma_a^2)$
- $\alpha(L)$, $\phi(L)$, $\theta(L)$: polynômes d'ordre respectifs d, p et q
- $\xi_i(L)$: polynôme caractérisant le point atypique qui se produit au temps $t = \tau$
- ω_i : impact sur la série
- $I_t(\tau)$: variable indicatrice (*dummy*) prenant la valeur de 1 au temps $t = \tau$ et 0 sinon.

Il existe principalement 5 types de points atypiques, définis de la manière suivante :

- *Additive Outliers* (AO) : affectent une seule observation à un moment du temps dans la série temporelle
$$\xi_{AO} = 1$$
- *Level Shifts* (LS) : effet permanent sur le niveau de la série
$$\xi_{LS} = 1/(1-L)$$
- *Temporary Changes* (TC) : effet temporaire la série qui retourne à son niveau précédent de manière exponentielle. Leur vitesse de retour dépend du paramètre δ dans le polynôme
$$\xi_{TC} = 1/(1-\delta L) \qquad \text{avec } 0 < \delta < 1$$
- *Innovative Outliers* (IO) : impact sur les innovations du modèle
$$\xi_{IO} = \theta(L)/\alpha(L)\phi(L)$$
- *Seasonal Outliers* (SO) (ou *seasonal level shifts*, SLS) : effet permanent mais seulement sur une période particulière (ex. tous les mois de décembre)
$$\xi_{SO} = 1/\Delta_S$$

Les points atypiques



Les points atypiques

Les points atypiques sont identifiés lors d'une procédure de détection séquentielle, comprenant une itération interne et une autre externe

- dans l'itération externe, en supposant qu'il n'y a pas des points atypiques, un modèle ARIMA (p, d, q) est estimé, donnant ainsi les résidus
- les résultats de l'itération externe sont alors utilisés dans l'itération interne pour identifier les points atypiques
- les statistiques de test pour les 5 types de points atypiques, $\hat{\tau}_i(\tau)$, $i = 1, \ldots, 5$, sont calculées pour chaque observation
- si $\hat{\tau}_{max} = max|\hat{\tau}_i(\tau)| > VC \Rightarrow$ un point atypique est identifié au temps $t = \tau$

Quand un outlier est détecté au temps $t = \tau_1$ alors la série est corrigée de la manière suivante :

$$X_t^* = X_t - f(t)_{i^*}$$

La procédure est répétée jusqu'à plus aucun point atypique ne soit détecté

Ces détections et corrections de points atypiques sont mises en œuvre dans les logiciels TRAMO et RegARIMA (pré-ajustement de la série à modéliser)

(détails de la procédure en annexe)

Figure: Détection des outliers pour la série de la production industrielle par TRAMO et RegARIMA

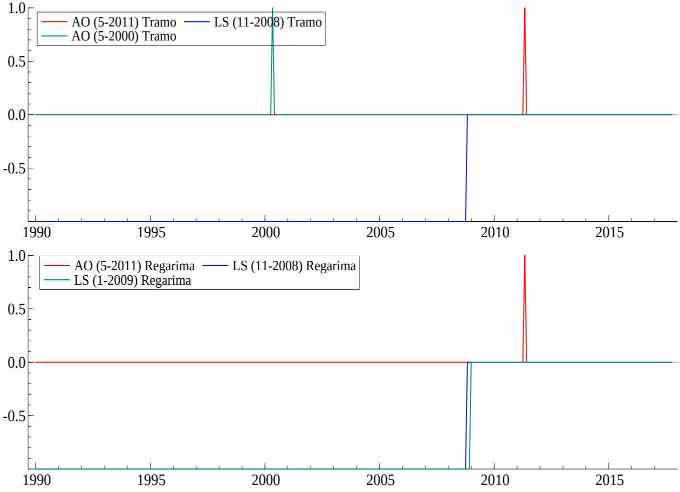
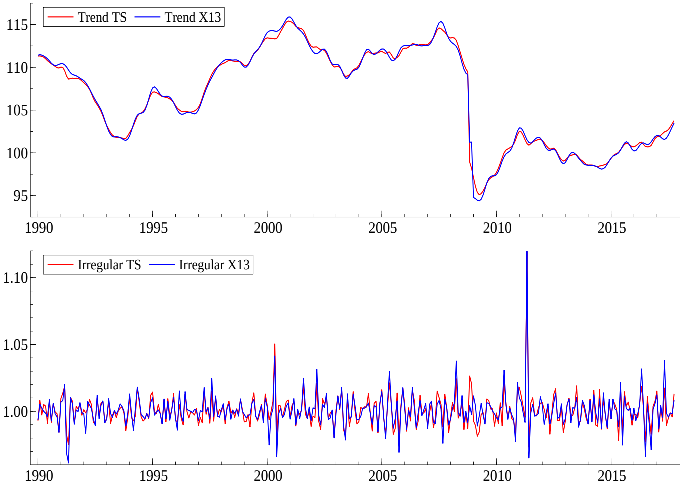


