

Modélisation et prévision

Séries chronologiques - Séance 1
*Décomposition d'une chronique*Frédéric Sur
École des Mines de Nancy<https://members.loria.fr/FSur/enseignement/modprev/>

Qu'est-ce qu'une série chronologique ?

Séries temporelles / Chroniques / *Time series*

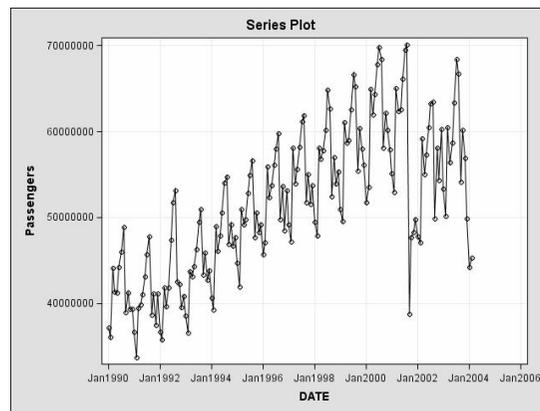
- suite d'observations d'une grandeur au cours du temps.
- temps *discret*.

Exemples :

- économétrie (taux de chômage),
- finance (cours d'action),
- écologie (pollution),
- démographie (population),
- météorologie (relevé de températures),
- astronomie (fluctuations de la magnitude d'un astre)...

Modèle sous-jacent, intrinsèque, inconnu liant les observations.

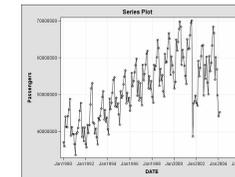
Exemple de série chronologique



Trafic aérien aux USA de 1990 à 2004.

But du cours : décrire l'évolution des chroniques à l'aide de modèles basés sur des propriétés statistiques.

Dans le modèle. . .



- **Perturbations ponctuelles** : variations forte amplitude. (*grève, krach boursier, attentats, mesure aberrante. . .*)
→ À traiter en premier. Cf « modèles d'intervention ».
- **Tendance T_t** : évolution globale.
- **Variations saisonnières S_t** : fluctuations périodiques.
Cf « données corrigées des variations saisonnières ».
→ moyenne nulle sur une période.
- **Composante résiduelle (irrégulière) u_t** : fluctuations irrégulières imprévisibles de faible amplitude.
→ moyenne nulle.

Pourquoi étudier les chroniques ?

- Description / explication d'un phénomène
chômage : variation tendancielle ou fluctuation saisonnière ?
- Préviation
consommation d'électricité, démographie, restauration rapide...
- Étude de la dynamique
cours d'actions.
- Impact d'un événement sur une série chronologique
« mesure » de l'impact des politiques publiques.

SG042 : techniques statistiques, aspects *temporels*.

Autres points de vue : aspects *fréquentiels*, traitement du signal.
→ cf cours électifs.

Modélisation et
prévision

F. Sur - ENSMN

Introduction

Modèles de
décomposition

Les trois composantes
Tendance
Composante
saisonnnière
Composante résiduelle

Décomposition

Typologie
Modèles
paramétriques
Filtrage

Conclusion

À faire avant chaque séance...

Polycopié du cours :
chaque chapitre correspond à une séance

→ **à lire avant le cours !**
(« pour en savoir plus » optionnel)

Données : fichier *chroniques.zip* sur Arche.

Pour approfondir :

- bibliographie du polycopié disponible à la bibliothèque.
- members.loria.fr/FSur/enseignement/modprev/
→ exercices à traiter en TP ;
→ slides + chapitres du polycopié à lire ;
→ suggestions de lectures complémentaires.

Modélisation et
prévision

F. Sur - ENSMN

Introduction

Modèles de
décomposition

Les trois composantes
Tendance
Composante
saisonnnière
Composante résiduelle

Décomposition

Typologie
Modèles
paramétriques
Filtrage

Conclusion

5/34

6/34

Séance 1

- 1 Introduction
- 2 **Modèles de décomposition**
 - Les trois composantes
 - Tendance
 - Composante saisonnière
 - Composante résiduelle
- 3 Décomposition
 - Typologie
 - Modèles paramétriques
 - Filtrage
- 4 Conclusion

