MOYENNES MOBILES CENTRÉES ET NON CENTRÉES CONSTRUCTION ET COMPARAISON

Michel GRUN-REHOMME, Dominique LADIRAY

L'analyse des séries temporelles a, de toute évidence, fait de gros progrès depuis une vingtaine d'années. Dans sa panoplie de méthodes pour aborder les problèmes du lissage et de la désaisonnalisation, le statisticien dispose aujourd'hui d'un outil un peu anachronique, qui résiste bien à l'usure du temps et aux innovations statistiques : **les moyennes mobiles**. Le succès de cet outil est essentiellement dû d'une part à son excellent rapport "qualité-prix" et, d'autre part, à l'hégémonie des logiciels de désaisonnalisation X11 et X11-ARIMA qui en font un large usage.

Les moyennes mobiles sont en effet très simples de principe, n'impliquent pas *a priori* l'utilisation de concepts ou de modèles sophistiqués et se révèlent d'application particulièrement souple: il est possible de construire une moyenne mobile possédant les propriétés souhaitées en termes de conservation de tendance, d'élimination de la saisonnalité, de réduction du bruit, etc, et s'adaptant ainsi au problème traité. Bâti sur de tels outils, le mythique logiciel de désaisonnalisation X11 défie le temps. Aujourd'hui encore, c'est une version de 1968 qui est utilisée et les améliorations importantes apportées à ce programme, notamment à travers le logiciel X11-ARIMA de Statistique Canada dans les années 75, n'en ont pas remis en cause le principe de base.

Force est de constater qu'aujourd'hui ces outils sont très largement utilisés, à tort selon certains qui pensent que de meilleures méthodes existent de nos jours, mais à raison selon les utilisateurs qui valident les résultats et emploient par exemple X11-ARIMA pour sa faculté de désaisonnaliser rapidement et correctement un grand nombre de séries.

Les moyennes mobiles sont de vieilles dames. De très importants efforts de recherche sur ce thème ont été faits au début de ce siècle par des noms aujourd'hui célèbres: Spencer, Henderson, Macaulay... et des résultats aujourd'hui oubliés ont été obtenus. Qui sait par exemple, que le fameux problème de la perte d'information aux extrémités de la série, et évoqué comme l'un des inconvénients majeurs de ces outils, a été étudié et en partie résolu par Henderson dans les années 20? Et, par la suite, Macaulay, Kendall, Musgrave (pour le logiciel X11), Bongard ... s'y sont attaqué avec plus ou moins de succès.

Dans un premier temps, cet article présente une synthèse des propriétés connues des moyennes mobiles. Celles-ci sont présentées dans un cadre méthodologique général plus actuel: leur mode de construction est vu comme un problème de minimisation d'une forme quadratique sous contraintes. Les moyennes mobiles traditionnelles, symétriques ou non-centrées sont alors replacées dans ce cadre et comparées entre elles.

Cette présentation unifiée permet alors de généraliser les modes de construction et d'association et de déduire quelques résultats nouveaux. Le rôle central des critères de Bongard (réduction de la variance résiduelle) et de Henderson (pouvoir de lissage) est mis en évidence. On montre ainsi l'équivalence des approches de Kendall et de Bongard. Un nouveau critère, mélange convexe des deux critères précédents, est défini et étudié. Les moyennes mobiles asymétriques de Musgrave sont généralisées, améliorées et des règles de construction de moyennes mobiles non-centrées, permettant de résoudre le problème de l'estimation des points aux extrémités des séries, sont présentées.

Ce travail s'inspire bien entendu des résultats obtenus par les grands noms déjà cités mais aussi de travaux plus récents, comme ceux de DOHERTY [2] et GRAY et THOMSON [5].

I - Rappels et notations

Ce chapitre est une brève présentation des notions de base sur les moyennes mobiles. Pour un exposé plus complet, le lecteur pourra consulter KENDALL [6] ou, pour un ouvrage en français, GOURIEROUX et MONFORT [4].

Dans la suite, on considèrera une série temporelle (X_t). X_t désignera alors la valeur de la série à l'instant t. $L(X_t)$ désigne la transformée de la série brute par un opérateur L et LX_t sera la valeur de cette nouvelle série à l'instant t.

Par ailleurs, on fera souvent référence aux problèmes du lissage ou de l'ajustement saisonnier dans lesquels la série brute est supposée *a priori* se décomposer additivement :

- en une tendance (T_t) et un bruit (ε_t) pour le lissage: $X_t = T_t + \varepsilon_t$
- en une tendance (T_t), une saisonnalité (S_t) et un bruit ($_t$) pour la désaisonnalisation: $X_t = T_t + \varepsilon_t$.