Figure: Outliers dans les composantes de tendance et irrégulière par TRAMO-SEATS et X13-RegARIMA



Le package R tsoutliers permet de détecter les points atypiques en séries temporelles à partir de l'approche de Chen et Liu (1993), notamment les IO, A), LS, TC et SLS.

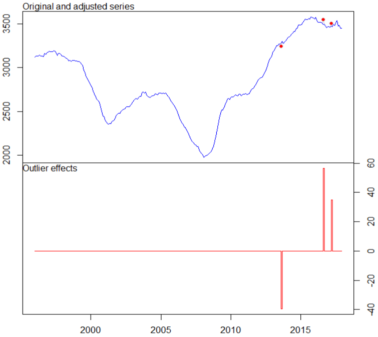
```
library(tsoutliers)
y <- read.table("c://R//data//demandcvs.txt")
yy <- ts(data = y, start=c(1996,01),frequency=12)
tso(yy)
fit <- tso(yy)
plot(fit)
adj <- fit$yadj
write(t(adj),file="adjusted-series.out",ncolumn=1,append=FALSE)
```

Le package R seasonal permet de détecter les points atypiques à partir du module RegARIMA de la méthode X13-ARIMA-SEATS, basée sur la méthode améliorée de Chen et Liu (1993) par Gómez et Maravall (2001).

```
library(seasonal)
y <- read.table("c://R//data//demandcvs.txt")
yy <- ts(data = y, start=c(1996,01),frequency=12)
seasX <- seas(yy)
# Linearized series or outlier-adjusted series: series.outlieradjorig or a19
adjseries = series(seasX, "a19")
```

Package R tsoutliers avec la fonction `tso`

```
Series: yy
Regression with ARIMA(1,1,2)(1,0,1)[12] errors

Coefficients:
         ar1      ma1     ma2    sar1     sma1     AO212    AO248    AO255
      0.9521  -0.7603  0.1996  0.5771  -0.8533  -39.3654  56.4502  35.0059
s.e.  0.0237   0.0649  0.0606  0.1233   0.0877    9.0728   9.8095   9.5972

sigma^2 estimated as 232.2:  log likelihood=-1087.98
AIC=2193.97   AICc=2194.68   BIC=2226.12

Outliers:
  type ind    time coefhat  tstat
1   AO 212 2013:08  -39.37 -4.339
2   AO 248 2016:08   56.45  5.755
3   AO 255 2017:03   35.01  3.647
```



Le package R forecast permet de détecter et corriger les points atypiques avec la fonction `tsclean` :

- cette méthode est basée sur décomposition à partir de la méthode STL
- la composante irrégulière (résidus) permet de détecter les outliers : $\pm 2(Q_{0.9} - Q_{0.1})$
- les outliers détectés sont remplacés par interpolation linéaire pour les séries non saisonnières (ou désaisonnalisées par la méthode STL)

```
library(forecast)
y <- read.table("c://R//data//demandcvs.txt")
yy <- ts(data = y, start=c(1996,01),frequency=12)
tsoutliers(yy)
yyc <- tsclean(yy)
plot(yyc)
```

Les logiciels de désaisonnalisation

Le logiciel TRAMO-SEATS est disponible sur le site de la Banque d'Espagne mais également dans des logiciels libres (Gretl) et commerciaux (EViews, SAS)

- BdE : https://www.bde.es/bde/en/secciones/servicios/Profesionales/Programas_estadi/Programas.html

Le logiciel X-13ARIMA-SEATS est la dernière version du US Bureau of the Census, incorporant la méthode SEATS dans le logiciel de désaisonnalisation X-12-ARIMA. Il est également disponible dans des logiciels libres (Gretl, R) et commerciaux (EViews, SAS, Matlab)

