

LA PROCÉDURE SMOOTH: une boîte à outils pour le lissage de séries temporelles

Dominique LADIRAY

L'analyse des séries temporelles pose dans la pratique un certain nombre de problèmes que les logiciels généraux de traitement ne peuvent que difficilement résoudre. Dans ce cas, le statisticien est amené à construire son propre outil d'analyse, ce qui ne peut se faire qu'à travers une programmation souvent assez fastidieuse. Je voudrais tout d'abord faire quelques remarques, de bon sens me semble-t-il, sur les problèmes rencontrés dans l'analyse des séries temporelles, avant de présenter une boîte à outils intégrée dans la version 6 du logiciel SAS et dont le but est d'aider l'analyste à construire son propre lisseur de série.

Quelques réflexions sur l'analyse des séries temporelles

Chaque année, un certain nombre d'étudiants de l'ENSAE, dans le cadre de stages ou de groupes de travail, ont à étudier des séries temporelles. Lorsqu'ils viennent me voir, je suis toujours très étonné de la simplicité de leur problème (au moins dans son exposé) et du peu de réponses utilisables de la théorie (au sens où il existe un programme mettant en œuvre la réponse éventuelle de la théorie).

Les exemples suivants vont essayer d'illustrer ces propos.

Il n'y a pas que la prévision dans la vie du statisticien !

La théorie statistique a fait, et fait encore, des gros progrès pour le problème de la prévision en séries temporelles, problème qui repose sur l'existence... d'un futur ou d'un passé ! Cette définition des séries temporelles (grandeur indexée par un paramètre appartenant à un ensemble ordonné non borné) est malheureusement restrictive : une pyramide des âges est, de mon point de vue, une série temporelle. À l'évidence, ce type de série ne relève pas du problème de la prévision qui reviendrait à chercher combien

il existe, à un instant donné, dans un pays, de personnes de 150 ans (1). Par contre, le statisticien peut être amené ici à *lisser* sa série pour en connaître l'allure générale. Le lissage est en effet un problème à part entière, au même titre que la prévision et la désaisonnalisation. La théorie est assez prolixe sur les méthodes de lissage mais, si on se tourne vers les logiciels généraux de traitement des séries temporelles, on s'aperçoit qu'il n'existe que très peu de choses : quelques logiciels proposent des lissages par "fonction spline" et encore ce phénomène est-il récent. SAS, par exemple, ne propose aucune procédure de lissage. Soyons juste : vous pouvez lisser graphiquement votre série grâce à la **PROC GPLOT...**, mais vous ne pouvez pas récupérer les valeurs de la série lissée ! Vous pouvez aussi obtenir une estimation de la tendance de votre série en utilisant une méthode de désaisonnalisation célèbre mise en œuvre dans la **PROC X11**. Mais, d'une part, rien n'oblige votre série à être saisonnière et, d'autre part, les méthodes de désaisonnalisation sont faites pour estimer la composante saisonnière et non la tendance : rien ne garantit que la séparation tendance-bruit obtenue à cette occasion soit "correcte".

Il n'y a pas que des séries mensuelles et trimestrielles !

Voilà encore une vérité de La Palice que les créateurs de logiciels, et les théoriciens, ont mis quelque temps à percevoir. Il est alors assez difficile de désaisonnaliser une série hebdomadaire ou journalière; la grande mode de l'étude des cours de la bourse va sans doute amener de nombreux chercheurs à faire progresser les choses en la matière mais, là encore, ce sera dans une optique de prévision, non de désaisonnalisation.

Une étudiante, en stage dans un grand magasin, cherchait à mesurer, sur une série hebdomadaire de chiffre d'affaires, l'impact de ventes promotionnelles. Aucun logiciel ne lui a permis de résoudre ce problème bien simple et il lui a fallu désaisonnaliser "à la main" sa série, en utilisant conjointement des moyennes mobiles et des méthodes de régression.

Il n'y a pas UN modèle !

C'est carrément ici un problème de fond. Beaucoup de méthodes reposent sur le fait que chaque série peut être représentée par UN modèle qu'il s'agit alors de déterminer. Ainsi on va supposer, par exemple, que la série se décompose en LA tendance, plus LA saisonnalité plus LE bruit. On essaiera ensuite de bâtir une méthode d'estimation "optimale" de chacune de ces composantes. Ces méthodes sont nombreuses : la

1. Pour nuancer mon propos, je dois reconnaître que la théorie de la prévision peut être utile même dans ce cas, pour estimer des données manquantes.

tendance pourra être modélisée par un polynôme, estimé alors par des méthodes classiques de régression, par un modèle ARIMA, ou estimée par des méthodes non-paramétriques (moyennes mobiles...).