I.1 Moyennes mobiles symétriques et asymétriques

On appelle **moyenne mobile de coefficients $\{\theta_i\}$**, l'opérateur noté MM $\{\theta_i\}$**, ou plus simplement MM, et défini par :**

$$\mathrm{MM}\ \{\theta_i\}\ X_t = \sum_{i=-p}^{+f} \theta_i\ X_{t+i}$$

La valeur à l'instant t de la série brute est donc remplacée par une moyenne pondérée de p valeurs "passées" de la série, de la valeur actuelle et de f valeurs "futures" de la série. La quantité p+f+1 est appelée **ordre de la moyenne mobile.**

Il est clair, pour des raisons de définition, que les p premiers points et les f derniers points de la série brute ne peuvent être transformés par l'opérateur MM. Lorsque p est égal à f, la moyenne mobile est dite **centrée**.

Si, en outre, on a $\theta_{-i} = \theta_i$ pour tout i, la moyenne mobile MM est dite **symétrique**.

Par la suite, on notera le vecteur de dimension (p+f+1,1) dont les coordonnées sont les coefficients de la moyenne mobile :

$$\theta = \begin{bmatrix} \theta_{-p} \\ \theta_{-p+1} \\ \cdot \\ \cdot \\ \cdot \\ \\ \cdot \\ \cdot \\ \cdot \\ \theta_{f-1} \\ \theta_f \end{bmatrix}$$

I.2. Propriétés simples des moyennes mobiles

I.2.1. En termes de conservation de tendances :

Il est facile de montrer que pour qu'une moyenne mobile MM conserve les polynômes de degré d, il faut et il suffit que ses coefficients vérifient :

$$\sum_{i=-p}^{i=+f} \theta_i = 1 \quad \text{et } \forall k \in \{1, 2, \ldots\ldots\} \sum_{i=-p}^{i=+f} \theta_i = 0$$

Matriciellement, ces contraintes s'écrivent : $C\theta = \alpha$, où C et α sont des matrices (d+1,p+f+1) et (d+1,1) valant :

$$\alpha = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 0 \end{bmatrix} \quad \text{et} \quad C = \begin{bmatrix} 1 & 1 & \ldots \ldots & 1 & 1 \\ -p & -p+1 & \ldots \ldots & f-1 & f \\ (-p)^2 & (-p+1)^2 & \ldots \ldots & (f-1)^2 & f^2 \\ \ldots & \ldots & \ldots \ldots & \ldots & \ldots \\ (-p)^{d-1} & (-p+1)^{d-1} & \ldots \ldots & (f-1)^{d-1} & f^{d-1} \\ (-p)^d & (-p+1)^d & \ldots \ldots & (f-1)^d & f^d \end{bmatrix}$$

On peut remarquer à cette occasion que toute moyenne mobile symétrique qui conserve les polynômes de degré 2d conserve aussi les polynômes de degré 2d+1.

I.2.2. En termes d'élimination de saisonnalités :

Les saisonnalités sont souvent "modèlisées" par des fonctions périodiques de période k (4 pour une série trimestrielle ....); dans le cas d'un modèle de composition additif, les coefficients saisonniers sont en outre supposés être de somme nulle. Ces fonctions engendrent alors un sous espace vectoriel de dimension k-1 dont il est facile d'exhiber une base. Ainsi, par exemple, dans le cas trimestriel, on trouve le sous espace engendré par :

$$\begin{bmatrix} 1 & -1 & 0 & 0 & 1 & -1 & 0 & 0 & \ldots \\ 1 & 0 & -1 & 0 & 1 & 0 & -1 & 0 & \ldots \\ 1 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & \ldots \end{bmatrix}$$

L'annulation de telles séries introduit donc des contraintes sur les coefficients de la moyenne mobile qui s'expriment matriciellement : $C\,\theta = \alpha$, où C est la matrice (k-1,p+f+1) dont les lignes sont les vecteurs de base ci-dessus et où α est la matrice nulle de dimensions (k-1,1).

Par ailleurs, nous verrons plus loin (II.2.2 et Annexe 1) qu'il est possible de traiter le cas de saisonnalités variant linéairement (ou polynômialement) avec le temps.

I.2.3. en termes de réduction du bruit :

Le résidu, dans la décomposition de la série brute, est souvent modélisé sous la forme d'un bruit blanc, suite de variables aléatoires (ε_t) d'espérance nulle, non corrélées, et de même variance σ^2. Ce bruit blanc est transformé par la moyenne mobile en une suite de variables aléatoires (ε_t^*), de même variance égale à :

$$\sigma^{*2} = \sigma^2 \sum_{i=-p}^{i=+f} \theta_i^2$$

Diminuer la composante irrégulière revient donc à diminuer la quantité :

$$\sum_{i=-p}^{i=+f} \theta_i^2$$

Cette quantité figurera, dans les tableaux et exemples présentés par la suite, sous le nom de Bongard.