- Census : https://www.census.gov/srd/www/x13as/
- Interface R : https://cran.r-project.org/web/packages/seasonal/

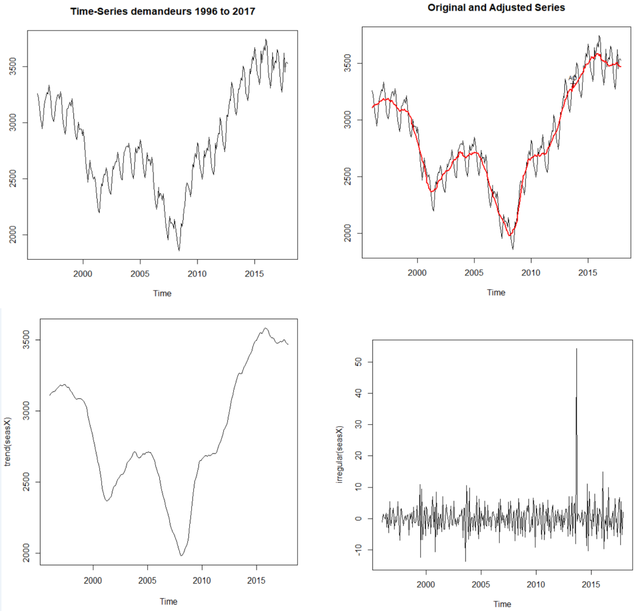
Le logiciel JDemetra+ est une interface proposée par Eurostat et la Banque Centrale de Belgique, pour l'utilisation des méthodes TRAMO-SEATS et X-13ARIMA-SEATS

- Eurostat : http://ec.europa.eu/eurostat/web/ess/-/jdemetra-officially-recommended-as-software-for-the-seasonal-adjustment-of-official-statistics

Le package seasonal avec la fonction seas permet de décomposer la série à partir de X13ARIMA-SEATS

```
library(season)
# By default, seas calls the SEATS adjustment procedure
# The X11 adjustment procedure: seasX11 <- seas(yy, x11 = "")
seasX <- seas(yy)
# Final CVS
final(seasX)
plot(seasX)
plot(irregular(seasX))
plot(trend(seasX))
summary(seasX)
```

```
Call:
seas(x = yy)

Coefficients:
                    Estimate Std. Error z value Pr(>|z|)
AO2013.Sep          54.32941   12.59425   4.314 1.60e-05 ***
MA-Nonseasonal-01    0.69654    0.04526  15.389  < 2e-16 ***
MA-Seasonal-12       0.43662    0.05987   7.293 3.04e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

SEATS adj.  ARIMA: (0 2 1)(0 1 1)  Obs.: 264  Transform: none
AICc:  2320, BIC:  2334  QS (no seasonality in final):    0
Box-Ljung (no autocorr.): 39.04 * Shapiro (normality): 0.9792 ***
Messages generated by X-13:
Warnings:
- At least one visually significant trading day peak has been found in one or more of
  the estimated spectra.
```


Time-Series demandeurs 1996 to 2017
2000
2500
3000
3500
2000
2005
2010
2015
Time
Original and Adjusted Series
AO
trend(seasX)
irregular(seasX)
-10
0
10
20
30
40
50

Forecasting on 3 years

```
series(seasX, "forecast.forecasts")
```

The udg function provides access to a large number of diagnostical statistics

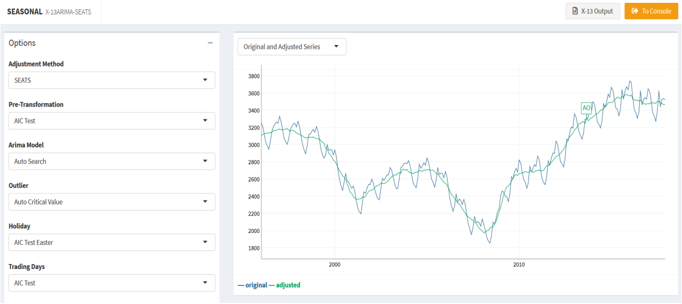
```
udg(seasX, "x13mdl")
```

The out function shows the content of the main output using the HTML version of X-13

```
out(seasX)
```