Modélisation et
prévision

F. Sur - ENSMN

Introduction

Modèles de
décomposition

Les trois composantes
Tendance
Composante
saisonnnière
Composante résiduelle

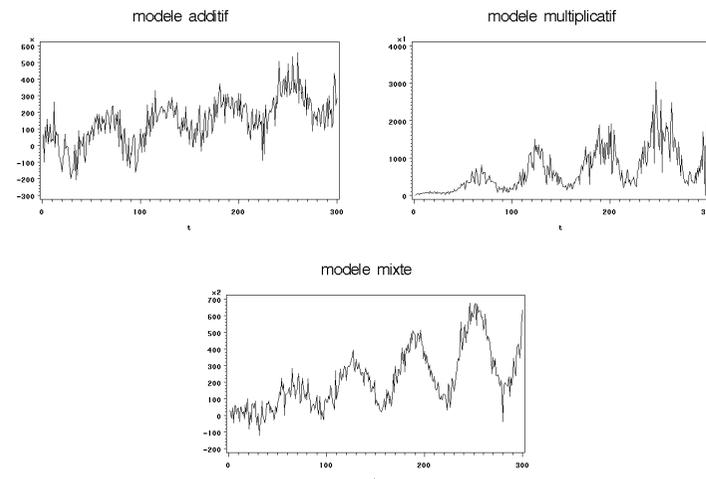
Décomposition

Typologie
Modèles
paramétriques
Filtrage

Conclusion

Modèles de décomposition

- Modèle additif : $X_t = T_t + S_t + u_t$
- Modèle multiplicatif : $X_t = T_t \cdot S_t \cdot u_t$
- Modèle mixte : $X_t = T_t \cdot S_t + u_t$



Modélisation et
prévision

F. Sur - ENSMN

Introduction

Modèles de
décomposition

Les trois composantes
Tendance
Composante
saisonnnière
Composante résiduelle

Décomposition

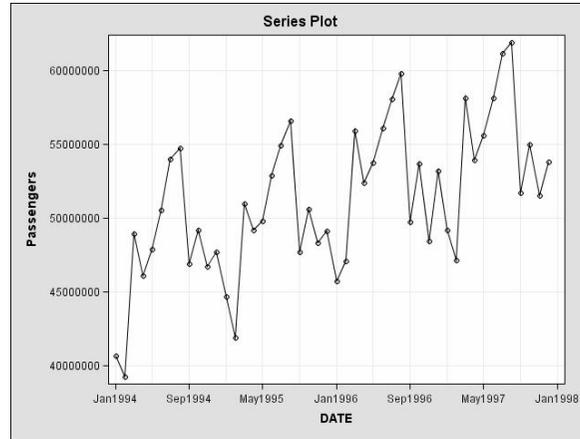
Typologie
Modèles
paramétriques
Filtrage

Conclusion

7/34

8/34

Exemple de décomposition



Trafic aérien aux USA de 1994 à 1997.

→ modèle mixte.

Modélisation et
prévision

F. Sur - ENSMN

Introduction

Modèles de
décomposition

Les trois composantes

Tendance

Composante
saisonnnière

Composante résiduelle

Décomposition

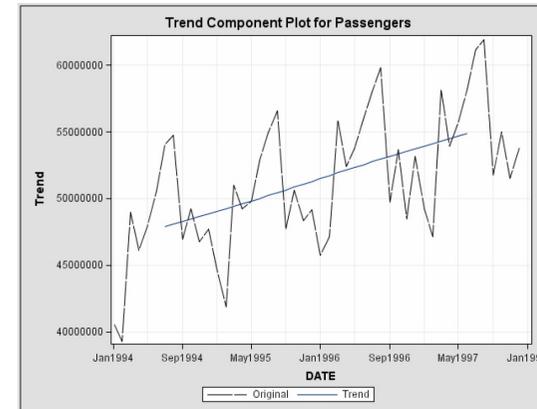
Typologie

Modèles
paramétriques

Filtrage

Conclusion

Exemple de décomposition



Tendance T_t .

Modélisation et
prévision

F. Sur - ENSMN

Introduction

Modèles de
décomposition

Les trois composantes

Tendance

Composante
saisonnnière

Composante résiduelle

Décomposition

Typologie

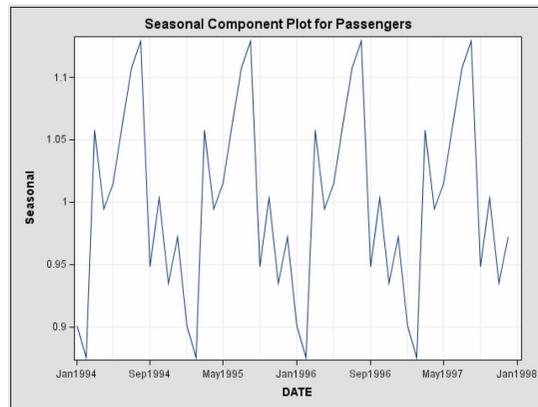
Modèles
paramétriques

Filtrage

Conclusion

10/34

Exemple de décomposition



Composante saisonnière S_t .