Les programmes, lorsqu'ils existent, vous permettent alors une chose merveilleuse : vous rentrez votre série et ils vous donnent LA décomposition.

Je crois, là encore, qu'il est bon de rappeler quelques vérités de simple bon sens. *Une série n'existe pas en tant que telle; elle ne prend sa réelle signification qu'accompagnée d'un problème* : J'ai vu la "même série" (c'est-à-dire la même succession de chiffres) traitée par les Comptes trimestriels et par la Conjoncture et, croyez-moi, les décompositions finales étaient loin d'être semblables.

La désaisonnalisation est, mathématiquement, un problème très simple, qui s'écrit par exemple :

$$\forall t \in \{1, T\} X_t = T_t + S_t + \varepsilon_t$$

C'est donc un système d'une équation à trois inconnues qui donc admet *a priori* une infinité de solutions. L'art du statisticien est de choisir parmi cette infinité de solutions une "sous-infinité" de solutions pas trop absurdes et c'est un sacré travail ! Bien évidemment, le caractère admissible de la solution dépend du problème initial : un conjoncturiste choisira plutôt parmi les solutions présentant des tendances pas trop lisses, contrairement à un historien.

On comprend bien qu'il ne peut pas exister de "définition mathématique" des composantes puisque cette définition est intimement liée au problème que cherche à résoudre le statisticien, problème généralement impossible à traduire en équations qui permettraient de chercher des solutions contraintes d'avoir un rapport avec la question.

Ce qui est plus grave, c'est que toutes les méthodes proposées conduisent à UNE solution nécessairement, donc en faisant des hypothèses fortes sur les composantes. Il me paraît clair que tout programme qui ne vous permet pas d'agir sur ces hypothèses, et ce de façon aisée, est dangereux. Bien sûr, en général ces hypothèses ne sont pas choisies au hasard et ont été testées longuement *empiriquement* ce qui fait que ces programmes donnent "statistiquement" (c'est-à-dire dans la grande majorité des cas) des solutions admissibles. Ainsi, le programme X11-ARIMA est "réglé" sur un ensemble de séries économiques canadiennes. Certaines hypothèses ont d'ailleurs été revues assez récemment. Rien ne permet de dire *a priori* que ces réglages sont corrects pour le grand magasin qui cherche à désaisonnaliser son chiffre d'affaires !

En conclusion de ce paragraphe, il m'a semblé qu'il y avait un besoin d'un programme qui permette à tout utilisateur de construire ses propres lisseurs à partir d'un ensemble

de lisseurs de base dont les propriétés sont bien connues; c'est ce qu'essaye de fournir la PROC SMOOTH.

Les lisseurs de base

L'utilisateur peut composer son propre outil à partir de cinq types de lisseurs de base : les moyennes mobiles, les médianes mobiles, les fonctions spline, un filtre "passe-bas" et la "régression LOWESS".

Il n'est pas question ici de détailler ces lisseurs mais d'exposer simplement leur principe et, le cas échéant, leurs principaux avantages et inconvénients. Le lecteur intéressé trouvera une description plus précise de ces outils dans la note (1).

Les moyennes mobiles

C'est sans doute le lisseur le plus connu. Son principe est simple : vous remplacez la valeur observée à l'instant t par une moyenne pondérée de la valeur présente et de valeurs passées et futures. La forme générale d'une valeur transformée par moyenne mobile est donc, si X_t désigne la valeur de la série à l'instant t :

$$X_t^* = \sum_{i=-m_1}^{i=m_2} \theta_i X_{t+i}$$

Dans ce cas, la moyenne mobile dépend de plusieurs paramètres :

- l'ordre de la moyenne mobile, c'est-à-dire le nombre de points sur lesquels porte la moyenne (m_1+m_2+1). Il est assez intuitif que le pouvoir "lissant" de la moyenne mobile sera d'autant plus fort que son ordre sera important ;
- les coefficients $\{\theta_i\}$.

