Chapitre 3
Prévisions de court terme : Méthodes de lissage exponentiel

Olivier DARNÉ

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

Le lissage exponentiel

Le lissage exponentiel (*exponential smoothing*) est une classe de méthodes de lissage de séries chronologiques dont l'objectif est la prévision à court terme.

Introduites par Holt en 1958 ainsi que par Winters en 1960 et popularisées par le livre de Brown en 1963, les méthodes de lissage constituent l'ensemble des techniques empiriques de prévision qui accordent plus ou moins d'importance aux valeurs du passé d'une série temporelle.

Ces méthodes sont fondées sur une hypothèse fondamentale : chaque observation à l'instant t dépend des observations précédentes $(t-1, t-2, \dots)$ et d'une variation accidentelle, et cette dépendance est plus ou moins stable dans le temps.

Les prévisions produites à l'aide de méthodes de lissage exponentiel sont des moyennes pondérées d'observations passées, les poids décroissant de façon exponentielle à mesure que les observations "vieillissent".
En d'autres termes, plus l'observation est récente, plus le poids associé est élevé.

Définition

Les méthodes de **lissage exponentiel non saisonnier** supposent que la chronique est structurée de la manière suivante :

$$X_t = f_t^{(k)} + \varepsilon_t$$

- ε_t : bruit blanc : $E[\varepsilon_t] = 0$; $E[\varepsilon_t^2] = \sigma_\varepsilon^2, \forall t$; $E[\varepsilon_t \varepsilon_{t'}] = 0, \forall t, \forall t', \forall t \neq t'$
- $f_t^{(k)}$: **fonction polynomiale de degré** k dont les paramètres dépendent du temps

Les méthodes de lissage se différencient entre elles selon le degré de $f_t^{(k)}$:

- **lissage exponentiel simple (LES)** : si $k = 0$ alors la série s'écrie

$$X_t = a_t + \varepsilon_t$$

avec $a_t = a$ une constante

- **lissage exponentiel double (LED)** : si $k = 1$ alors la série s'écrie

$$X_t = a_t + b_t t + \varepsilon_t$$

avec $a_t = a$ et $b_t = b$

- **lissage exponentiel généralisé (LEG)** : si k est quelconque ($k > 1$)

Le lissage exponentiel simple (LES)

Le lissage exponentiel simple (LES) permet d'effectuer des prévisions pour des séries chronologiques dont la tendance est constante et sans saisonnalité

On cherche à faire une prévision de la série (X_t), avec $t = 1, \ldots, T$, à l'horizon h, cad prévoir la valeur

$$X_{T+h} = \widehat{X}_T(h)$$

Soit un réel β tel que $0 < \beta < 1$, on cherche une prévision $\widehat{X}_T(h)$ sous la forme de la constante qui s'ajuste le mieux au sens des moindres carrés pondérés au voisinage de T, cad la solution du problème de minimisation

$$\min_a \sum_{j=0}^{T-1} \beta^j (X_{T-j} - a)^2$$

On remarque que dans l'expression à minimiser l'influence des observations décroit lorsqu'on s'éloigne de la date T

Prévision

La **prévision** de la série à l'horizon h, $\widehat{X}_T(h)$, fournie par le **lissage exponentiel simple (LES)** est donnée par

$$\widehat{X}_T(h) = (1-\beta)\sum_{j=0}^{T-1} \beta^j X_{T-j}$$

où β : **coefficient de lissage**

$\widehat{X}_T(h)$ est une **moyenne des observations passées** où le poids de chaque observation décroît de façon exponentielle avec la distance.

On peut remarquer que $\widehat{X}_T(h)$ ne dépend pas de h :

$$\widehat{X}_T(h) \equiv \widehat{X}_T$$

Dérivation de la formule du LES : On cherche la constante telle que

$$\min_a \sum_{j=0}^{T-1} \beta^j (X_{T-j} - a)^2 \equiv \min_a S$$

La solution du problème de minisation permet d'obtenir le **prédicteur du LES** (cf. annexe) :

$$\widehat{a}_T = \left(\frac{1-\beta}{1-\beta^T} \right) \sum_{j=0}^{T-1} \beta^j X_{T-j}$$

On peut obtenir alors la **formule de mise à jour** ou **formule adaptative** (cf. annexe)

$$\widehat{X}_T(h) = \widehat{X}_{T-1}(h) + (1-\beta)\left(X_T - \widehat{X}_{T-1}(h)\right)$$

Cette formule permet de calculer directement (à partir de la prévision $\widehat{X}_{T-1}$ à la date $T-1$) une nouvelle prévision $\widehat{X}_T$ lorsqu'une nouvelle observation X_T est effectuée.

Cette équation permet donc de mettre à jour les prévisions à l'horizon h à partir de la dernière prévision de manière extrêmement simple

L'initialisation de la récurrence (car X_1 connu et $\widehat{X}_1$ inconnu) est généralement faite de 2 manières :

- soit $\widehat{X}_1 = X_1$
- soit $\widehat{X}_1 = (1/T)\sum_{t=1}^{T} X_t$

La **valeur optimale du coefficient de lissage** β, avec $0 < \beta < 1$, résulte d'un arbitrage entre l'inertie liée à l'intégration des données lointaines et la sensibilité aux valeurs récentes