(modèle multiplicatif, donc S_t centré sur 1 et pas 0)

Modélisation et
prévision

F. Sur - ENSMN

Introduction

Modèles de
décomposition

Les trois composantes

Tendance

Composante
saisonnnière

Composante résiduelle

Décomposition

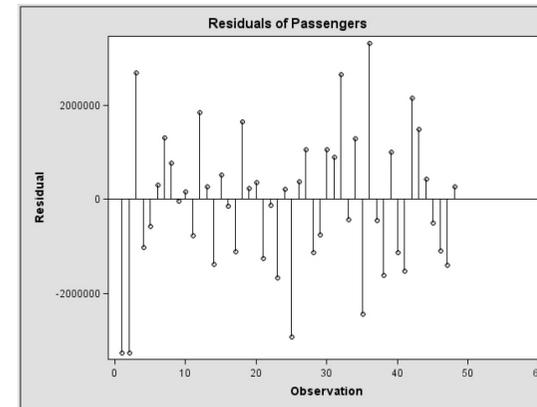
Typologie

Modèles
paramétriques

Filtrage

Conclusion

Exemple de décomposition



Résidus u_t .

(pas de structure apparente)

Modélisation et
prévision

F. Sur - ENSMN

Introduction

Modèles de
décomposition

Les trois composantes

Tendance

Composante
saisonnnière

Composante résiduelle

Décomposition

Typologie

Modèles
paramétriques

Filtrage

Conclusion

12/34

9/34

11/34

Composante tendancielle T_t

Exemples de composante tendancielle T_t paramétriques :

- linéaire : $T_t = at + b$
- quadratique : $T_t = at^2 + bt + c$
- exponentielle : $T_t = T_0 e^{at}$
- ...

avec les paramètres $a, b, c, T_0 \dots$ à déterminer.

Modélisation et
prévision

F. Sur - ENSMN

Introduction

Modèles de
décomposition

Les trois composantes

Tendance

Composante

saisonnière

Composante résiduelle

Décomposition

Typologie

Modèles

paramétriques

Filtrage

Conclusion

Composante saisonnière S_t

S_t périodique, de période p .

→ éventuellement plusieurs composantes périodiques superposées.

Exemple : comp. annuelle + comp. trimestrielle...

Question : comment trouver les périodes ?

→ « ça se voit », expertise.

→ analyse fréquentielle (périodogramme, proc spectra).

Modélisation et
prévision

F. Sur - ENSMN

Introduction

Modèles de
décomposition

Les trois composantes

Tendance

Composante

saisonnière

Composante résiduelle

Décomposition

Typologie

Modèles

paramétriques

Filtrage

Conclusion

13/34

14/34

Composante résiduelle u_t (le bruit)

→ pas de tendance ou de phénomène périodique dans u_t .

Modélisation : (u_t) réalisation d'un processus aléatoire (ε_t) stationnaire.

(i.e. propriétés statistiques *invariantes au cours du temps*)

Définition : (ε_t) stationnaire (au second ordre)

- $\forall t, \mathbb{E}(\varepsilon_t) = m$ (moyenne constante)
- $\forall t, s, \text{cov}(\varepsilon_t, \varepsilon_s) = \gamma(|t - s|)$
(autocovariance symétrique, invariante par translation)

En particulier $\forall t, \text{var}(\varepsilon_t) = \gamma(0)$ (variance constante).

→ les moments d'ordre 1 et 2 ne varient pas au cours du temps.

Modélisation et
prévision

F. Sur - ENSMN

Introduction

Modèles de
décomposition

Les trois composantes

Tendance

Composante

saisonnière

Composante résiduelle

Décomposition

Typologie

Modèles

paramétriques

Filtrage

Conclusion

Cas particuliers de processus stationnaires

- Bruit blanc faible : $m = 0$ et $\forall h > 0, \gamma(h) = 0$.
→ absence d'autocorrélation temporelle.
- Bruit blanc fort : $m = 0$ et (ε_t) i.i.d.
→ indépendance temporelle.
- Bruit blanc gaussien : (ε_t) i.i.d. $\sim \mathcal{N}(0, \sigma^2)$.
→ cf résidus dans la régression.