The view function is a graphical tool for choosing a seasonal adjustment model

```
install.packages("seasonalview")
view(seasX)
```



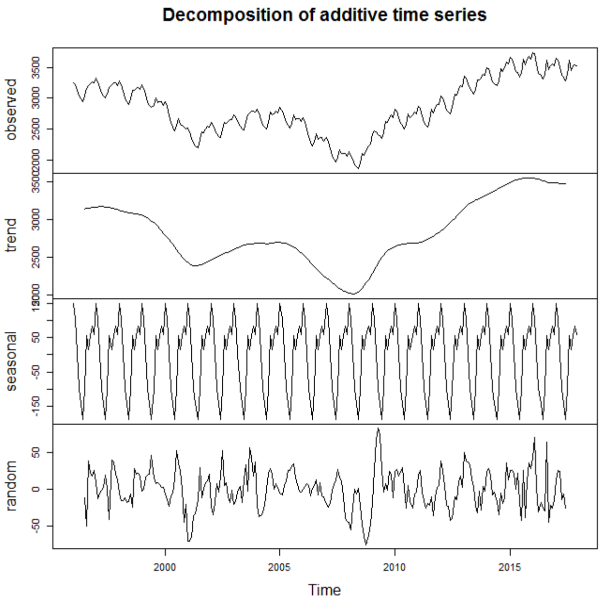
Analyse de la désaisonnalisation

A l'inverse de nombreux packages R le logiciel JDemetra+ et le package proposent un grand nombre de diagnostiques pour évaluer la qualité de la désaisonnalisation et de la décomposition de la série originale par TRAMO-SEATS et X13-ARIMA

- Pre-processing (Tramo ou Regarima)
  - Forecast
  - Regressors (outliers, trading days, easter effect . . .)
  - Arima
  - Pre-adjustment series
  - Residuals
  - Likelihood
- Decomposition
- Benchmarking
- Diagnostics
  - Seasonality tests
  - Spectral analysis
  - Sliding spans (rolling-window estimation)
  - Revisions history
  - Model stability

Le package stats avec la fonction `decompose` permet de décomposer la série à partir de **moyennes mobiles**, avec des schémas additif ou multiplicatif

```
y <- read.table("c://R//data//demandbrut.txt")
yy <- ts(data = y, start=c(1996,01),frequency=12)
fit <- decompose(yy)
plot(fit)
show(yy)
```

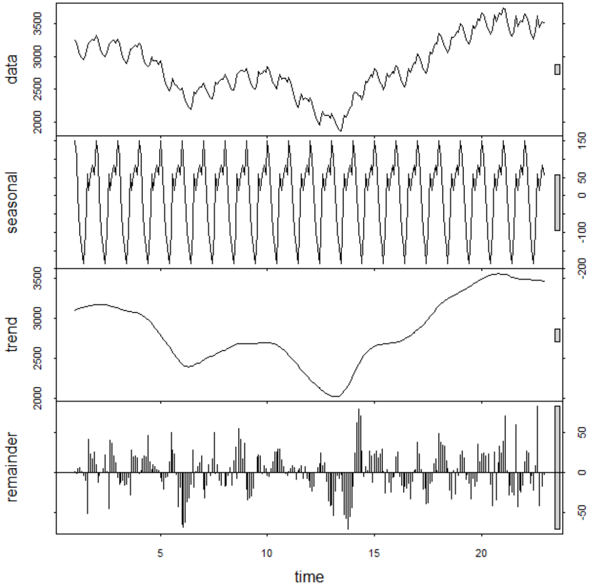

Decomposition of additive time series
observed
trend
seasonal
random
2000
2500
3000
3500
2000
2500
3000
3500
-150
-50
50
150
-50
0
50
2000
2005
2010
2015
Time