- Si β est proche de 0 $\Rightarrow$ **prévision souple** : fortement influencée par les observations les plus récentes
  (β^j devenant négligeable pour les grandes valeurs de j)
- Si $\beta = 0 \Rightarrow$ prévision égale à la dernière valeur observée
- Si β est proche de 1 $\Rightarrow$ **prévision rigide** : l'influence des observations passées est d'autant plus importante et remonte loin dans le passé. Peu sensible aux fluctuations exceptionnelles (conjoncturelles)
- Si $\beta = 1 \Rightarrow$ toutes les prévisions sont identiques
- En pratique, on prend $\beta \in]0; 1[$ afin d'exclure les 2 cas extrêmes

Le **choix du** β est en général très subjectif et varie selon le contexte de l'étude et/ou du type de prévision souhaité.
En pratique :

- si on veut une **prévision rigide**, on choisira $\beta \in [0.7; 0.99]$
- si on veut une **prévision souple**, on choisira $\beta \in [0.01; 0.3]$
- on peut déterminer β à partir des données (*data-driven*), en **minimisant** la somme des carrés des erreurs de prévision (MSE, *mean squared error*) aux dates $1, \ldots, T-h$:

$$\min \sum_{t=1}^{T-h} e_t^2 = \sum_{t=1}^{T-h} \left(X_{t+h} - \widehat{X}_t(h) \right)^2$$

Le package R stat estime le LES avec la fonction `HoltWinters`

```
library(seasonal)
y <- read.table("c://R//data//demandbrut.txt")
yy <- ts(data = y, start=c(1996,01),frequency=12)
seasX <- seas(yy)
cvs <- final(seasX)

# LES automatique
m <- HoltWinters(cvs,beta=FALSE,gamma=FALSE)

# LES en fixant a = 0.5
m <- HoltWinters(cvs,alpha=0.5,beta=FALSE,gamma=FALSE)

summary(m)
plot(m)
plot(m$fitted[,1])


# prévision sur un horizon h = 50 et intervalle à 95%
p=predict(m,n.ahead=50,prediction.interval = TRUE)
plot(m,p)
show(p)
prevp = p[1]
```

Figure: Lissage LES avec $\beta = 0.1$, 0.5 et 0.9

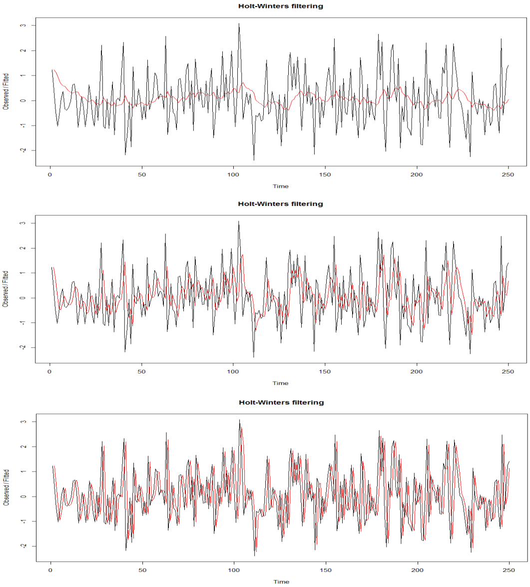


Figure: Prévision LES avec $\beta = 0.1$, 0.5 et 0.9 pour un horizon de $h = 50$

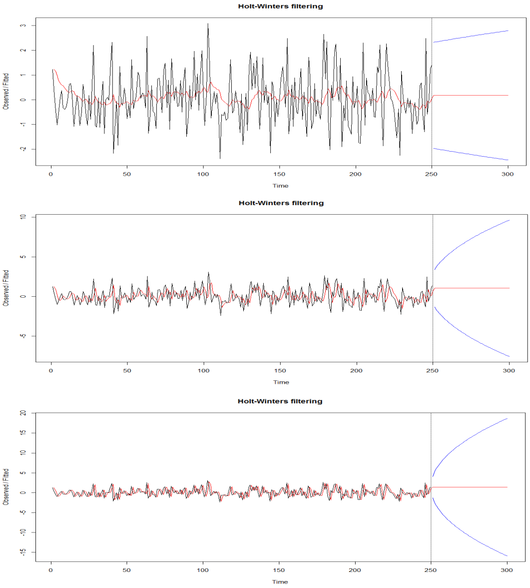
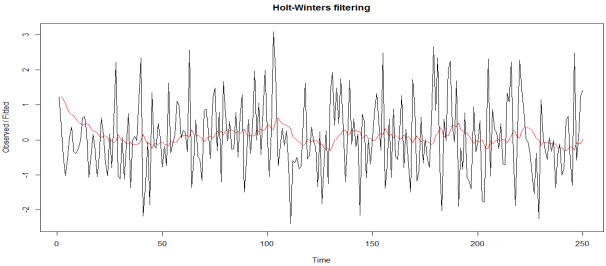
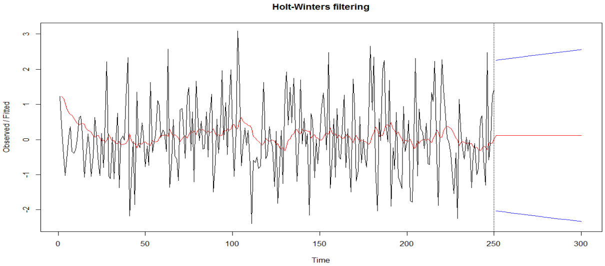


Figure: Lissage et prévision LES en minimisant la somme des carrés des erreurs de prévision





Smoothing parameter: $\beta = 0.07835199 \Rightarrow$ prévision très souple : forte influence des observations les plus récentes

Le lissage exponentiel double (LED)

