

ANALYSE DES SÉRIES TEMPORELLES AVEC R



3 - Décomposition d'une série temporelle

ALAIN QUARTIER-LA-TENTE
Insee

1. Schéma de décomposition
2. STL
3. X-13-ARIMA et TRAMO-SEATS
4. Conclusion

Objectifs de cette séquence

Présenter les principales méthodes de décomposition

Questions de positionnement

Pourquoi désaisonnaliser une série ?

Quelles sont les principales méthodes décomposition ?

Quelles sont les différents schémas de décomposition ?

Sommaire

1. Schéma de décomposition
2. STL
3. X-13-ARIMA et TRAMO-SEATS
4. Conclusion

Rappels sur les différentes composantes

- Tendance-cycle : évolution de moyen/long terme de la série (TC_t)
➔ tendance et cycle estimés conjointement : évite problèmes de définition (limite entre cycles courts et longs ?) et difficultés sur séries courtes
- Saisonnalité et autres effets de calendrier (jours ouvrables, fêtes mobiles, etc.) S_t
- Irrégulier : composante d'erreur (idéalement faible) I_t

Rappels sur les différentes composantes

- Tendance-cycle : évolution de moyen/long terme de la série (TC_t)
➔ tendance et cycle estimés conjointement : évite problèmes de définition (limite entre cycles courts et longs ?) et difficultés sur séries courtes
- Saisonnalité et autres effets de calendrier (jours ouvrables, fêtes mobiles, etc.) S_t
- Irrégulier : composante d'erreur (idéalement faible) I_t

Objectif de la désaisonnalisation : enlever S_t pour mieux analyser le signal conjoncturel et faire des comparaisons temporelles/spatiales

Rappels sur les différentes composantes

- Tendance-cycle : évolution de moyen/long terme de la série (TC_t)
➔ tendance et cycle estimés conjointement : évite problèmes de définition (limite entre cycles courts et longs ?) et difficultés sur séries courtes
- Saisonnalité et autres effets de calendrier (jours ouvrables, fêtes mobiles, etc.) S_t
- Irrégulier : composante d'erreur (idéalement faible) I_t

Objectif de la désaisonnalisation : enlever S_t pour mieux analyser le signal conjoncturel et faire des comparaisons temporelles/spatiales

$Y_t = f(S_t, TC_t, I_t)$: quelle relation entre les différentes composantes ?

Les différents schémas de décomposition

- Schéma additif :

$$X_t = TC_t + S_t + I_t$$

- Schéma multiplicatif :

$$X_t = TC_t \times S_t \times I_t$$

- Schéma pseudo-additif (, suppose) :

$$X_t = TC_t \times (S_t + I_t - 1)$$

➔ lorsque certains mois ont des valeurs très faibles, suppose saisonnalité et irrégulier indépendants

- Schéma log-additif :

$$\log(X_t) = \log(TC_t) + \log(S_t) + \log(I_t)$$

Sommaire

1. Schéma de décomposition
- 2. STL**
3. X-13-ARIMA et TRAMO-SEATS
4. Conclusion

Seasonal and Trend decomposition using Loess (STL)

Application successive de Loess (régression locale robuste) et de moyennes mobiles

Avantages :

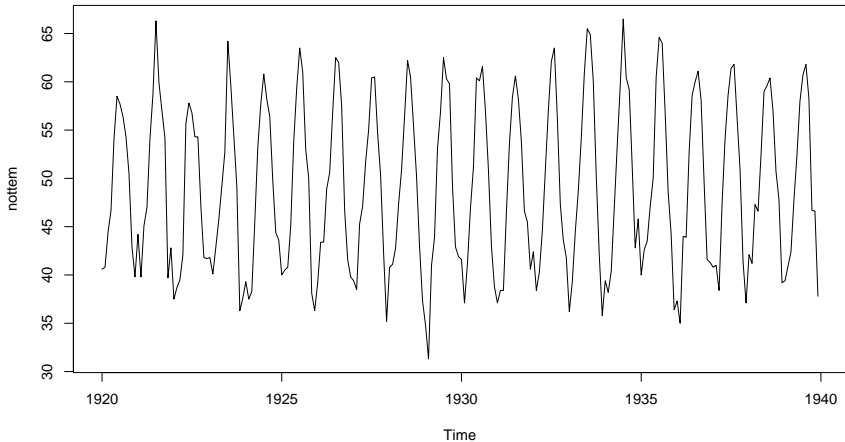
- Applicable sur toutes les fréquences
- Robuste aux points atypiques