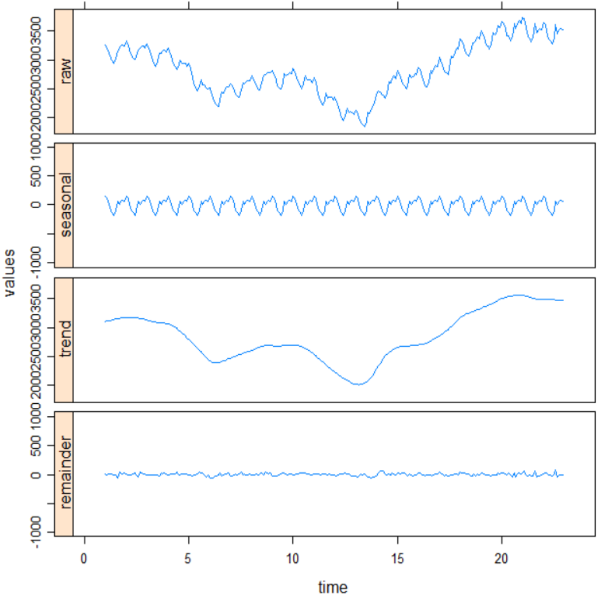
Le package stats avec la fonction `stl` et le package stlplus avec la fonction `stlplus` décomposent la série à partir de l'**approche STL** (*Seasonal-Trend decomposition based on Loess*) proposé par Cleveland et alii (1990)

```
y <- read.table("c://R//data//demandbrut.txt")
yy <- ts(y[1:264,1],frequency=12)
# NB: utile quand on a un message d'erreur avec ts() avec "univarié"
decomp <- stl(yy, s.window="per")
plot(decomp)
show(plot)
```

```
library(stlplus)
decomp <- stlplus(yy, s.window="per")
plot(decomp)
show(plot)
plot_trend(decomp)
tendance = trend(decomp)
```

Les packages R de désaisonnalisation

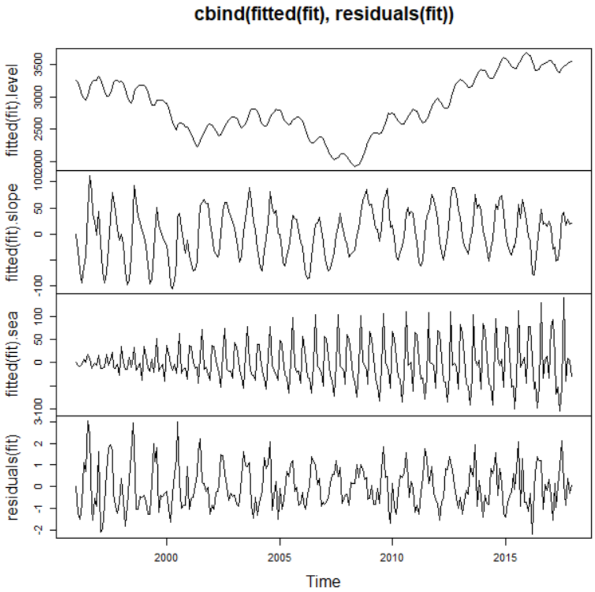



raw
seasonal
trend
remainder
values
time
0
5
10
15
20
2000
2500
3000
3500
-1000
0
500
1000

Le package stats avec la fonction `StrucTS` décompose la série à partir de l'**approche structurelle des séries temporelles (STS)** (*Structural Time Series*) avec le modèle BSM

```
y <- read.table("c://R//data//demandbrut.txt")
yy <- ts(y[1:264,1],frequency=12)
fit <- StructTS(yy)
plot(cbind(fitted(fit), residuals(fit)))
show(fit)
```

```
Call:
StructTS(x = yy)

Variances:
level    slope    seas    epsilon
48.51    803.04   83.47   0.00
```


cbind(fitted(fit), residuals(fit))
fitted(fit).level
2000
2500
3000
3500
fitted(fit).slope
-100
0
50
100
fitted(fit).sea
-100
0
50
100
residuals(fit)
-2
-1
0
1
2
3
2000
2005
2010
2015
Time