Le lissage exponentiel double (LED) généralise l'idée du LES au cas où la série peut être ajustée par une droite au voisinage de T

$$X_t = a + b(t - T)$$

On cherche dans ce cas une prévision à l'horizon h, $\widehat{X}_T(h)$ de la forme :

$$\widehat{X}_T(h) = \widehat{a}_T + \widehat{b}_T h$$

où le couple $(\widehat{a}_T, \widehat{b}_T)$ minimise la fonction

$$\min_{a,b} \sum_{j=0}^{T-1} \beta^j \left[X_{T-j} - (a + bj)\right]^2 \equiv \min_{a,b} Q$$

La solution du problème de minisation permet d'obtenir les prédicteurs du LED :

$$\begin{aligned}\widehat{a}_T &= 2S_1(T) - S_2(T)\\ \widehat{b}_T &= \frac{1-\beta}{\beta}(S_1(T) - S_2(T))\end{aligned}$$

avec

- la série lissée, appelée LES de la série initiale

$$S_1(t) = (1-\beta)\sum_{j=0}^{t-1}\beta^j X_{t-j}$$

- la série doublement lissée, appelée LES de la série lissée

$$S_2(t) = (1-\beta)^2\sum_{j=0}^{t-1}(j\beta^j X_{t-j}) + (1-\beta)S_1(t)$$

Prévision

La **prévision** de la série à l'horizon h, $\widehat{X}_T(h)$, fournie par le **lissage exponentiel double (LED)** est donnée par

$$\widehat{X}_T(h) = \widehat{a}_T + \widehat{b}_T h$$

où β est la **coefficient de lissage**, et le couple $(\widehat{a}_T, \widehat{b}_T)$ est donné par

$$\begin{aligned} \widehat{a}_T &= 2S_1(T) - S_2(T) \\ \widehat{b}_T &= \frac{1-\beta}{\beta}\left(S_1(T) - S_2(T)\right) \end{aligned}$$

Le lissage exponentiel double

Les formules de mise à jour ou formules adaptatives s'obtiennent à partir des expressions suivantes :

$$\begin{aligned}
\widehat{a}_T &= \widehat{a}_{T-1} + \widehat{b}_{T-1} + (1-\beta)^2 \left(X_T - \widehat{X}_{T-1}(1) \right) \\
\widehat{b}_T &= \widehat{b}_{T-1} + (1-\beta)^2 \left[X_T - \widehat{X}_{T-1}(1) \right]
\end{aligned}$$

L'initialisation de la récurrence est généralement obtenue par :

$$\begin{aligned}
\widehat{a}_2 &= X_2 - X_1 \\
\widehat{b}_2 &= X_2
\end{aligned}$$

Lissage exponentiel généralisé (LEG)

On peut généraliser cette technique de lissage avec le lissage exponentiel généralisé (LEG) pour traiter des séries sans saisonnalité et présentant une tendance polynômiale de degré supérieur à 1 : $f_t^{(k)}$ avec $k > 1$

Les résultats font intervenir, dans ce cas, les opérateurs de lissage d'ordre k, $S_k(t)$, itérés d'ordre k de $S_1(t)$

Lissage Holt-Winters

Un méthode un peu différente a été introduite par Holt et Winters, adaptée aux séries sans saisonnalité pouvant être ajustées par une droite au voisinage de T (comme pour le lissage exponentiel double).

La différence entre la méthode de Holt-Winters et le lissage exponentiel double porte sur les formules de mise à jour.

Holt et Winters ont proposé une version des formules de mise à jour où les pondérations ne dépendent pas que d'un seul paramètre (β) mais de 2 paramètres (α, γ)

Propriété

Les méthodes de **Holt-Winters** et **LED** sont identiques si on prend

$$\begin{aligned} \alpha &= \beta^2 \\ \gamma &= 1 - \frac{(1-\beta)^2}{1-\beta^2} = \frac{2\beta}{1-\beta^2} \end{aligned}$$

L'initialisation peut se faire comme dans le cas du LED.

L'avantage de cette approche est d'avoir une plus grande flexibilité mais la contrepartie est de devoir régler deux paramètres.

- si α et γ sont proches de 1 $\Rightarrow$ prévision lisse (fort poids du passé)
- si β et γ sont petits $\Rightarrow$ α grand $\Rightarrow$ lissage fort (on tient compte du passé lointain)
- le choix des coefficients de lissage est encore plus subjectif et peut être réglé automatiquement en utilisant un critère des moindres carrés

La prévision à l'horizon h par cette méthode est donnée

$$\widehat{X}_T(h) = \widehat{a}_T + \widehat{b}_T h$$

Le package R **holtwinters** avec la fonction `HoltWinters`

 # LED HW automatique
```
m <- HoltWinters(cvs,gamma=FALSE)
```

 # LED HW en fixant a et b (paramètre $1-a'$)
 # avec $a = 1-(a')^2$, $b = (1-a')/(1+a')$
```
m <- HoltWinters(cvs,alpha=a,beta=b,gamma=FALSE)

summary(m)
plot(m)
plot(m$fitted[,1])
```