Inconvénients :

- Pas de choix automatique du schéma de décomposition
- Pas de correction de jours ouvrables
- Pas de choix automatique des paramètres (longueur des filtres à spécifier)

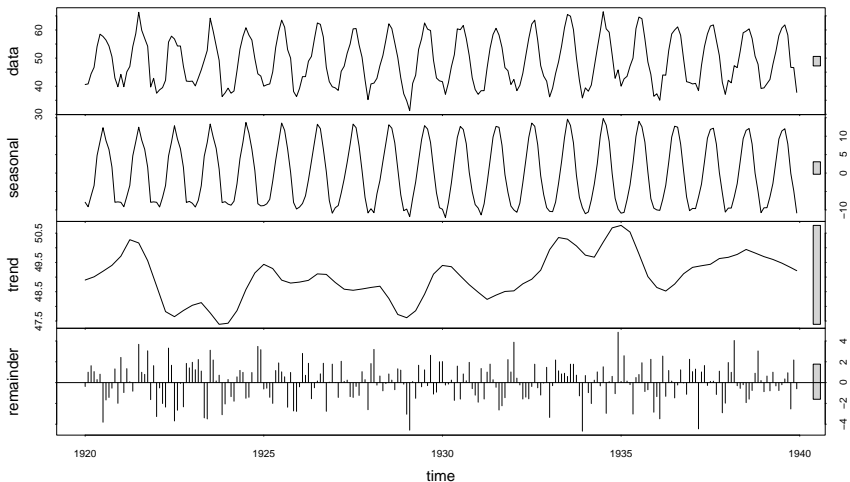
Exemple (1)

```
plot(nottem)
```



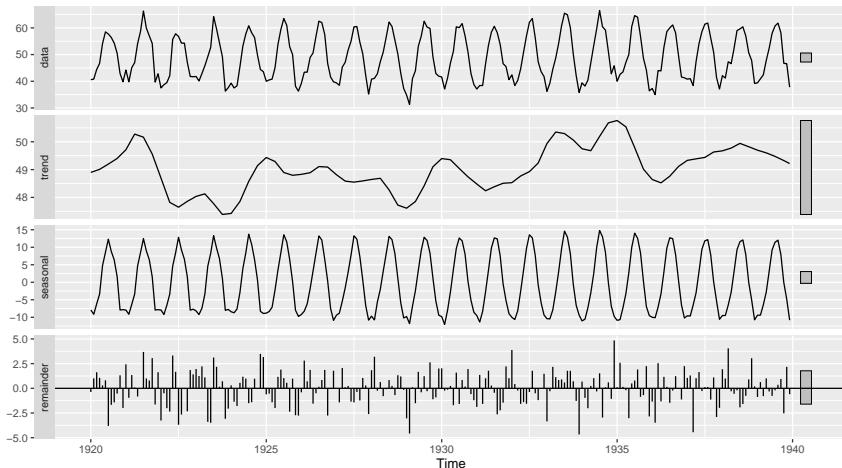
Exemple (2)

```
plot(stl(nottem, s.window = 7))
```



Exemple (3) : autoplot

```
library(forecast)
stl(nottem, s.window = 7) %>% autoplot()
```



Sommaire

1. Schéma de décomposition
2. STL
- 3. X-13-ARIMA et TRAMO-SEATS**
4. Conclusion

Méthodes utilisées dans les INS

Dans les INS, les méthodes principalement utilisées sont X-13-ARIMA et TRAMO-SEATS :