I.3. Quelques effets indésirables des moyennes mobiles

I.3.1 L'effet Slutsky-Yule

La transformation du bruit blanc (ε_t) par une moyenne mobile donne un processus aléatoire (ε_t^*) corrélé dans le temps. En effet, le coefficient d'autocorrélation d'ordre k (k entier naturel non nul) est :

$$\rho(k) = \frac{\mathrm{Cov}(\varepsilon_t^*, \varepsilon_{t+k}^*)}{\theta_t^{*2}} = \frac{\sum_{i=-p}^{i=+f} \sum_{j=-p}^{j=+f} \theta_i \theta_j \mathrm{Cov}(\varepsilon_{t+i}, \varepsilon_{t+k+j})}{\sigma^2 \left(\sum_{j=-p}^{i=+f} \theta_i^2 \right)}$$

$$= \frac{\sum_{i=-p}^{i=+f} \sum_{j=-p}^{j=+f} \theta_i \; \theta_j \, E(\varepsilon_{t+i}, \varepsilon_{t+k+j})}{\sigma^2 \left(\sum_{i=-p}^{i=+f} \theta_i^2 \right)}$$

$$= \begin{cases} \dfrac{\sum_{j=-p}^{j=f-k} \theta_j \theta_{j+k}}{\sum_{i=-p}^{i=+f} \theta_i^2} & \text{si } \; k \leq p+f \\ \text{O sinon} & \end{cases}$$

Cette corrélation entre p+f termes consécutifs du processus (ε_t^*) introduit des oscillations parasites. Il est impossible de les éliminer mais on peut les atténuer en réduisant la variance résiduelle. Par ailleurs, la "périodicité" de ces oscillations est aléatoire; il est cependant possible d'en estimer la valeur moyenne qui figurera, dans les tableaux et exemples présentés par la suite, sous le nom de PÉRIODE.

I.3.2 Effet d'amplitude et effet de phase

La transformation par une moyenne mobile d'une série géométrique d'amplitude variable produit des effets "parasites" d'amplitude et de phase.

Soit la série: $X_t = \rho^t \; e^{i\,\omega\,t}$, où $\omega = \frac{2\pi}{T}$ avec $T > 0$

alors on a, par transformation par la moyenne mobile MM :

$$X^*_t = MMX_t = \sum_{k=-p}^{k=f} \theta_k \, X_{t+k} = \sum_{k=-p}^{k=f} \theta_k \, \rho^k e^{i\omega k} \rho^t \, e^{i\omega t}$$

$$= \left(\sum_{k=-p}^{k=f} \theta_k \, \rho^k e^{i\omega t} \right) X_t$$

Si a et φ désignent respectivement le module et l'argument de $\sum_{k=-p}^{k=f} \theta_k \, \rho^k e^{i\omega k}$, on obtient : $X_t^* = a \; \rho^t e^{i(\omega t + \varphi)}$

On a donc un effet d'amplitude (a) et un effet de phase (φ). L'effet de phase est inhérent aux moyennes mobiles non symétriques, sauf à ajouter, dans leur construction, une contrainte sur leurs coefficients de la forme :

$$\sum_{k=-p}^{k=f} \theta_k \; \rho^k \sin \omega \, k = 0$$

Pour une moyenne mobile symétrique (p=f=m) avec $\rho = 1$, on a :

$$X_t^* = \sum_{k=-p}^{k=f} \theta_k \, e^{i\omega(t+k)} = (\theta_0 + 2 \sum_{k=1}^{k=m} \theta_k \; \cos \omega \, k) \, e^{i\omega t}$$

Dans ce cas, on a $\varphi = 0$, c'est-à-dire pas d'effet de phase, ou $\varphi = \Pi$, c'est-à-dire opposition de phase, et l'effet d'amplitude vaut $\left| (\theta_0 + 2 \sum_{k=1}^{k=m} \theta_k \; \cos \; \omega \, k) e^{i\omega t} \right|$

I.4. Résolution d'un problème de minimisation d'une forme quadratique sous contraintes

Soit θ le vecteur de dimension (p+f+1,1) dont les coordonnées sont les coefficients inconnus de la moyenne mobile recherchée. Soient en outre les matrices connues w , Ω, C , et α, et, de tailles respectives (p+f+1,1), (p+f+1,p+f+1), (k,p+f+1) et (k,1).