Modélisation et
prévision

F. Sur - ENSMN

Introduction

Modèles de
décomposition

Les trois composantes

Tendance

Composante

saisonnière

Composante résiduelle

Décomposition

Typologie

Modèles

paramétriques

Filtrage

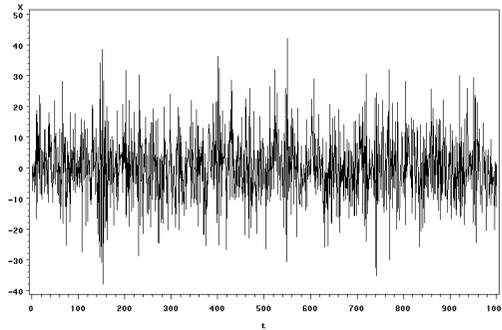
Conclusion

15/34

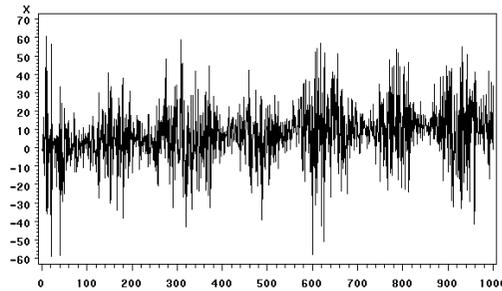
16/34

Exemples

bruit blanc
gaussien



processus non
stationnaire



Caractérisation des processus stationnaires

Processus stationnaires considérés :

$$X_t = m + \sum_{j=0}^{+\infty} a_j \varepsilon_{t-j}$$

avec $\sum_{j=0}^{+\infty} |a_j| < +\infty$ et (ε_t) **bruit blanc** (faible)

Remarque : (ε_t) sera même généralement un bruit blanc *gaussien*.

Par définition, (X_t) caractérisé par

- moyenne m ,
- fonction de covariance γ .

→ comment estimer m et γ ?

Estimation de la moyenne m

Estimateur de m :

$$\bar{X}_T = \frac{1}{T} \sum_{t=1}^T X_t$$

Estimation de l'autocovariance γ

Définition : fonction d'autocorrélation

$$\forall h \geq 0, \rho(h) = \frac{\text{Cov}(X_t, X_{t+h})}{\sqrt{\text{Var}(X_t)\text{Var}(X_{t+h})}} = \frac{\gamma(h)}{\gamma(0)}$$

Fonction d'autocorrélation empirique :

$$\hat{\rho}(h) = \frac{T}{T-h} \frac{\sum_{t=h+1}^T (X_t - \bar{X}_T)(X_{t-h} - \bar{X}_T)}{\sum_{t=1}^T (X_t - \bar{X}_T)^2}$$

Graphe de $\hat{\rho}(h)$: *corrélogramme*
(**ACF** = *auto-correlation function*).

Intérêt pratique du corrélogramme

Proposition

$$\text{Si } X_t = m + \sum_{j=0}^k a_j \varepsilon_{t-j}, \quad \text{alors } \forall h > k, \rho(h) = 0.$$

$$\text{Preuve : } \text{Cov}(X_t, X_{t+h}) = \sum_{i=0}^k \sum_{j=0}^k a_i a_j \text{Cov}(\varepsilon_{t-i}, \varepsilon_{t+h-j}).$$

$$\text{Or } \text{Cov}(\varepsilon_{t-i}, \varepsilon_{t+h-j}) = 0 \quad \text{si } h \neq j - i \text{ (car } (\varepsilon_t) \text{ b.b.)}.$$