- Méthodes recommandées par Eurostat
- Permettent de faire des CJO et de prendre en compte les spécificités calendaires
- Choix automatique du schéma de décomposition (additive ou multiplicatif)
- Prise en compte de différents types de ruptures
- Applicables que sur des séries au plus mensuelles (travaux en cours pour étendre ces méthodes)

Désaisonnalisation en deux modèles

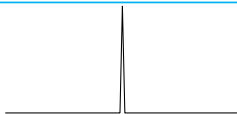
1. Linéarisation de la série : correction des points aberrants, des ruptures et des effets calendriers par un modèle RegARIMA
2. Décomposition de la série linéarisée : par une méthode non-paramétrique (X-11, application successive de moyennes mobiles) ou paramétrique (SEATS)

Les principaux types d'outliers

Choc ponctuel

Additive outlier (AO)

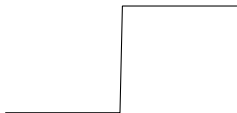
Affecte l'Irrégulier



Changement de niveau

Level Shift (LS)

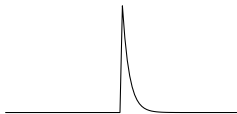
Affecte la Tendence



Changement de niveau transitoire

Transitory Change (TC)

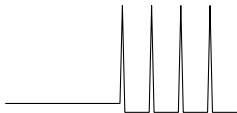
Affecte l'Irrégulier



Rupture de profil saisonnier

Seasonal Outlier (SO)

Affecte la Composante Saisonnière



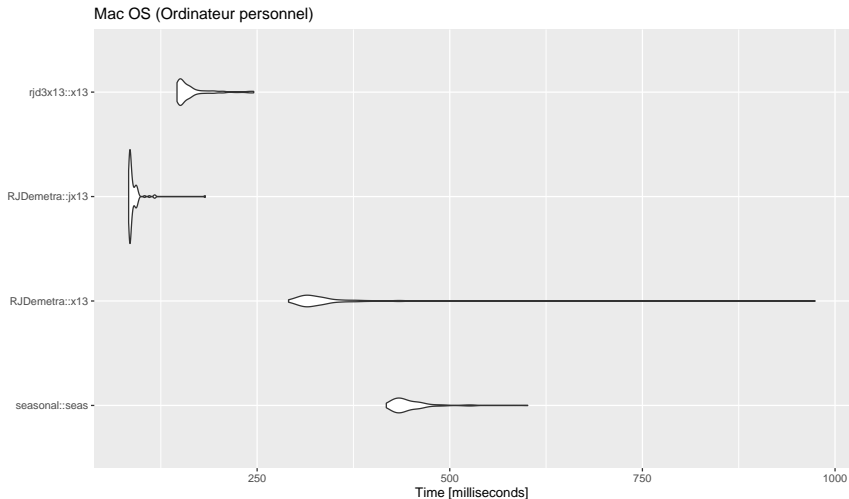
Package à utiliser

Pour X-13ARIMA deux packages peuvent être utilisés :

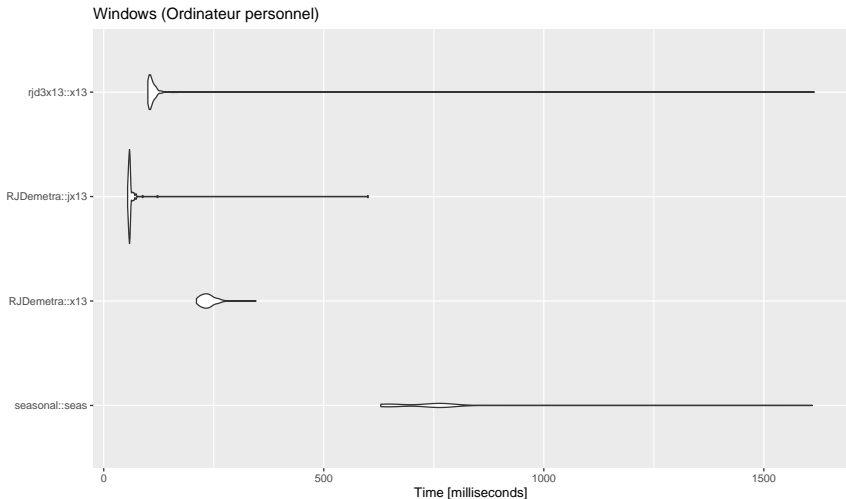
- `seasonal` basé sur les programmes du Census Bureau
- `RJDemetra` basé sur les `JDemetra+` (logiciel recommandé par Eurostat)