On montre qu'il est possible de choisir le système de coefficients de telle sorte que la moyenne mobile possède certaines propriétés (voir par exemple "les fonctions spline"). Ainsi, la moyenne mobile d'ordre 5, dont les coefficients sont égaux à $(1/8, 1/4, 1/4,$

(1). Savez-vous que les moyennes mobiles utilisées dans X11 et ses avatars n'ont jamais été revues depuis ... 25ans!

1/4, 1/8) élimine les saisonnalités de période 4 tout en conservant les tendances linéaires, ce qui en fait un outil de base dans la désaisonnalisation des séries trimestrielles.

En théorie, il est donc possible de calculer la moyenne mobile qui convient "le mieux" à son problème. En pratique, de nombreux types de moyennes mobiles ont déjà été étudiés et certaines se révèlent plus efficaces que d'autres; les plus connues sont intégrées dans la **PROC SMOOTH** : moyennes de Henderson, de Spencer...

Ces moyennes souffrent essentiellement de deux défauts. Tout d'abord, comme toute moyenne, elles sont peu robustes et sensibles donc à la présence de "points aberrants". Ensuite, et c'est leur défaut majeur, leur utilisation entraîne une perte d'information au début et en fin de série puisque certaines valeurs ne pourront être lissées faute de suffisamment de points dans le passé ou le futur. À l'heure actuelle, il faut bien le reconnaître, il n'y a pas de solution satisfaisante à ce problème.

Nombre de logiciels de désaisonnalisation, dont X11 et X11-ARIMA, sont construits sur de tels outils.

Les médianes mobiles

Leur principe est assez semblable à celui des moyennes mobiles: au lieu de synthétiser un ensemble de valeurs de la série par une moyenne pondérée, on les résumera par leur médiane. Ces outils ont été remis au goût du jour par J.W Tukey dans la fin des années 60 .

Elles ne souffrent pas des défauts des moyennes mobiles. La médiane est en effet robuste par nature et n'est donc que peu sensible à la présence de "points aberrants". D'autre part, il existe des procédures naturelles pour reconstituer les début et fin de série.

Ce sont donc, par construction, des lisseurs aux propriétés complémentaires des moyennes mobiles. Un logiciel de désaisonnalisation, SABL, développé par le "Bell's Laboratory" repose d'ailleurs sur une association astucieuse de ces outils de base.

Les fonctions spline

L'idée de ce type de lisseur est assez simple: on découpe *a priori* la série en plusieurs tranches et, sur chaque partie, on ajuste un polynôme de degré fixé à l'avance. *A priori*, la fonction obtenue, polynomiale par morceaux, présentera des points de discontinuité ou de non-dérivabilité à chaque extrémité de tranche. Il est néanmoins possible de faire ce même travail sous la contrainte que le "raccordement se passe bien", c'est-à-dire que

la fonction résultante soit dérivable en tout point, ce qui lui assurera un "aspect lisse". L'utilisateur peut donc "régler" son lisseur grâce au découpage retenu et aussi grâce au degré des polynômes.

En pratique, la série est découpée en tranches de longueur quasi-égales et les polynômes sont choisis de même degré, souvent 3, d'une part pour des raisons de calcul numérique et aussi pour la grande variété de types de tendance que ce choix permet de restituer.

Lisseur "passe-bas" basé sur la transformée de Fourier

Dans la **PROC SMOOTH**, ce lisseur agit en plusieurs étapes :

- la partie linéaire de la série est enlevée ;
- la série résiduelle est analysée par la transformée de Fourier, ce qui permet de repérer les hautes fréquences qui rentrent dans sa composition ;
- ces hautes fréquences sont alors enlevées, dans la mesure où elles sont supposées correspondre aux saisonnalités ou au bruit. C'est en fait à ce moment que l'utilisateur peut régler la force de son filtre : plus il enlèvera de fréquences, plus la courbe résultante sera lisse ;
- la partie linéaire précédemment enlevée est alors rajoutée à la partie "basses fréquences" pour conduire à la courbe lissée.

La régression LOWESS

Cette méthode est due à W.S. Cleveland et repose sur le principe suivant : à chaque instant t , on considère un ensemble de p valeurs de la série autour de cet instant sur lesquelles on ajuste une droite ; la valeur prédite par cette droite à l'instant t sera la valeur lissée. Tout le "secret" de cette méthode résulte évidemment dans la façon dont on trouve la droite :

- une première régression pondérée est faite sur les p points les plus proches de l'instant t , le poids de chaque point décroissant avec son éloignement de l'instant t ;
- une seconde régression pondérée est faite sur ces p points avec un poids qui décroît cette fois avec leur éloignement de la première droite ajustée. Ce procédé est itéré afin de déterminer le système de poids.