Ω est supposée symétrique, définie, positive et C est supposée de plein rang. On s'intéresse à la résolution du problème de minimisation sous contraintes du type :

$$\begin{cases} \underset{\theta}{\mathrm{Min}} & {}^t(\theta - w)\,\Omega\,(\theta - w) \\ \text{sous} & C\theta = \alpha \end{cases}$$

La solution unique de ce problème classique est :

$$\theta = \Omega^{-1}\,{}^tC\,(C\Omega^{-1}\,{}^tC)^{-1}(\alpha - Cw) + w$$

Bien entendu, la résolution de ce problème suppose que l'on n'ait pas plus d'équations que d'inconnues et donc que le nombre de contraintes indépendantes sur les coefficients résumées dans la matrice C (soit k) est inférieur ou égal à l'ordre de la moyenne mobile (soit p+f+1). Dans le cas où il y aurait égalité, la contrainte détermine pleinement la solution et il n'y a plus de problème d'optimisation ; d'ailleurs C est alors une matrice inversible et $\theta = C^{-1}\alpha$.

II - Génération de moyennes mobiles symétriques classiques

II.1. Les moyennes mobiles arithmétiques simples

L'exemple de moyenne mobile le plus simple est classiquement obtenu en supposant que celle-ci conserve les constantes tout en diminuant au maximum l'importance de la perturbation. Cette réduction de la composante irrégulière est mesurée par la quantité $\sum_{i=-p}^{i=+f} \theta_i^2$. Cela équivaut, dans les termes du problème d'optimisation évoqué, à prendre une matrice Ω égale à la matrice identité. Par ailleurs, on a ici :

$$C = [1\ 1 \ldots\ldots 1\ 1], \quad w = \begin{bmatrix} 0 \\ 0 \\ . \\ . \\ . \\ 0 \\ 0 \end{bmatrix} \text{ et } \alpha = [1]$$

On obtient alors les moyennes arithmétiques simples, dont tous les coefficients sont égaux à l'inverse de l'ordre de la moyenne mobile. La symétrie de la solution, obtenue dès que la moyenne recherchée est supposée centrée, permet en plus d'affirmer que ces moyennes mobiles conservent aussi les droites.

On peut aisément vérifier qu'une moyenne mobile arithmétique simple d'ordre m élimine les fonctions périodiques de période m, et qu'elle peut-être obtenue, en supposant toujours que la moyenne des coefficients saisonniers est nulle, avec :

$$C=\begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & -1 & 0 & 0 & \dots & 0 \\ 1 & 0 & -1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & 0 & 0 & \dots & 0 \end{bmatrix} \quad \text{et} \quad \alpha=\begin{bmatrix} 1 \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

C'est cette optique qu'il est préférable de privilégier, comme nous allons le voir immédiatement, si on veut généraliser le mode de construction de ces moyennes.

II.2. Les compositions de moyennes mobiles simples

II.2.1. Les compositions de moyennes mobiles simples de type mxn

Dans le logiciel Census X11, des moyennes composées à partir de moyennes mobiles arithmétiques simples sont utilisées pour estimer la série des coefficients saisonniers; elles sont notées 3x3, 3x5, 3x9 ...ces notations indiquant qu'une moyenne mobile arithmétique simple d'ordre 3 est appliquée à la série puis, sur cette série lissée, une moyenne mobile simple d'ordre 3 ou 5 ou 9.

Il suffit alors, pour construire ces moyennes, de choisir le même critère de réduction de la variance que précédemment en mettant dans la matrice C les contraintes d'annulation de saisonnalités *ad hoc*. On aboutit alors, par exemple, à des moyennes comme celles présentées dans le *tableau 1*.

II.2.2. Annulation de saisonnalités variant polynômialement avec le temps

L'hypothèse de stabilité de la saisonnalité dans le temps est parfois peu justifiée et on est amené à modéliser cette évolution de la façon suivante :

$$S = \left(a_0 + a_1 t^2 + \ldots\ldots\ldots\ldots\ldots + a_d t^d\right) u_t$$

où (u_t) désigne une fonction périodique de période b. Si l'on cherche alors une moyenne mobile d'ordre p+f+1 annulant ce type de saisonnalités, on montre (voir annexe 1 pour un calcul général détaillé) que les coefficients de la moyenne mobile doivent vérifier, en supposant p+f+1=nb :

$$\forall k = f - b + 1, f - b + 2, \;\ldots\ldots\; , f - 1$$

$$\begin{cases} \displaystyle\sum_{j=0}^{j=n-1} \theta_{k-jb} - \sum_{j=0}^{j=n-1} \theta_{f-jb} = 0 \\ \displaystyle\sum_{j=0}^{j=n-1} (k-jb)\, \theta_{k-jb} - \sum_{j=0}^{j=n-1} (f-jb)\, \theta_{f-jb} = 0 \\ \vdots \\ \displaystyle\sum_{j=0}^{j=n-1} (k-jb)^d\, \theta_{k-jb} - \sum_{k=0}^{j=n-1} (f-jb)^d\, \theta_{k-jb} = 0 \end{cases}$$

L'annulation de fonctions périodiques de période 4 variant quadratiquement avec le temps conduira donc à 9 contraintes. Le tableau 1 présente quelques exemples de telles moyennes mobiles composées à partir de moyennes mobiles simples. Trois moyennes notées M4x4x5 figurent dans ce tableau. Les deux premières ont les mêmes propriétés d'élimination des saisonnalités d'ordre 4 variant linéairement avec le temps et les saisonnalités d'ordre 5 : seul leur nombre de termes diffère. La troisième, outre ces mêmes propriétés, conserve les polynômes de degré 3 (voir ci-après).