TRAMO-SEATS : uniquement implémenté dans `RJDemetra`

Comparaison des performances (1)



Comparaison des performances (2)



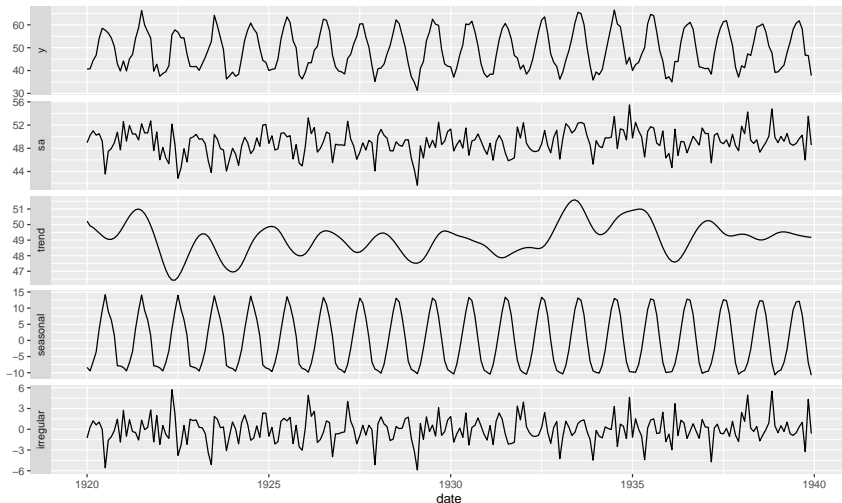
Spécifications de base

Spécification	Transformation	Pre-adjustment for leap-year	Working days	Trading days	Easter effect	Outliers	ARIMA model
RSA0	no	no	no	no	no	no	(0,1,1)(0,1,1)
RSA1	test	no	no	no	no	test	(0,1,1)(0,1,1)
RSA2	test	no	test	no	test	test	(0,1,1)(0,1,1)
RSA3	test	no	no	no	no	test	AMI
RSA4	test	no	test	no	test	test	AMI
RSA5	test	no	no	test	test	test	AMI
RSAfull	test	no	test		test	test	AMI
X11	no	no	no	no	no	no	(0,1,1)(0,1,1)
RSA1	test	no	no	no	no	test	(0,1,1)(0,1,1)
RSA2c	test	test	test	no	test	test	(0,1,1)(0,1,1)
RSA3	test	no	no	no	no	test	AMI
RSA4c	test	test	test	no	test	test	AMI
RSA5	test	test	no	test	test	test	AMI

Paramètre spec dans `RJDemetra::x13()` et `RJDemetra::tramoSeats()`

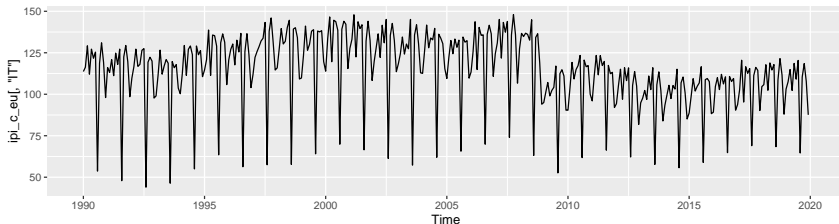
Exemple

```
library(RJDemetra);library(ggdemetra);autoplot(x13(nottem))
```