 # prévision sur un horizon $h = 50$
```
p=predict(m,n.ahead=50,prediction.interval = TRUE)
plot(m,p)
```

Figure: Lissage LED par la méthode Holt-Winters avec $\beta = 0.1$, 0.5 et 0.9

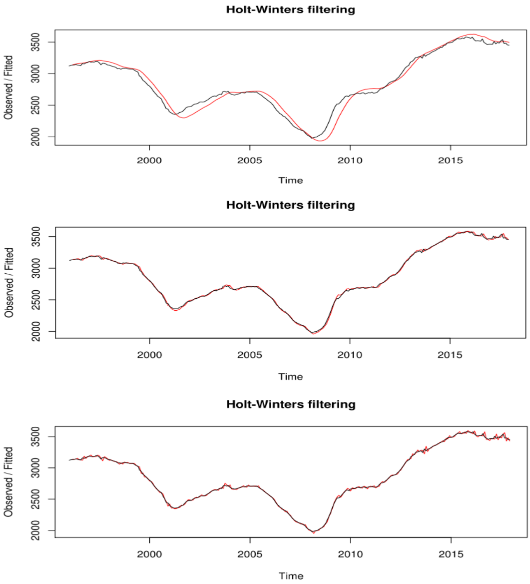


Figure: Prévision LED avec $\beta = 0.1$, 0.5 et 0.9 pour un horizon de $h = 50$

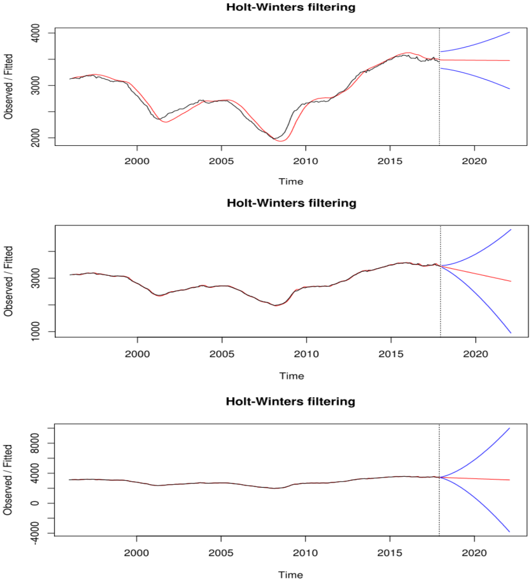
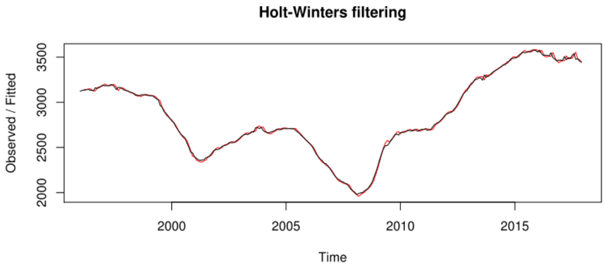
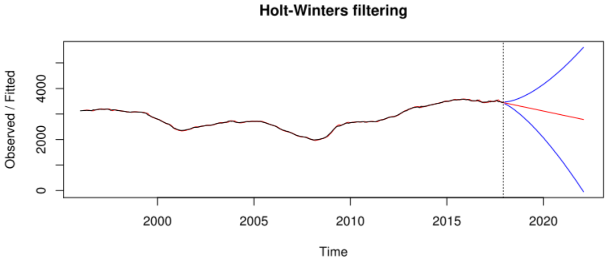


Figure: Lissage et prévision LED en minimisant la somme des carrés des erreurs de prévision





Smoothing parameters: $\alpha = 0.7947475$ et $\gamma = 0.4852147$

Table: Paramètres estimés pour la constante *a* et la tendance *b*.

Holt-Winters R param $(1-\alpha')$	0.1	0.5	0.9	min
α	0.190	0.750	0.990	0.795
γ	0.053	0.333	0.818	0.485
$\widehat{a}$	3486.5	3450.6	3451.2	3449.1
$\widehat{b}$	-0.207	-11.34	-6.974	-13.32

Notes : $\alpha = 1-\alpha'^2$ et $\gamma = (1-\alpha')/(1+\alpha')$.

$\Rightarrow$ prévision lisse (fort poids du passé)

La méthode de Holt et Winters saisonnière

La méthode saisonnière additive

Définition

On suppose que la série $X_1, \dots, X_T$ peut être approchée, au voisinage de T, par le **modèle additif**

$$a + b(t - T) + S_t + \varepsilon_t$$

où S_t : la **saisonnalité** de période P.

La **méthode saisonnière additive de Holt-Winters** propose pour l'estimation de a, b et S_t les formules de mise à jour suivantes

$$\begin{aligned}
\widehat{a}_T &= (1-\beta)\widehat{a}_{T-1} + \beta\left(\widehat{b}_T - \widehat{b}_{T-1}\right) \\
\widehat{b}_T &= \alpha\left(X_T - \widehat{S}_{T-P}\right) + (1-\alpha)\left(\widehat{b}_{T-1} - \widehat{a}_{T-1}\right) \\
\widehat{S}_T &= \gamma\left(X_T - \widehat{b}_T\right) + (1-\gamma)\widehat{S}_{T-P}
\end{aligned}$$

où $\alpha, \beta, \gamma \in]0;1[$: **coefficients de lissage**

Le choix des coefficients de lissage dans la pratique peut s'effectuer en minimisant la somme des carrés des erreurs de prévision

On doit calculer les valeurs initiales pour utiliser les formules de mise à jour.

Cela peut se faire en 3 étapes :

1. Appliquer une moyenne mobile adaptée à la série afin de supprimer la saisonnalité. On obtient ainsi une estimation de la tendance linéaire.
2. Appliquer un opérateur adapté à la série transformée afin de récupérer le coefficient a. L'estimer. En déduire une estimation du coefficient b.
3. En déduire une estimation du saisonnier

La méthode saisonnière multiplicative

Définition

On suppose que la série $X_1, \ldots, X_T$ peut être approchée, au voisinage de T, par le **modèle multiplicatif**

$$(a + b(t - T))\, S_t$$

où S_t : la **saisonnalité** de période P.