Mnxp désigne la composée de deux moyennes mobiles arithmétiques simples d'ordre n et d'ordre p, Sp15 désigne la moyenne mobile de Spencer sur 15 termes, conservant les polynômes de degré 3 et annulant les saisonnalités d'ordre 4 variant linéairement avec le temps et d'ordre 5.

La quantité notée Henderson désigne la valeur du critère de Henderson (voir II-4).

Tableau 1 :

Moyennes mobiles simples composées éliminant diverses saisonnalités

i	M3x3	M3x5	M3x7	M4x4x5	M4x4x5 (1)	Sp15	M4x4x5 (2)
- 7						- 0,009	- 0,038
- 6					0,045	- 0,019	0,010
- 5				0,013	0,058	- 0,016	- 0,016
- 4			0,048	0,038	0,083	0,009	0,009
- 3		0,067	0,095	0,075	0,120	0,066	0,123
- 2	0,111	0,133	0,143	0,125	0,080	0,144	0,115
- 1	0,222	0,200	0,143	0,163	0,072	0,209	0,180
0	0,333	0,200	0,143	0,175	0,084	0,231	0,231
1	0,222	0,200	0,143	0,162	0,072	0,209	0,180
2	0,111	0,133	0,143	0,125	0,080	0,144	0,115
3		0,067	0,095	0,075	0,120	0,066	0,123
4			0,048	0,037	0,083	0,009	0,009
5				0,013	0,058	- 0,016	- 0,016
6					0,045	- 0,019	0,010
7						- 0,009	- 0,038
Henderson	0,148	0,036	0,018	0,002	0,066	0,006	0,318
Bongard	0,235	0,164	0,125	0,129	0,084	0,193	0,179
Période	11,030	15,496	18,936	20,107	16,284		11,739

Mn x p désigne la composée de deux moyennes mobiles arithmétiques simples d'ordre n et d'ordre p. Sp15 désigne la moyenne mobile de Spencer sur 15 termes, conservant les polynômes de degré 3 et annulant les saisonnalités d'ordre 4 variant linéairement avec le temps et d'ordre 5. La quantité notée henderson désigne la valeur du critère de Henderson (voir 2-4)

II.3. Les moyennes mobiles de Kendall et Stuart

De façon tout à fait naturelle, on peut souhaiter généraliser le premier procédé de construction des moyennes mobiles arithmétiques simples en cherchant des moyennes qui conservent les polynômes de degré d, Il faut tout d'abord remarquer que cette propriété de conservation de polynôme est **locale**: il suffit qu'une série soit localement, c'est-à-dire sur toute période de longueur l'ordre de la moyenne mobile, assimilable à un polynôme de degré d pour qu'elle soit conservée par la moyenne mobile.

Kendall et Stuart se sont intéressés à ce problème en le resolvant par une technique de "régressions mobiles", Sur tout ensemble de 2p+1 points consécutifs de la série, on ajuste un polynôme de degré d : la valeur ajustée au centre de cet ensemble de points représentera la valeur de la série lissée à cet instant. Ils montrent par ailleurs que cette méthode revient à appliquer à la série de départ une moyenne mobile ad hoc, dont les coefficients ne dépendent que du degré du polynôme choisi et du nombre p.

En fait, et c'est l'objet de l'annexe 2, on peut montrer que cette approche est complètement équivalente à résoudre un problème de minimisation, le critère étant celui de

Bongard, de minimisation de la variance résiduelle, et les contraintes celles inhérentes à la conservation d'un polynôme de degré d. Comme précédemment donc la matrice est la matrice identité et on a par ailleurs, pour une moyenne centrée d'ordre 2p+1 :

$$\alpha=\begin{bmatrix}1\\0\\0\\.\\.\\.\\.\\.\\.\\0\\0\end{bmatrix},w=\begin{bmatrix}1\\0\\0\\.\\.\\.\\.\\.\\.\\0\\0\end{bmatrix}\quad C=\begin{bmatrix}1 & 1 & \ldots\ldots & 1 & 1\\ -p & -p+1 & \ldots\ldots & p-1 & f\\ (-p)^2 & (-p+1)^2 & \ldots\ldots & (p-1)^2 & f^2\\ \ldots & \ldots & \ldots\ldots & \ldots & \ldots\\ (-p)^{d-1} & (-p+1)^{d-1} & \ldots\ldots & (p-1)^{d-1} & p^{d-1}\\ (-p)^d & (-p+1)^d & \ldots\ldots & (p-1)^d & p^d\end{bmatrix}$$

Là encore, on pourrait imposer en outre à la moyenne mobile d'éliminer par exemple certaines saisonnalités ; on retrouverait alors l'approche développée par Bongard en 1962 dont les résultats précédents sont des cas particuliers.