Ce lisseur, très gourmand en temps de calcul, a été utilisé par le Bell's Laboratory pour mettre au point le logiciel de désaisonnalisation STL.

La procédure SMOOTH

La procédure **SMOOTH** est développée en Fortran et intégrée dans la version 6 du logiciel SAS grâce au module **SAS/TOOLKIT**.

Actuellement, tous les lisseurs décrits ci-dessus ne sont pas encore disponibles. La procédure **SMOOTH** n'imprime rien mais crée en sortie un tableau SAS comprenant les séries brutes et les séries lissées que l'on peut alors visualiser et traiter par les autres procédures SAS (**PLOT**, **PRINT**...).

Syntaxe de la procédure

```
PROC SMOOTH / <option>  
VAR liste de variables;  
COPY liste de variables;  
MAVERAGE nom = liste de coefficients /<option> ;  
SMOOTHER liste de lisseurs élémentaires /<option> ;
```

Options de l'instruction PROC

Trois options sont disponibles à ce niveau :

- l'option **DATA** = nom de table SAS qui permet de préciser le nom de la table SAS contenant les séries à lisser. Si cette option est absente, c'est la dernière table SAS créée qui sera utilisée par défaut ;
- l'option **OUT** = nom de table SAS qui permet d'attribuer un nom à la table SAS qui sera **AUTOMATIQUEMENT** créée par la procédure. Si cette option est absente, la table sera créée dans la base temporaire **WORK** avec un nom de type **DATAn** ;
- une option **NOLABEL**: dans la table SAS créée, les noms des séries lissées sont générés automatiquement (voir plus loin) et peuvent ne pas être parlants. À chaque série lissée est associé par défaut un label. Si par exemple le nom de votre série est **INDICE** et que vous la lissez par une moyenne mobile de **HENDERSON** sur 9 termes nommée **HEND9**, le label associé à la série lissée sera **INDICE(HEND9)**. Ce label pourra être utilisé ultérieurement dans d'autres procédures et apparaît notamment

lorsque vous faites une **PROC CONTENTS**. Vous pouvez décider de ne pas créer ce label en précisant **NOLABEL**, ou plus simplement **NOL**;

Instruction VAR

L'instruction **VAR** permet de préciser le nom des séries que l'on veut lisser. La procédure ne traite, par nature, que des variables numériques. Par conséquent, lorsque l'instruction **VAR** est présente, elle doit préciser une liste de noms valides de variables SAS numériques de la table d'entrée.

Lorsque cette instruction est absente, la procédure lisse TOUTES les variables numériques de la table SAS d'entrée.

Instruction COPY

Cette instruction permet de récupérer dans la table de sortie, des variables de la table d'entrée autres que les séries à lisser de l'instruction **VAR**. Vous pouvez ainsi recopier par exemple une variable "date" qui permettra le tracé de courbes par **PROC PLOT**.

Les variables de l'instruction **VAR** se retrouvant systématiquement dans la table en sortie, vous NE DEVEZ PAS les repréciser dans l'instruction **COPY** (erreur syntaxique sinon).

Instruction MAVERAGE

La procédure **SMOOTH** permet d'utiliser directement toute une série de moyennes mobiles qui sont "pré-programmées" (voir plus loin).

Vous pouvez néanmoins définir vos propres moyennes mobiles en utilisant pour chaque nouvelle moyenne une instruction **MAVERAGE**.

La syntaxe générale de l'instruction **MAVERAGE** est :

MAVERAGE nom = liste de coefficients / option ;

vous pouvez écrire indifféremment **MAVERAGE** ou **MAV**.

- c'est par ce nom que vous pourrez dans la suite de la procédure faire référence à ce nouveau lisseur. Ce nom doit être un nom SAS valide et doit être différent des noms des lisseurs pré-programmés (voir plus loin).
- cette liste de coefficients permet de définir précisément la moyenne mobile.