La **méthode saisonnnière multiplicative de Holt-Winters** propose pour l'estimation de a, b et S_t les formules de mise à jour suivantes

$$\begin{aligned}
\widehat{a}_T &= (1-\beta)\widehat{a}_{T-1} + (1-\beta)\left(\widehat{b}_T - \widehat{b}_{T-1}\right) \\
\widehat{b}_T &= \alpha \frac{X_T}{\widehat{S}_{T-P}} + (1-\alpha)\left(\widehat{b}_{T-1} - \widehat{a}_{T-1}\right) \\
\widehat{S}_T &= \gamma \frac{X_T}{\widehat{b}_T} + (1-\gamma)\widehat{S}_{T-P}
\end{aligned}$$

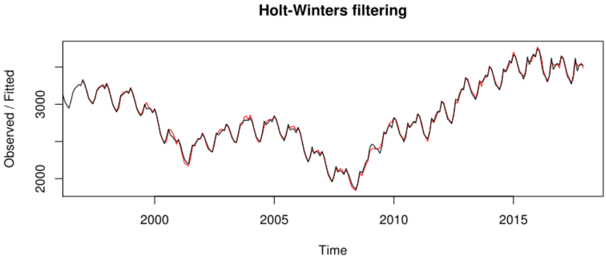
où $\alpha, \beta, \gamma \in]0;1[$: **coefficients de lissage**

Les valeurs initiales sont déterminées de la même manière que pour le cas saisonnier additif.

Le package R **holtwinters** avec la fonction `HoltWinters`

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

# lissage exponentiel Holt-Winters additif automatique
m <- HoltWinters(yy,seasonal=add)

# lissage exponentiel Holt-Winters additif en fixant a, b et c
m <- HoltWinters(yy,alpha=a,beta=b,gamma=c,seasonal=add)

# lissage exponentiel Holt-Winters multiplicatif automatique
m <- HoltWinters(yy,seasonal=mul)

# lissage exponentiel Holt-Winters multiplicatif en fixant a, b et c
xlisse <- HoltWinters(yy,alpha=a,beta=b,gamma=c,seasonal=mul)
```

Figure: Lissage Holt-Winters avec saisonnalité additive et multiplicative



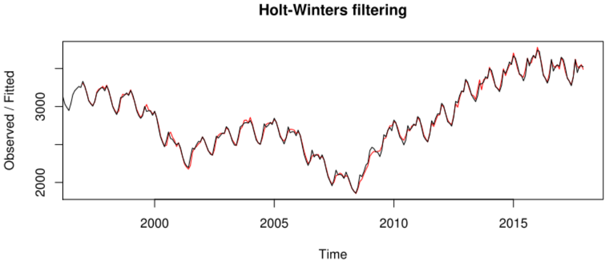


Figure: Prévision Holt-Winters avec saisonnalité additive et multiplicative pour un horizon de $h = 50$

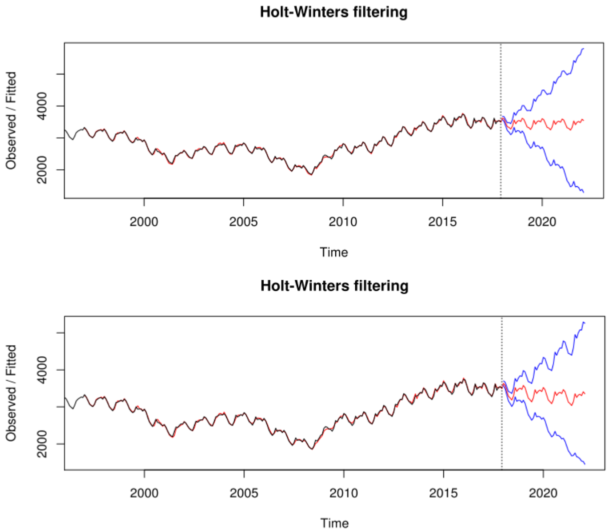


Table: Paramètres estimés de la méthode saisonnière Holt-Winters.

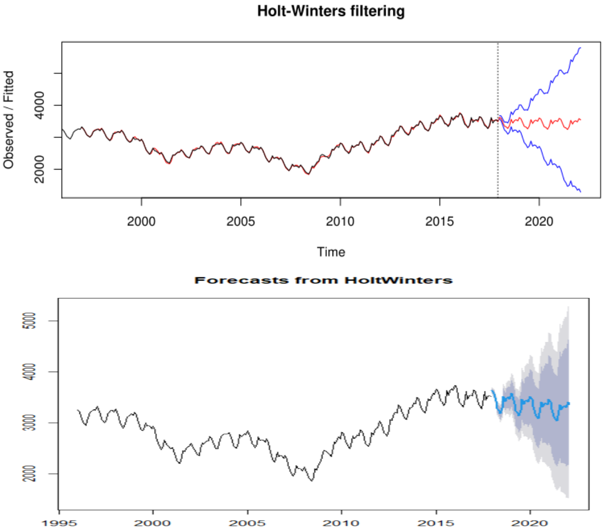
Smoothing parameters	α	β	γ	$\widehat{a}$	$\widehat{b}$	
additif	0.813	0.225	1.00	3472.7	-0.889	
multiplicatif	0.886	0.136	1.00	3447.4	-4.483	
Seasonal parameters	$\widehat{a}_1$	$\widehat{a}_2$	$\widehat{a}_3$	$\widehat{a}_4$	$\widehat{a}_5$	$\widehat{a}_6$
additif	154.7	117.1	-3.25	-118.3	-137.6	-192.0
multiplicatif	1.057	1.040	0.998	0.958	0.946	0.931
Seasonal parameters	$\widehat{a}_7$	$\widehat{a}_8$	$\widehat{a}_9$	$\widehat{a}_{10}$	$\widehat{a}_{11}$	$\widehat{a}_{12}$
additif	-101.5	99.8	-19.3	60.9	85.0	50.9
multiplicatif	0.965	1.032	1.003	1.023	1.032	1.022

Remarque coef. saisonniers : $\widehat{a}_i < 0$ cas additif $\Leftrightarrow$ $\widehat{a}_i < 1$ cas multiplicatif

Le package R forecast permet également de prévoir à partir des différentes méthodes de lissage avec la fonction `forecast`