Séance 1

1 Introduction

2 Modèles de décomposition

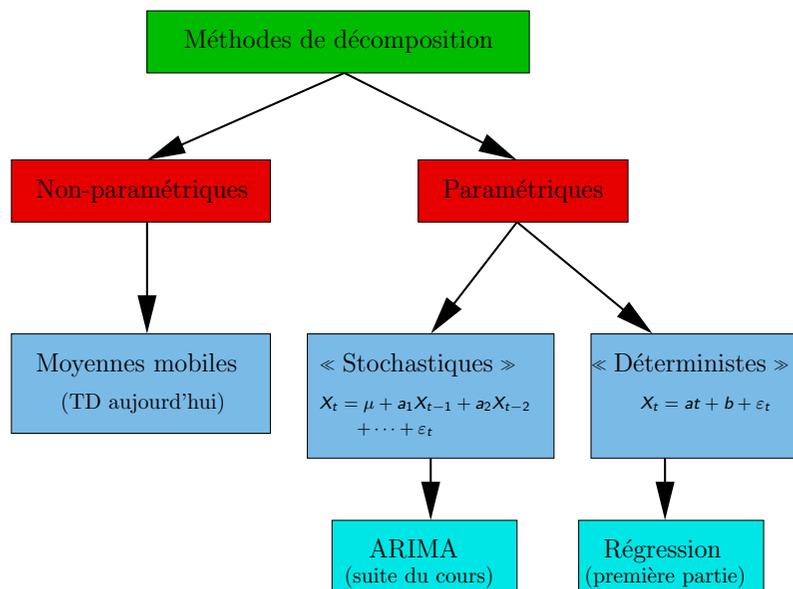
- Les trois composantes
- Tendance
- Composante saisonnière
- Composante résiduelle

3 Décomposition

- Typologie
- Modèles paramétriques
- Filtrage

4 Conclusion

Décomposition des chroniques



Modèles paramétriques déterministes

(X_t) : chronique mensuelle (trimestrielle dans le poly.)

Modèle paramétrique :

$$X_t = \underbrace{a_1 + a_2 t}_{T_t} + \underbrace{\sum_{i=1}^{12} b_i \delta_i(t)}_{S_t} + u(t)$$

δ_i : indicatrice du mois i .

Estimation des paramètres aux moindres carrés des résidus :

$$\begin{cases} \min_{a_i, b_j} \sum_{t=1}^T \left(X_t - a_1 - a_2 t - \sum_{i=1}^{12} b_i \delta_i(t) \right)^2 \\ \text{t.q. } \sum_{i=1}^{12} b_i = 0 \end{cases}$$

→ table de Buys-Ballot.

Filtrage par moyennes mobiles

Définition : moyenne mobile

Soit (X_t) une chronique, et
 $Y_t = M(X_t)$ la chronique telle que

$$\forall t, Y_t = \sum_{i=-m_1}^{m_2} \theta_i X_{t+i}$$

(Y_t) est obtenue de (X_t) par *filtrage par moyenne mobile*.

Remarque : M est un opérateur linéaire.

Décomposition d'une chronique par filtrage

Premier objectif : décomposition de $X_t = T_t + S_t + u_t$.

On cherche un filtre tel que

- $M(T_t) = T_t$;
- $M(S_t) = 0$;
- $M(u_t) \ll \text{aussi petit que possible} \gg$.

de sorte que : $M(X_t) \simeq T_t$

Les moyennes mobiles arithmétiques

Filtre symétrique : $m_1 = m_2 (= m)$, $\theta_i = \theta_{-i}$

$$M(X_t) = \sum_{i=-m}^m \theta_i X_{t+i}$$

Propriété

Une moyenne mobile symétrique telle que $\sum_{i=-m}^m \theta_i = 1$
conserve les chroniques affines $X_t = at + b$.

Propriété

Parmi ces moy. mob., celles qui « minimisent » $M(u_t)$ avec
 (u_t) réalisation d'un b.b. faible (ε_t) vérifient : $\forall i, \theta_i = \frac{1}{2m+1}$.

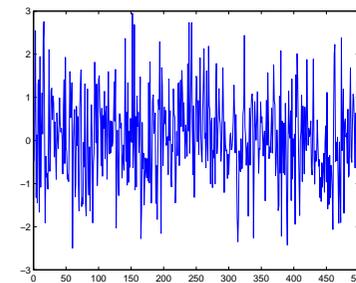
Preuve : $\text{Var}(M(\varepsilon_t)) = \text{Var}(\sum_{i=-m}^m \theta_i \varepsilon_{t+i}) = \sigma^2 \sum_{i=-m}^m \theta_i^2$,
à minimiser, sous contrainte $\sum_{i=-m}^m \theta_i = 1$.

Définition : moyenne mobile arithmétique

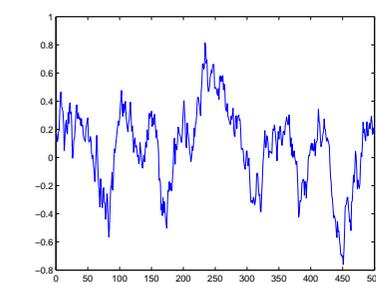
$$\forall i \in \{-m, \dots, m\}, \theta_i = \frac{1}{2m+1}$$

Effet de Slutsky-Yule

Attention à l'apparition éventuelle d'une périodicité
« artificielle ».



bruit blanc gaussien, $\sigma = 1$



lissage arithmétique, $2m + 1 = 19$

Remarque : l'écart-type du bruit est divisée par $\sqrt{2m+1}$
($\simeq 4.36$).