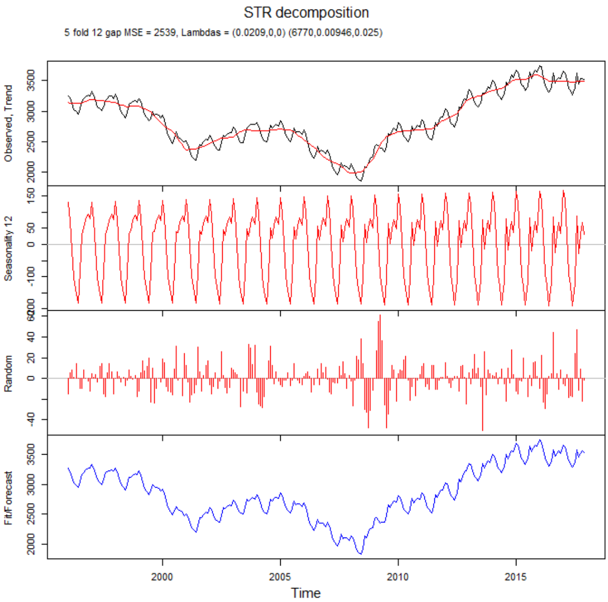
Le package stR avec la fonction `AutoSTR` décompose la série en utilisant une **procédure basée sur la régression** proposée par Dokumentov et Hyndman (2015) : STR (*Seasonal-Trend Decomposition Procedure Based on Regression*) :

$$y_t = T_t + S_t + I_t$$

- Smooth trend : $\Delta^2 T_t = T_{t+1} - 2T_t + T_{t-1}$
- S_t : composante saisonnière pouvant être fractionnaire, multiple, complexe
- $I_t \sim \mathcal{N}(0, \sigma_I^2)$

```
library(stR)
y <- read.table("c://R//data//demandbrut.txt")
yy <- ts(data = y, start=c(1996,01),frequency=12)
decomp <- AutoSTR(yy)
plot(decomp)
show(decomp)
```

Référence : Dokumentov et Hyndman (2015), STR: A Seasonal-Trend Decomposition Procedure Based on Regression


STR decomposition
5 fold 12 gap MSE = 2539, Lambdas = (0.0209,0,0) (6770,0.00946,0.025)
Observed, Trend
2000
2500
3000
3500
Seasonality 12
-200
-100
0
50
150
Random
-40
0
20
40
60
Fit/Forecast
2000
2500
3000
3500
2000
2005
2010
2015
Time

Le package forecast avec la fonction `tbats` décompose la série à partir d'un **modèle d'espace d'état avec lissage exponentiel et erreurs ARMA** proposé par De Livera, Hyndman et Snyder (2011) (cf. chapitre 3).

```
library(forecast)
y <- read.table("c://R//data//demandbrut.txt")
yy <- ts(data = y, start=c(1996,01),frequency=12)
decomp <- tbats(yy)
plot(decomp)
comp <- tbats.components(decomp)
```

Référence : De Livera, Hyndman et Snyder (2011), Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing

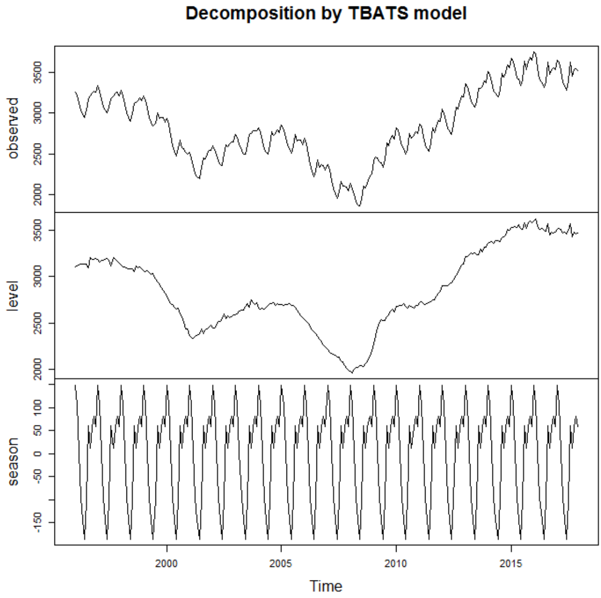

Decomposition by TBATS model
observed
level
season
3500
3000
2500
2000
100
50
0
-50
-150
2000
2005
2010
2015
Time