```
y <- read.table("c://R//data//demandbrut.txt")
yy <- ts(data = y, start=c(1996,01),frequency=12)
m <- HoltWinters(yy,seasonal="mul")

# prévision sur un horizon h = 50 et intervalle à 80% et 95%
library(forecast)
fit = forecast(m, h=50)
plot(fit)
show(fit)

# Point forecasts
prevf = fit$mean
show(prevf)

# A list containing information about the fitted model
mod = fit$model
show(mod)
```

Figure: Prévision Holt-Winters avec saisonnalité multiplicative pour un horizon de $h = 50$ avec les fonctions `predict` et `forecast`



Lissage exponentiel sous forme de composantes

Une représentation alternative des méthodes de lissage exponentiel est la représentation sous forme de composantes.

Cette représentation comprend une équation de prévision et une équation de lissage pour chaque composante incluse dans la méthode.

La **représentation sous forme de composantes du LES** est donnée par

$$\begin{aligned} \widehat{y}_{t+h|t} &= l_t \\ l_t &= \alpha y_t + (1-\alpha) l_{t-1} \end{aligned}$$

où l_t est la composante de niveau (ou valeur lissée).

La **représentation sous forme de composantes du LED** est donnée par

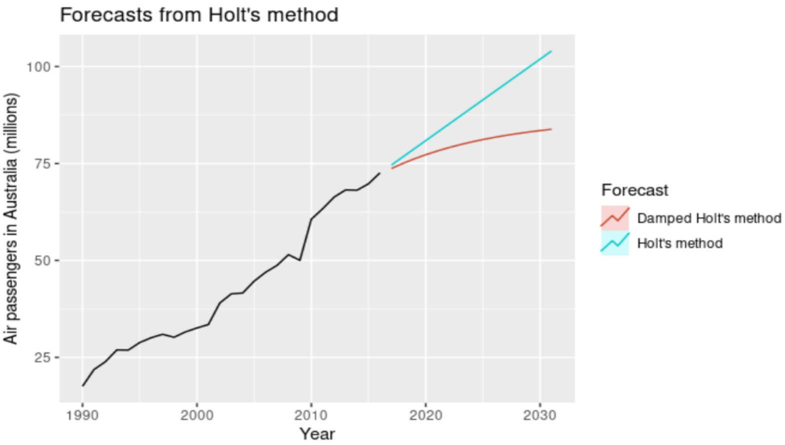
$$\begin{aligned} \widehat{y}_{t+h|t} &= l_t + hb_t \\ l_t &= \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1}) \\ b_t &= \beta^*(l_t - l_{t-1}) + (1-\beta^*)b_{t-1} \end{aligned}$$

- l_t : la composante de niveau, avec α son paramètre de lissage $0 \leq \alpha \leq 1$
- b_t : la composante de tendance (pente), avec β^* son paramètre de lissage $0 \leq \beta^* \leq 1$.

La **représentation sous forme de composantes du LED "damped"** : pour améliorer les prévisions du LED Gardner et McKenzie (1985) ont introduit un paramètre ϕ qui "atténue" (***damped***) la tendance à une ligne plate dans le futur :

$$\begin{aligned} \widehat{y}_{t+h|t} &= l_t + (\phi + \phi^2 + \cdots + \phi^h)b_t \\ l_t &= \alpha y_t + (1-\alpha)(l_{t-1} + \phi b_{t-1}) \\ b_t &= \beta^*(l_t - l_{t-1}) + (1-\beta^*)\phi b_{t-1} \end{aligned}$$

- le paramètre *damped* $0 < \phi < 1$
- si $\phi = 1 \Rightarrow$ LED
- en pratique $0.8 \leq \phi \leq 0.98$


Forecasts from Holt's method
Air passengers in Australia (millions)
100
75
50
25
1990
2000
2010
2020
2030
Year
Forecast
Damped Holt's method
Holt's method

La **représentation sous forme de composantes du lissage exponentiel de Holt-Winters** pour la méthode additive est donnée par

$$\begin{aligned}
\widehat{y}_{t+h|t} &= l_t + hb_t + s_{t+h-m(k+1)} \\
l_t &= \alpha(y_t - s_{t-m}) + (1-\alpha)(l_{t-1} + b_{t-1}) \\
b_t &= \beta^*(l_t - l_{t-1}) + (1-\beta^*)b_{t-1} \\
s_t &= \gamma(y_t - l_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m}
\end{aligned}$$

- s_t : composante saisonnière
- m : les périodes de la saisonnalité ($m = 4$ ou 12)
- k : la partie entière de $(h-1)/m$: $\lfloor (h-1)/m \rfloor$
- le paramètre de lissage : $0 \leq \gamma \leq 1-\alpha$

Une représentation de la méthode multiplicative est également proposée.

Trend Component		Seasonal Component N (None)	A (Additive)	M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

Short hand	Method
(N,N)	Simple exponential smoothing
(A,N)	Holt's linear method
(A_d,N)	Additive damped trend method
(A,A)	Additive Holt-Winters' method
(A,M)	Multiplicative Holt-Winters' method
(A_d,M)	Holt-Winters' damped method

La méthode ETS

Hyndman et alii (2008) proposent des versions sous forme d'espace d'état de la méthode saisonnière de Holt-Winters (additive et multiplicative)

Ces modèles sont nommés modèles ETS (*Error, Trend, Seasonal* ou *ExponenTial Smoothing*)

Rappel : la représentation sous forme de composantes du LES est donnée par

$$\widehat{y}_{t+h|t} = l_t \tag{1}$$

$$l_t = \alpha y_t + (1-\alpha) l_{t-1} \tag{2}$$

L'équation de lissage (2) peut se ré-écrire de la manière suivante :

$$l_t = l_{t-1} + \alpha(y_t - l_{t-1}) = l_{t-1} + \alpha e_t$$

où e_t sont les résidus définis par

$$e_t = y_t - l_{t-1} = y_t - \widehat{y}_{t+h|t-1} \qquad \Rightarrow \qquad y_t = l_{t-1} + e_t$$

Afin de mettre ces équations sous forme d'espace d'état, ils supposent que les résidus $e_t = \varepsilon_t \sim i.i.d.\mathcal{N}(0;\sigma^2)$

Les équations peuvent alors se ré-écrire avec une équation de mesure (3) (ou équation d'observation) et une équation d'état (4) (ou équation de transition) :

$$y_t = l_{t-1} + \varepsilon_t \quad (3)$$