II.4. Les moyennes mobiles de Henderson

Les moyennes mobiles de type "Henderson" sont surtout utilisées pour **lisser** une série, L'estimation de la tendance doit donc être une courbe lisse, Une base de l'ensemble des séries étant constituée des séries définies par :

$$X_t(t_0)=\begin{cases}1 & \text{si } t=t_o\\ 0 & \text{si } t\neq t_o\end{cases}.$$

il suffit d'imposer que les transformées de ces séries par la moyenne mobile soient souples, Ces séries transformées sont, à une translation des temps près, égales à la série des coefficients de la moyenne mobile :

$$\ldots\ 0\ 0\quad \theta_{-p}\quad \theta_{-p+1}\ \ldots\ \theta_{p-1}\quad \theta_p\ \ 0\ \ 0\ \ldots$$

Il suffit donc d'imposer à la courbe des coefficients de la moyenne mobile d'être souple, Henderson a proposé d'utiliser comme critère de "souplesse" la quantité :

$$\sum_{i=-\infty}^{i=+\infty} (\nabla^3 \theta_i)^2$$

où ∇ représente l'opérateur différence première,($\nabla X_t = X_t - X_{t-1}$)

Plus cette quantité est faible, plus la série transformée par la moyenne mobile est jugée souple. Cette quantité est notée Henderson dans les tableaux proposés. De plus, les moyennes mobiles de Henderson sont supposées restituer correctement des polynômes de faible degré. Ainsi, dans le logiciel Census X11, des moyennes mobiles de Henderson symétriques sur 5, 7, 9, 13 ou 23 termes, et conservant les polynômes de degré 2, sont utilisées selon les cas (et en particulier selon la périodicité de la série).

Si on cherche par exemple à retrouver la moyenne mobile de Henderson symétrique sur 2p +1 termes, qui conserve donc les polynômes de degré inférieur ou égal à 3, les paramètres de notre problème de minimisation seront :

$$\underset{(3,2p+1)}{C} = \begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ -p & -p+1 & \dots & p-1 & p \\ (-p)^2 & (-p+1)^2 & \dots & (p-1)^2 & p^2 \end{bmatrix} \quad \underset{(3,1)}{\alpha} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \underset{(2p+1)}{w} = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

$$\underset{(2p+1,2p+1)}{\Omega} = \begin{bmatrix} 20 & -15 & 6 & \dots & 0 & 0 & 0 \\ -15 & 20 & -15 & \dots & \dots & \dots & \dots \\ 6 & -15 & 20 & \dots & 6 & -1 & 0 \\ -1 & 6 & -15 & \dots & -15 & 6 & -1 \\ 0 & -1 & 6 & \dots & 20 & -15 & 6 \\ \dots & \dots & \dots & \dots & -15 & 20 & -15 \\ 0 & 0 & 0 & \dots & 0 & -15 & 20 \end{bmatrix}$$

II.5. De Bongard à Henderson

On peut avoir l'idée (comme dans [5]) de mélanger en quelque sorte les deux critères et de construire des moyennes mobiles "conciliant" la réduction de bruit et le pouvoir de lissage de la série, C'est ce que nous avons fait en considérant une combinaison convexe des deux critères, et en résolvant donc le problème :

$$\begin{cases} \underset{\theta}{\text{Min}} & {}^t(\theta - w)\,[k\,\Omega + (1-k)I](\theta - w) \\ \text{sous} & C\,\theta = \alpha \end{cases}$$

où k est une constante comprise entre 0 et 1, I et Ω désignant les matrices associées respectivement aux critères de Bongard et Henderson. Ce programme se résoud de la même façon que précédemment et, à titre d'exemple, le *tableau* 2 présente, pour des valeurs de k augmentant de 0,1 en 0,1, les moyennes mobiles centrées d'ordre 9 conservant les polynômes de degré 2. Sur cet exemple, on constate que le pouvoir de "lissage" représenté par le critère de HENDERSON s'améliore très rapidement dès que le degré de "contamination" du critère de BONGARD atteint 10 à 20%.

La notation Sp signifie que la moyenne mobile a été obtenue avec une valeur de k égale à p/10. La valeur i= 0 correspond à l'instant t présent.