Figure: Séries CVS par R packages, Tramo-Seats, X13ARIMA et Stamp

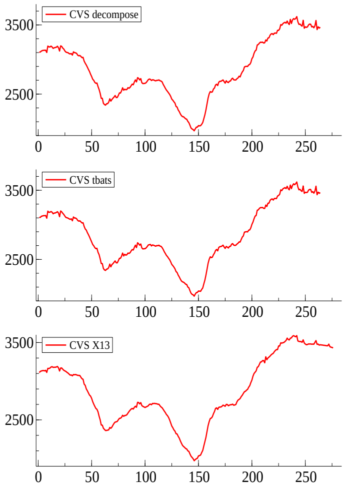


Figure: Tendance-cycle par R packages, Tramo-Seats, X13ARIMA et Stamp

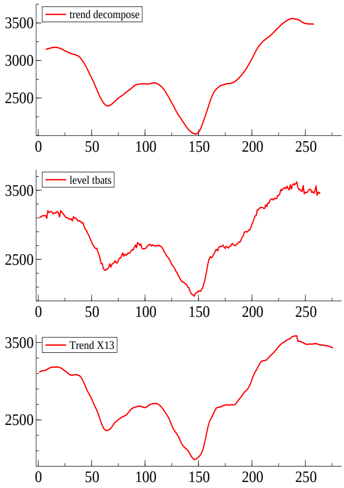


Figure: Composante saisonnière par R packages, Tramo-Seats, X13ARIMA et Stamp

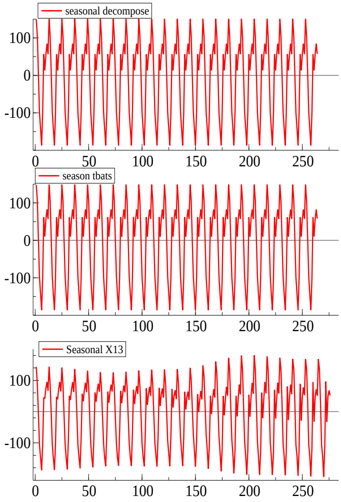
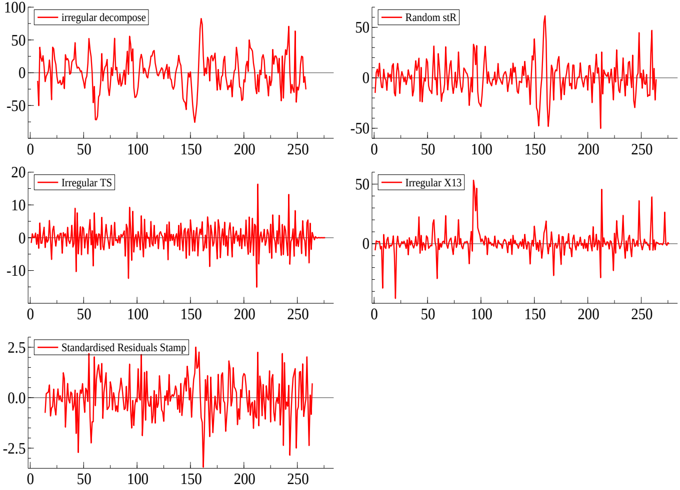


Figure: Composante irrégulière par R packages, Tramo-Seats, X13ARIMA et Stamp



Les prévisions

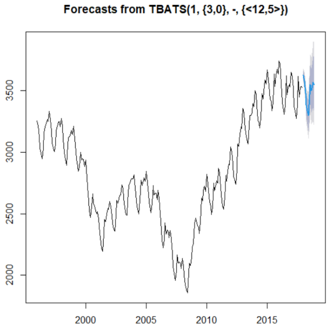
Le package R seasonal permet de faire des **prévisions** sur 3 années (par défaut) à partir de la méthode X13-ARIMA-SEATS

```
library(seasonal)
y <- read.table("c://R//data//demandbrut.txt")
yy <- ts(data = y, start=c(1996,01),frequency=12)
seasX <- seas(yy)
fore = series(seasX, "forecast.forecasts")
# or
fore = series(seasX, "fct")
# one-ahead forecast
fore[1]
```

	forecast	lowerci	upperci
Jan 2018	3636.420	3588.4369	3684.403
Feb 2018	3603.655	3524.8252	3682.485
Mar 2018	3480.880	3370.6104	3591.149
Apr 2018	3339.444	3196.0490	3482.838
May 2018	3300.944	3122.4886	3479.399
Jun 2018	3233.297	3017.8063	3448.787
Jul 2018	3325.934	3071.4632	3580.406
Aug 2018	3551.805	3256.4597	3847.150
Sep 2018	3396.800	3058.7456	3734.853
Oct 2018	3468.845	3086.3112	3851.380
Nov 2018	3486.330	3057.5980	3915.062
Dec 2018	3462.927	2986.3357	3939.519