Vous devez tous les donner, même si votre moyenne est symétrique, ceci justement pour permettre l'utilisation de moyennes mobiles non symétriques dans les coefficients. La procédure norme automatiquement la somme des coefficients et ne permet d'utiliser et de définir que des moyennes mobiles centrées (c'est-à-dire faisant intervenir autant de points dans le passé et dans le futur). Seule l'option **NOCENTER** ci-dessous permet de moduler légèrement cette contrainte.

- Dans le cas où vous ne souhaitez pas centrer vos moyennes mobiles d'ordre pair, vous devez préciser l'option **NOCENTER** (ou **NOC**) au moment où vous la définissez. Ainsi l'instruction:

MAVERAGE M4T = 1 1 1 1 / NOC;

crée une moyenne mobile sur 4 termes définie par

M4T = 0.25 0.25 0.25 0.25 ;

et dans ce cas, on a :

M4T (X_t) = 0.25 X_{t-1} + 0.25X_t + 0.25 X_{t+1} + 0.25 X_{t+2}

(l'opération de centrage revient à appliquer à toute moyenne mobile d'ordre pair la moyenne M2 = 0.5 0.5)

Instruction SMOOTHER

L'instruction (ou les instructions) **SMOOTHER** permet de définir le (ou les) lisseurs que vous désirez appliquer à vos séries.

Ces lisseurs peuvent être composés de façon quelconque à partir d'un ensemble de lisseurs élémentaires: moyennes mobiles (définies par une instruction **MAVERAGE** ou pré-programmées), médianes mobiles, opérateurs spécialisés (Twicing, Splitting), fonction spline...

La syntaxe générale de l'instruction **SMOOTHER** est :

SMOOTHER liste de lisseurs élémentaires / <options> ;

Les lisseurs élémentaires utilisables sont :

- les **MOYENNES MOBILES** qui sont caractérisées par un nom, celui donné dans une instruction **MAVERAGE** si vous avez défini vous-même cette moyenne, ou celui associé à une moyenne pré-programmée ;
- les **MÉDIANES MOBILES** caractérisées par leur ordre: '3' désigne ainsi la médiane mobile sur 3 termes. Si vous souhaitez répéter cette médiane mobile jusqu'à ce que la série lissée soit invariante par ce lisseur, vous utiliserez l'opérateur R en codifiant '**3R**'. L'opérateur R ne peut être utilisé qu'avec une médiane mobile ;
- le **SPLITTING** , soit seul en codifiant S, soit doublé, en codifiant '**SS**', soit répété jusqu'à stabilité en codifiant '**SR**';
- les **FONCTIONS SPLINE**, mises en œuvre en codifiant '**SPxx**' où **xx** désigne le nombre de "tronçons" à faire et sur lesquels des polynômes de degré 3 seront ajustés. Le "filtre passe-bas" utilisant la transformée de Fourier est mis en oeuvre en codifiant '**SMxx**' où **xx**, variant de 00 à 99, permet de régler l'intensité du lissage: 00 ne lisse pas la série alors que 99 la transforme en droite ;

La **RÉGRESSION LOWESS** est mise en oeuvre par '**LOxx**' où **xx** désigne le pourcentage de points de la série intervenant dans chaque régression locale.

Trois options sont disponibles pour cette instruction :

- l'option **INT** qui vous permet de récupérer dans la table en sortie l'ensemble des résultats des lissages intermédiaires ;
- l'option **TWICE** (ou **TW**) qui permet d'obtenir la version "twicée" du lisseur dans le cas de mélange de moyennes et médianes mobiles;
- l'option **END** qui permet d'obtenir des estimations des points de début ou de fin de séries lorsqu'on utilise un lisseur qui "perd de l'information".

Exemple :

SMOOTHER '4' '3R' 'SR' '2' 'H' /int tw;

Dans ce cas, chaque série sera lissée de la façon suivante :

- on applique une médiane mobile sur 4 termes ('4') ;
- puis une médiane mobile sur 3 termes jusqu'à stabilité ('3R') ;

- puis le SPLITTING jusqu'à stabilité de la série lissée ('SR');
- puis une médiane mobile d'ordre 2 ;
- et enfin une moyenne mobile pré-programmée, de nom H ('H').