Exemple (1)

```
autoplot(ipi_c_eu[, "IT"])
```



```
mod = x13(ipi_c_eu[, "IT"])
```

```
# Pre-ajustement : donne des informations sur le schéma de décomposition  
# (log ou non), sur la CJO, les points atypiques et le modèle ARIMA  
mod$regarima
```


Exemple (2)

```
## y = regression model + arima (3, 1, 0, 0, 1, 1)
## Log-transformation: no
## Coefficients:
##           Estimate Std. Error
## Phi(1)      0.57871      0.055
## Phi(2)      0.41329      0.059
## Phi(3)      0.09131      0.055
## BTheta(1) -0.45726      0.048
##
##           Estimate Std. Error
## Monday      0.8197      0.287
## Tuesday     0.8457      0.286
## Wednesday   1.5247      0.286
## Thursday    0.6638      0.285
## Friday      1.2974      0.288
## Saturday   -2.6035      0.285
## Leap year   3.3513      0.890
## Easter [1] -3.0191      0.565
## LS (12-2008) -16.4492      2.149
## LS (8-2009)  9.5156      2.154
##
```

Exemple (3)

```
##
## Residual standard error: 3.023 on 332 degrees of freedom
## Log likelihood = -877.9, aic = 1786 aicc = 1787, bic(corrected for length)
# Informations sur la décomposition (non traitée dans ce cours)
mod$decomposition

## Monitoring and Quality Assessment Statistics:
##      M stats
## M(1)  0.039
## M(2)  0.031
## M(3)  1.018
## M(4)  0.324
## M(5)  0.852
## M(6)  0.142
## M(7)  0.095
## M(8)  0.160
## M(9)  0.072
## M(10) 0.114
## M(11) 0.101
## Q     0.285
## Q-M2  0.316
```

Exemple (4)

```
##
## Final filters:
## Seasonal filter: 3x5
## Trend filter: 13 terms Henderson moving average
```

```
# Séries finales estimées :
mod$final
```

```
## Last observed values
```

##		y	sa	t	s	i
##	Jan 2019	103.1	106.8108	106.5337	-3.71079247	0.27706044
##	Feb 2019	107.5	107.5520	106.4284	-0.05202971	1.12366733
##	Mar 2019	114.9	105.6270	106.2661	9.27304074	-0.63912431
##	Apr 2019	102.3	105.5645	106.0825	-3.26453259	-0.51799031
##	May 2019	118.6	105.8104	105.8478	12.78961029	-0.03744185
##	Jun 2019	109.1	106.4235	105.6324	2.67648686	0.79112529
##	Jul 2019	120.4	104.6128	105.4636	15.78719927	-0.85075490
##	Aug 2019	64.6	105.6598	105.2320	-41.05983503	0.42784829
##	Sep 2019	110.4	105.0552	104.9091	5.34484540	0.14607297
##	Oct 2019	118.5	104.4491	104.5092	14.05094470	-0.06015490
##	Nov 2019	107.8	104.6479	104.1309	3.15211067	0.51695828
##	Dec 2019	87.6	102.7603	103.8287	-15.16034328	-1.06833688

Exemple (5)

```
##
## Forecasts:
##           y_f      sa_f      t_f      s_f      i_f
## Jan 2020 100.60122 103.6937 103.6240 -3.0924959 0.06974302
## Feb 2020 104.88729 104.1412 103.5059  0.7461336 0.63521163
## Mar 2020 115.35780 102.9935 103.4731 12.3643224 -0.47967022
## Apr 2020 100.71674 103.6347 103.4547 -2.9179581 0.17994685
## May 2020 108.88911 103.5382 103.4107  5.3508925 0.12754385
## Jun 2020 112.99300 103.2766 103.3331  9.7163730 -0.05643767
## Jul 2020 119.00943 103.0564 103.2277 15.9529868 -0.17125131
## Aug 2020  58.52488 103.3292 103.1321 -44.8042957 0.19705844
## Sep 2020 112.54290 103.0286 103.0125  9.5143004 0.01610625
## Oct 2020 113.24225 102.7349 102.8926 10.5073243 -0.15771395
## Nov 2020 105.65780 102.9839 102.7820  2.6738591 0.20193998
## Dec 2020  91.03785 102.4851 102.6915 -11.4472191 -0.20644338
```

```
# Diagnostics sur la désaisonnalisation
```

```
mod$diagnostics
```