$$l_t = l_{t-1} + \alpha\varepsilon_t \quad (4)$$

Ce modèle sera noté ETS(A,N,N) : lissage exponentiel simple (LES) avec erreurs additives

De manière similaire les méthodes de lissage exponentiel peuvent se ré-écrire sous la forme d'espace d'état

Trend	Seasonal N	Seasonal A	Seasonal M
N	$\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}$ $\hat{y}_{t+h\|t} = \ell_t$	$\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)\ell_{t-1}$ $s_t = \gamma(y_t - \ell_{t-1}) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = \ell_t + s_{t-m+h_m^+}$	$\ell_t = \alpha(y_t / s_{t-m}) + (1-\alpha)\ell_{t-1}$ $s_t = \gamma(y_t / \ell_{t-1}) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = \ell_t s_{t-m+h_m^+}$
A	$\ell_t = \alpha y_t + (1-\alpha)(\ell_{t-1} + b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$ $\hat{y}_{t+h\|t} = \ell_t + hb_t$	$\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$ $s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = \ell_t + hb_t + s_{t-m+h_m^+}$	$\ell_t = \alpha(y_t / s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$ $s_t = \gamma(y_t / (\ell_{t-1} - b_{t-1})) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = (\ell_t + hb_t)s_{t-m+h_m^+}$
A$_d$	$\ell_t = \alpha y_t + (1-\alpha)(\ell_{t-1} + \phi b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)\phi b_{t-1}$ $\hat{y}_{t+h\|t} = \ell_t + \phi_h b_t$	$\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + \phi b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)\phi b_{t-1}$ $s_t = \gamma(y_t - \ell_{t-1} - \phi b_{t-1}) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = \ell_t + \phi_h b_t + s_{t-m+h_m^+}$	$\ell_t = \alpha(y_t / s_{t-m}) + (1-\alpha)(\ell_{t-1} + \phi b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)\phi b_{t-1}$ $s_t = \gamma(y_t / (\ell_{t-1} - \phi b_{t-1})) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = (\ell_t + \phi_h b_t)s_{t-m+h_m^+}$
M	$\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}b_{t-1}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}$ $\hat{y}_{t+h\|t} = \ell_t b_t^h$	$\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)\ell_{t-1}b_{t-1}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}$ $s_t = \gamma(y_t - \ell_{t-1}b_{t-1}) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = \ell_t b_t^h + s_{t-m+h_m^+}$	$\ell_t = \alpha(y_t / s_{t-m}) + (1-\alpha)\ell_{t-1}b_{t-1}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}$ $s_t = \gamma(y_t / (\ell_{t-1}b_{t-1})) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = \ell_t b_t^h s_{t-m+h_m^+}$
M$_d$	$\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}b_{t-1}^\phi$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}^\phi$ $\hat{y}_{t+h\|t} = \ell_t b_t^{\phi_h}$	$\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)\ell_{t-1}b_{t-1}^\phi$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}^\phi$ $s_t = \gamma(y_t - \ell_{t-1}b_{t-1}^\phi) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = \ell_t b_t^{\phi_h} + s_{t-m+h_m^+}$	$\ell_t = \alpha(y_t / s_{t-m}) + (1-\alpha)\ell_{t-1}b_{t-1}^\phi$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}^\phi$ $s_t = \gamma(y_t / (\ell_{t-1}b_{t-1}^\phi)) + (1-\gamma)s_{t-m}$ $\hat{y}_{t+h\|t} = \ell_t b_t^{\phi_h} s_{t-m+h_m^+}$

La méthode ETS

L'estimation des divers paramètres se fait par le maximum de vraisemblance.

Dans la représentation sous forme de composantes les paramètres doivent respecter

$$0 < \alpha, \beta^*, \gamma^*, \phi < 1$$

Dans la représentation sous forme d'espace d'état les paramètres de lissage des composantes de niveau (α), de tendance (β), de saisonnalité (γ) et *damped* (ϕ) ont les contraintes suivantes :

$$\begin{aligned}
\beta &= \alpha\beta^* \\
\gamma &= (1-\alpha)\gamma^* \\
& 0 < \alpha < 1 \\
& 0 < \beta < \alpha \\
& 0 < \gamma < 1-\alpha \\
& 0.8 < \phi < 0.98 \qquad \text{(R package forecast)}
\end{aligned}$$

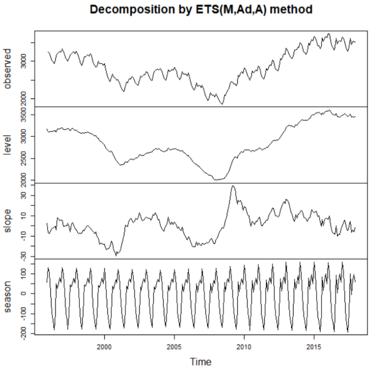
Dans le cadre des modèles ETS il est possible d'utiliser un critère de sélection de modèles (AIC, AIC_c ou BIC) pour déterminer le modèle ETS le plus approprié

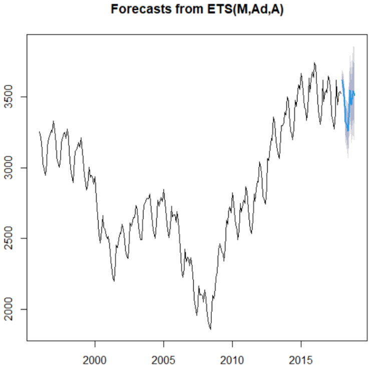
Le package forecast permet de **décomposer** et de **prévoir** à partir de la méthode ETS avec la fonction `ets`

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