Remarquez que chaque lisseur élémentaire doit figurer entre quotes ; il est alors facile de faire la différence entre: '4' '2' et '42'. Les options **INT** et **TW** permettent de récupérer dans la table de sortie les résultats des lissages par : **4, 43R, 43RSR, 43RSR2, 43RSR2H et 43RSR2H, twice.**

Moyennes mobiles pré-programmées

Vous pouvez utiliser directement, c'est-à-dire sans les définir, 20 moyennes mobiles qui ont été pré-programmées Il s'agit des moyennes suivantes :

- moyennes mobiles arithmétiques simples centrées :

Elles sont notées Mx où x désigne l'ordre de la moyenne et peut varier de 3 à 12. Par exemple, M5 désigne la moyenne mobile arithmétique simple de coefficients (1/5 , 1/5 , 1/5 , 1/5 , 1/5).

Ces moyennes mobiles simples pré-programmées sont centrées et par conséquent, la moyenne M4 est définie par les coefficients (1/8 , 1/4 , 1/4 , 1/8) et non par les coefficients (1/4 , 1/4 , 1/4 , 1/4); pour utiliser cette dernière moyenne, il faudrait la définir dans une instruction **MAVERAGE** avec en outre l'option **NOCENTER**.

- moyennes mobiles de HENDERSON sur 5, 9, 13 ou 15 points :
Ces moyennes sont appelées respectivement **HEND5, HEND9, HEND13 et HEND15** ;
- moyennes mobiles de SPENCER sur 15 et 21 points :
Ces moyennes sont appelées respectivement **SPEN15 et SPEN21**;
- et 4 autres moyennes mobiles.

Le HANNING et les moyennes X et Y utilisées fréquemment dans le cas de lissage par médianes mobiles - voir 5). Ces moyennes sont notées respectivement H, X et Y. Et enfin, une moyenne dite "binomiale" sur 15 termes notée **BINO15**. Pour utiliser ces moyennes, il suffit de les nommer dans une instruction **SMOOTHER**. Ainsi le programme :

```
PROC SMOOTH DATA=NEW OUT=LISS;  
VAR BANK;  
SMOOTHER HEND9;
```

permet le lissage de la variable **BANK** par la moyenne mobile de HENDERSON sur 9 termes.

Structure de la table SAS en sortie

La table de sortie contient d'une part les variables de l'instruction **COPY**, d'autre part les séries brutes - spécifiées dans l'instruction **VAR-**, et enfin les séries lissées correspondant à chaque "smoother" - avec ou sans les lissages intermédiaires selon la présence de l'option **INT**.

Le nom des séries brutes est celui de la variable concernée.

Le nom des séries lissées est de la forme **VnSm** ou **VnSm_p** :

- n désigne le numéro de la série brute (ordre de la variable SAS concernée dans la liste **VAR**).
- m désigne le numéro de l'instruction **SMOOTHER** définissant le lisseur, ces instructions étant numérotées dans l'ordre où elles apparaissent dans le programme.
- p désigne le numéro du lisseur intermédiaire dans le cas où l'option **INT** a été spécifiée.

Si l'option **NOLABEL** n'est pas précisée à l'appel de la procédure, un label est créé pour chaque série lissée. Ce label est de la forme :

nomvar (nomliss)
où nomvar est le nom de la série brute et
nomliss le nom du lisseur appliqué.

Exemple :

```
PROC SMOOTH DATA=DONNEES OUT=LISS;  
VAR SER1 D11;  
SMOOTHER 'HEND9';  
SMOOTHER '3R' H/INT TW;
```

La table en sortie contiendra les variables:

NOM	LABEL
SER1	
V1S1	SER1(HEND9)
V1S2_1	SER1(3R)
V1S2_2	SER1(3RH)
V1S2_3	SER1(3RH,tw)
D11	
V2S1	D11(HEND9)
V2S2_1	D11(3R)
V2S2_2	D11(3RH)
V2S2_3	D11(3RH,tw)

BIBLIOGRAPHIE

- *Exploratory data analysis* de J.W. TUKEY, 1977, Addison Wesley.
- "Moderns Methods of data analysis" de J. FOX et J.S. LONG, 1990. *Sage publications*.
- "Séries temporelles et modèles dynamiques" de C. GOURIEROUX et A. MONFORT, *Économica* 1990.
- "Locally weighted regression: an approach to regression analysis by local fitting" W.S. CLEVELAND et S.J. DEVLIN, *Jasa*, 1988.
- "Lissage robuste de séries chronologiques : Une étude empirique" D. LADIRAY et N. ROTH, *Annales d'Économie et de Statistique*, 1987.