II.6. Les moyennes mobiles de Spencer

Une moyenne mobile comme l'une des deux premières M4x4x5 présentées ci-dessus a de bonnes propriétés en termes d'élimination de saisonnalités mais, en termes de restitution de tendance, ne restitue que les droites. Au début de ce siècle, le problème du lissage préoccupait beaucoup les actuaires qui cherchaient en particulier à lisser les courbes de mortalité et ce, de façon simple, les contraintes matérielles de temps de calcul étant, à cette époque, fondamentales. Ainsi, Spencer a utilisé ce principe de composition de moyennes mobiles à coefficients simples, ce qui assure une succession de calculs simples, et a cherché une moyenne mobile qui, appliquée après une 4x4x5, corrigerait les défauts, en termes de lissage, de cette dernière. Remarquons en outre que si on utilise une moyenne symétrique qui permettrait à la composée de tout cela de conserver les paraboles, cette résultante conserverait naturellement aussi les cubiques.

Tableau 2 :

Moyennes mobiles centrées de Bongard et Henderson

I	S0	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
-4	-0,091	-0,063	-0,055	-0,050	-0,047	-0,045	-0,044	-0,043	-0,042	-0,041	-0,041
-3	0,061	0,015	0,005	0,000	-0,004	-0,005	-0,006	-0,008	-0,009	-0,009	-0,010
-2	0,169	0,154	0,142	0,135	0,130	0,127	0,124	0,122	0,121	0,119	0,118
-1	0,234	0,252	0,257	0,260	0,262	0,264	0,264	0,265	0,266	0,266	0,267
0	0,255	0,284	0,300	0,310	0,316	0,320	0,323	0,326	0,328	0,330	0,331
1	0,234	0,252	0,257	0,260	0,262	0,264	0,264	0,265	0,266	0,266	0,267
2	0,169	0,154	0,142	0,135	0,130	0,127	0,124	0,122	0,121	0,119	0,118
3	0,061	0,015	0,005	0,000	-0,004	-0,005	-0,006	-0,008	-0,009	-0,009	-0,010
4	-0,091	-0,063	-0,055	-0,050	-0,047	-0,045	-0,044	-0,043	-0,042	-0,041	-0,041
HENDERSON	0,402	0,126	0,091	0,078	0,073	0,070	0,069	0,068	0,068	0,068	0,067
BONGARD	0,255	0,263	0,269	0,273	0,276	0,278	0,279	0,281	0,282	0,283	0,283
MIXTE	0,255	0,249	0,233	0,214	0,194	0,174	0,153	0,132	0,110	0,089	0,067
PÉRIODE	10,13	11,27	11,32	11,26	11,20	11,15	11,10	11,06	11,03	11,00	10,98

On peut montrer (voir [4] pour un calcul analogue) que si on cherche une moyenne d'ordre 3, symétrique, de coefficients (a,b,a) et qui, appliquée à notre M4x4x5, conduit à une moyenne mobile générale conservant les cubiques, il faut que a et b vérifient :

$$\begin{cases} 2a+b=1 \\ 22a+9b=0 \end{cases}$$

Ce qui conduit à la solution unique $\frac{1}{4}[-9, 22, -9]$ correspondant à la moyenne notée M4x4x5 (1) du *tableau 1*.

Les propriétés de lissage de cette moyenne n'ont pas paru suffisantes à SPENCER qui a cherché une moyenne mobile d'ordre 5 conduisant aux mêmes propriétés. Si cette moyenne, symétrique, admet pour coefficients (a,b,c,b,a), on peut montrer (voir [4] pour un calcul analogue) que a, b et c doivent vérifier le système :

$$\begin{cases} 2a+2b+c=1 \\ 34a+22b+9c=0 \end{cases}$$

Ce qui conduit à un ensemble moyennes mobiles de coefficients :

$$\frac{1}{12}[2c-11, 17-8c, 12c, 17-8c, 2c-11]$$

parmi lesquelles Spencer a privilégié la moyenne $\frac{1}{12}[-9, 9, 12, 9, -9]$ obtenue pour c=1. La moyenne de Spencer sur 15 points ainsi obtenue figure dans le tableau 1, A titre de comparaison figure, dans ce même tableau, sous le nom M4x4x5 (2), la moyenne mobile sur 15 points ayant les mêmes propriétés que la moyenne de Spencer, et minimisant le critère de Bongard : elle est obtenue pour c voisin de -13. De même, la moyenne mobile sur 15 points ayant les mêmes propriétés que la moyenne de Spencer, et minimisant le critère de Henderson est obtenue pour c voisin de 0,7, ce qui montre la proximité des approches de Spencer et Henderson.