Exemple (6)

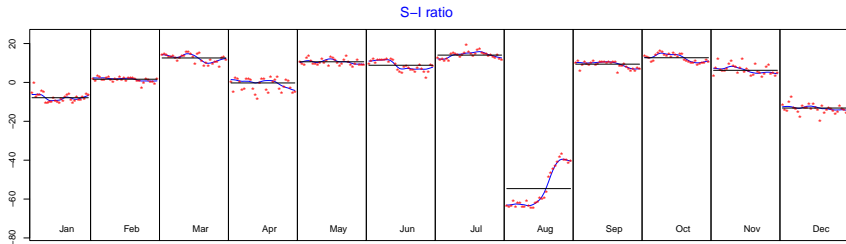
```
## Relative contribution of the components to the stationary
## portion of the variance in the original series,
## after the removal of the long term trend
## Trend computed by Hodrick-Prescott filter (cycle length = 8.0 years)
##           Component
## Cycle           2.275
## Seasonal       89.957
## Irregular       0.740
## TD & Hol.       2.242
## Others          3.886
## Total          99.100
##
## Combined test in the entire series
## Non parametric tests for stable seasonality
##
##                                     P.value
## Kruskal-Wallis test                 0
## Test for the presence of seasonality assuming stability 0
## Evolutive seasonality test          0
##
## Identifiable seasonality present
##
```

Exemple (7)

```
## Residual seasonality tests
##                                     P.value
## qs test on sa                      0.085
## qs test on i                       0.012
## f-test on sa (seasonal dummies)    0.542
## f-test on i (seasonal dummies)    0.254
## Residual seasonality (entire series) 0.438
## Residual seasonality (last 3 years) 0.999
## f-test on sa (td)                  0.671
## f-test on i (td)                   0.997
```

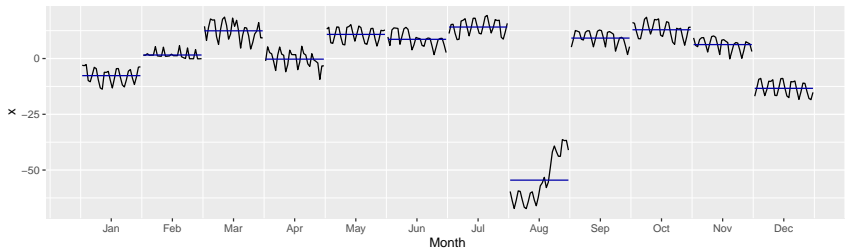
```
# Pour tracer les coefficients saisonniers (hors CJO) et irrégulier
plot(mod$decomposition)
```

Exemple (8)



```
# Coefficients saisonniers + CJO : plus erratiques
ggmonthplot(ggdemetra::seasonal(mod))
```

Exemple (9)



Sommaire

1. Schéma de décomposition
2. STL
3. X-13-ARIMA et TRAMO-SEATS
- 4. Conclusion**

Conclusion

- La désaisonnalisation permet de faire des comparaisons temporelles et spatiales
- Plusieurs schémas de décomposition possibles en fonction des relations entre les composantes :
 - Additif : $X_t = TC_t + S_t + I_t$ saisonnalité et irrégulier indépendant du niveau
 - Multiplicatif : $X_t = TC_t \times S_t \times I_t$ saisonnalité et irrégulier proportionnels au niveau

Bibliographie

Hyndman, R.J., & Athanasopoulos, G. (2018) *Forecasting : principles and practice*, 2nd edition, OTexts : Melbourne, Australia. [OTexts.com/fpp2](https://otexts.com/fpp2). Accessed on nov. 2022.