```
ETS(M,Ad,A)

Call:
 ets(y = yy)

  Smoothing parameters:
    alpha = 0.8561
    beta  = 0.1282
    gamma = 0.1439
    phi   = 0.9403
```

Modèle ETS(M,Ad,A) : modèle HW avec tendance *damped* additive, saisonnalité additive et des erreurs multiplicative


Decomposition by ETS(M,Ad,A) method
observed
level
slope
season
3000
2000
3500
3000
2500
2000
30
10
-10
-30
100
0
-100
-200
2000
2005
2010
2015
Time


Forecasts from ETS(M,Ad,A)
3500
3000
2500
2000
2000
2005
2010
2015

La méthode TBATS

De Livera, Hyndman et Snyder (2011) proposent de modifier les modèles ETS pour prendre en compte les modèles non linéaires, l'autocorrélation des résidus et une plus grande variété de comportements saisonniers :

- une transformation Box-Cox :

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

- les erreurs ε_t suivent un processus ARMA(p, q)

$$\varepsilon_t = \sum_{i=1}^{p} \varphi_i \varepsilon_{t-i} + \sum_{i=1}^{q} \theta_i \upsilon_{t-i} + \upsilon_t$$

- une composante saisonnière plus variée :

$$\sum_{i=1}^{T} s_{t-m_i}^{(i)}$$

où $m1, \ldots, m_T$ sont les périodes saisonnières

Ces modèles sont notés modèles BATS (*Box-Cox transform, ARMA errors, Trend, and Seasonal components*) : BATS$(\lambda, \phi, p, q, m_1, m2, \ldots, m_T)$:

- BATS$(1, 1, 0, 0, m_1)$: méthode (simple) Holt-Winters saisonnier additif
- BATS$(1, 1, 1, 0, m_1, m_2)$: méthode double Holt-Winters saisonnier additif avec ereurs AR(1) (Taylor, 2001, 2008)

De Livera, Hyndman et Snyder (2011) modifient les modèles BATS en introduisant une représentation trigonométrique des composantes saisonnières avec le modèle de Harvey (1989), nommés modèles TBATS (*trigonometric BATS*) :

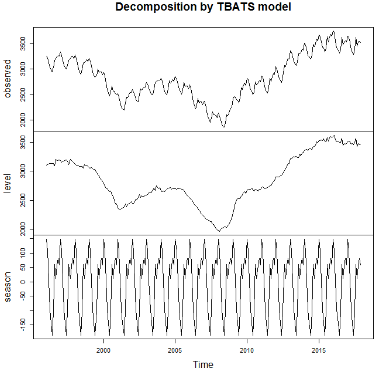
- TBATS$(\lambda, \phi, p, q, \{m_1, k_1\}, \{m_2, k_2\}, \ldots, \{m_T, k_T\})$
- k_j $(j = 1, \ldots, T)$: le nombre d'harmoniques dans les modèles trigonométriques saisonniers

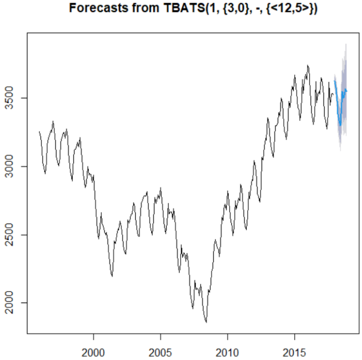
Le package forecast permet de **décomposer** et de **prévoir** à partir de la méthode TBATS avec la fonction `tbats`

Les paramètres du modèle TBATS sont : TBATS$(\alpha, p, q, \phi, \{m_1, k_1\}, \{m_2, k_2\}, \ldots, \{m_T, k_T\})$ avec $\alpha = \lambda$

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

```
TBATS(1, {3,0}, -, {<12,5>})

Call: tbats(y = yy)

Parameters
  Alpha: 1.069024
  Gamma-1 Values: 0.0002268646
  Gamma-2 Values: 9.576097e-05
  AR coefficients: -0.093992 0.29299 0.385744
```


Decomposition by TBATS model
observed
level
season
Time
2000
2005
2010
2015


Forecasts from TBATS(1, {3,0}, -, {<12,5>})
2000
2500
3000
3500
2000
2005
2010
2015