Saisonnalité et moyennes mobiles arithmétiques

$$M_{2m+1}(X_t) = \frac{1}{2m+1} (X_{t-m} + X_{t-m+1} + \dots + X_{t+m}).$$

Propriété

Les composantes saisonnières S_t de période $2m+1$ et de moyenne nulle sur une période sont éliminées par M_{2m+1} :
 $M_{2m+1}(S_t) = 0.$

Généralisation : composantes saisonnières de période $2m$:

$$M_{2m}(X_t) = \frac{1}{2m} \left(\frac{1}{2} X_{t-m} + X_{t-m+1} + \dots + X_{t+m-1} + \frac{1}{2} X_{t+m} \right).$$

Discussion détaillée du filtrage : cf traitement du signal.

Modélisation et
prévision

F. Sur - ENSMN

Introduction

Modèles de
décomposition

Les trois composantes

Tendance

Composante

saisonnière

Composante résiduelle

Décomposition

Typologie

Modèles

paramétriques

Filtrage

Conclusion

Récapitulatif

Hypothèses : (X_t) chronique à décomposer sous la forme :

$$X_t = T_t + S_t + u_t,$$

avec

- T_t linéaire
- S_t de période p et moyenne nulle sur la période.

Alors :

- $M_p(X_t) = M_p(T_t) + M_p(S_t) + M_p(u_t) \simeq T_t.$
- $M_k(S_t + u_t) \simeq S_t$ si $k \ll p$
Problème : pas forcément de moyenne nulle sur une période.
- $M_k(S_t + u_t) - M_p(M_k(S_t + u_t))$ est de moyenne nulle.

Modélisation et
prévision

F. Sur - ENSMN

Introduction

Modèles de
décomposition

Les trois composantes

Tendance

Composante

saisonnière

Composante résiduelle

Décomposition

Typologie

Modèles

paramétriques

Filtrage

Conclusion

29/34

30/34

Un algorithme de décomposition par filtrage

Décomposition de $X_t = T_t + S_t + u_t.$

- 1 périodogramme, connaissances \rightarrow période p de $S_t.$
- 2 estimation de la tendance : $T_t = M_p(X_t)$
- 3 estimation de $\Sigma_t = S_t + u_t$ par $\Sigma_t = X_t - T_t$
- 4 estimation de la composante saisonnière : $S_t = M'(\Sigma_t)$
avec $M'(\Sigma_t) = M_k(\Sigma_t) - M_p M_k(\Sigma_t)$ et $k \ll \text{petit}$
- 5 estimation de la série *corrigée des variations saisonnières* (cvs) : $X'_t = X_t - S_t$
- 6 composante résiduelle $u_t = X'_t - T_t$

Modélisation et
prévision

F. Sur - ENSMN

Introduction

Modèles de
décomposition

Les trois composantes

Tendance

Composante

saisonnière

Composante résiduelle

Décomposition

Typologie

Modèles

paramétriques

Filtrage

Conclusion

Généralisation

- moyennes mobiles non-arithmétiques pour tendance de degré $> 1.$
- filtrage itératif

Remarque : c'est l'idée du programme **Census X11** (1965).

\rightarrow tendance, saisonnalité, données c.v.s., résidus.

Modélisation et
prévision

F. Sur - ENSMN

Introduction

Modèles de
décomposition

Les trois composantes

Tendance

Composante

saisonnière

Composante résiduelle

Décomposition

Typologie

Modèles

paramétriques

Filtrage

Conclusion

31/34

32/34

Séance 1

- 1 Introduction
- 2 Modèles de décomposition
 - Les trois composantes
 - Tendance
 - Composante saisonnière
 - Composante résiduelle
- 3 Décomposition
 - Typologie
 - Modèles paramétriques
 - Filtrage
- 4 Conclusion

Conclusion

Décomposition (additive / multiplicative / mixte) d'une chronique :

- tendance,
- composante saisonnière (période?),
- composante résiduelle / irrégulière (processus stationnaire).

Dans ce cours :

- filtrage (moyennes mobiles, *Census X11*)
→ en TD aujourd'hui.
- processus ARIMA
→ jusque la fin du cours.