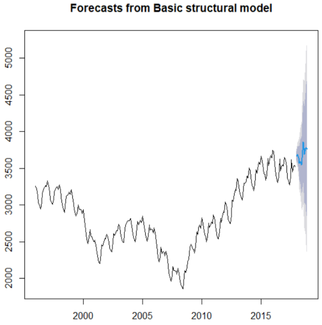
Parmi les R packages seul le package forecast permet de faire des **prévisions** avec la fonction forecast mais pour différentes approches comme avec la méthode TBATS

```
library(forecast)
y <- read.table("c://R//data//demandbrut.txt")
yy <- ts(data = y, start=c(1996,01),frequency=12)
fittbats <- tbats(yy)
# 6-ahead forecasts
prevtbats <- forecast(fittbats,12)
plot(prevtbats)
```



Le package forecast permet de faire des **prévisions** à partir de la méthode STS avec le modèle BSM

```
library(forecast)
y <- read.table("c://R//data//demandbrut.txt")
yy <- ts(data = y, start=c(1996,01),frequency=12)
fitsts <- StructTS(yy)
# 6-ahead forecasts
prevsts <- forecast(fitsts,12)
plot(prevsts)
```



Le package forecast permet de faire des **prévisions** à partir de la méthode STL

```
library(forecast)
y <- read.table("c://R//data//demandbrut.txt")
yy <- ts(data = y, start=c(1996,01),frequency=12)
fitstl <- stlm(yy)
# 6-ahead forecasts
prevstl <- forecast(fitstl,12)
plot(prevstl)
# or
yy <- ts(y[1:264,1],frequency=12)
prevstl = stlf(yy,12)
```

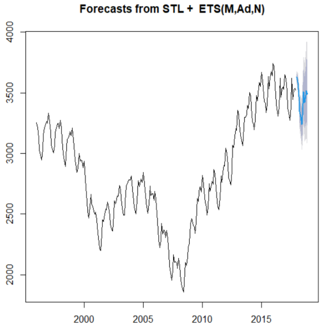


Figure: Comparaison des prévisions par X13-ARIMA-SEAT, STL, STS et TBATS

Prévisions	janv-18	févr-18	mars-18	avr-18
X13-ARIMA-SEATS	3 636.42	3 603.66	3 480.88	3 339.44
STL	3 628.69	3 587.48	3 470.21	3 345.75
STS	3 665.72	3 685.35	3 639.58	3 569.22
TBATS	3 624.80	3 578.57	3 483.59	3 385.26
réalisée	**3 636.60**	**3 565.90**	**3 441.20**	**3 324.90**
Erreurs de prévision				
X13-ARIMA-SEATS	0.18	-37.76	-39.68	-14.54
STL	7.91	-21.58	-29.01	-20.85
STS	-29.12	-119.45	-198.38	-244.32
TBATS	11.80	-12.67	-42.39	-60.36

Figure: Evolution des prévisions par X13-ARIMA-SEAT, STL, STS et TBATS

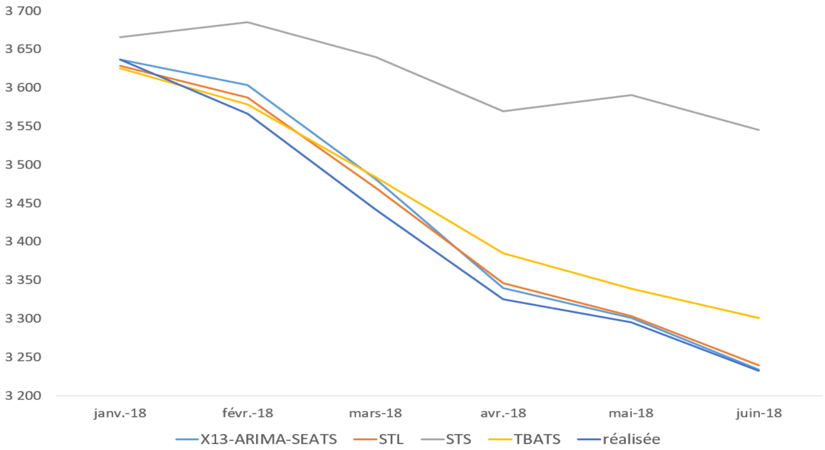


Figure: Comparaison des estimations révisées de la série CVS par Tramo-Seats