III - Génération de moyennes mobiles non centrées

Dans la pratique, les moyennes mobiles symétriques sont préférées aux moyennes mobiles asymétriques parce qu'elles présentent, dans leur application, moins d'effets pervers, notamment en ce qui concerne les effets d'amplitude et de phase. Malheureusement, si on lisse une série avec une moyenne mobile centrée d'ordre 2p+1, on ne disposera pas d'estimation de la série lissée pour les p premiers et les p derniers instants, ce qui est pour le moins génant. Dans la pratique, on serait donc conduit à utiliser des moyennes non centrées pour effectuer ces estimations. Notons, dès à présent, que les problèmes d'amplitude et de phase évoqués ci-dessus sont dans ce cas moins cruciaux dans la mesure où il s'agit, non de lisser la série dans son ensemble, mais d'estimer quelques points. En fait, à notre connaissance, il n'y a que dans le programme de désaisonnalisation Census X11 et ses dérivés que des moyennes mobiles non centrées sont utilisées pour résoudre ce problème ; et encore faut-il reconnaitre que leur mode de génération n'a été que très récemment retrouvé (voir [2]), même s'il existe de bonnes études sur les propriétés de ces moyennes (voir LAROQUE [7] et WALLIS [11]).

III.1.Deux générations "naturelles" de moyennes mobiles non centrées

Une première idée très naturelle est d'utiliser, pour ces estimations, des moyennes mobiles centrées (et même symétriques) d'ordre inférieur, Par exemple, si notre série a été lissée par une moyenne mobile de Henderson sur 9 termes et si T est le dernier instant de la série, on pourrait estimer la valeur de la série lissée en T-4 par une Henderson sur 7 termes, la valeur en T-3 par une Henderson sur 5 termes . Malheureusement, cette idée se heurte à plusieurs problèmes. Tout d'abord, le cas du dernier instant d'observation n'est pas résolu et nous n'aurons toujours pas d'estimation pour cette date. Ensuite, si notre exemple permet d'utiliser des moyennes mobiles d'ordre

Tableau 3 :

Moyennes mobiles d'ordre 9 de Henderson conservant les polynômes de degré 2 et 3

I	H4_4	D5_3	D6_2	D7_1	D8_0	T5_3	T6_2	T7_1	T8_0
-8					0,143				-0,132
-7				0,041	0,118			0,024	0,069
-6			-0,024	0,015	-0,121		0,055	0,012	0,278
-5		-0,051	-0,041	-0,071	-0,346	0,015	-0,026	-0,046	0,104
-4	-0,041	-0,049	-0,015	-0,114	-0,302	-0,037	-0,129	-0,086	-0,302
-3	-0,010	0,048	0,065	-0,025	0,046	-0,048	-0,064	-0,025	-0,403
-2	0,118	0,192	0,173	0,179	0,466	0,083	0,173	0,151	0,067
-1	0,267	0,292	0,260	0,369	0,622	0,292	0,389	0,344	0,672
0	0,331	0,289	0,279	0,393	0,373	0,398	0,393	0,396	0,648
1	0,267	0,195	0,211	0,214		0,292	0,197	0,231	
2	0,118	0,077	0,092			0,065	0,013		
3	-0,010	0,007				-0,059			
4	-0,041								
HENDERSON	0,067	0,042	**0,029**	0,209	0,889	0,205	0,257	0,220	3,684
BONGARD	0,283	0,257	**0,235**	0,389	1,006	0,347	0,399	0,362	1,240
PÉRIODE	10,975	11,860	12,881	9,166	7,570	9,169	8,695	9,245	5,810

inférieur ayant les mêmes propriétés en termes de conservation de tendances, il n'en serait pas de même pour d'autres types de moyennes mobiles ; ainsi, des moyennes mobiles arithmétiques simples éliminent des saisonnalités différentes selon leur ordre. Et enfin, il est assez intuitif de comprendre que plus l'ordre de la moyenne mobile est grand, plus l'effet de rabot de la moyenne sera important; changer cet ordre conduit donc à moins lisser la fin de la série que la partie centrale et il n'est pas évident *a priori* que ce soit souhaitable.

Une seconde idée découle du fait que la résolution du problème de minimisation d'une forme quadratique sous contraintes qui permet de déterminer les coefficients d'une moyenne mobile ne nécessite pas l'hypothèse que la moyenne mobile soit centrée, KENDALL (voir [6]) puis BONGARD (voir [1]) avaient d'ailleurs déjà utilisé cette idée et imaginé de telles moyennes mobiles non-centrées, À titre d'exemple, le tableau 3 présente les coefficients des moyennes mobiles de Henderson non centrées d'ordre 9 qui permettraient d'estimer les 4 derniers points d'une série lissée avec une Henderson centrée sur 9 termes ; on a choisi d'y faire figurer des moyennes conservant les polynômes de degré 2 (notés D) ou 3 (notés T). La moyenne mobile centrée de Henderson conservant les polynômes de degré 2 étant symétrique, elle conserve aussi les cubiques, Il est curieux de constater que, sur cet exemple, il vaut mieux utiliser pour lisser une courbe, du point de vue des critères de Henderson et Bongard, une moyenne de type D6_2, donc non centrée, plutôt qu'une moyenne centrée.

La notation p_f signifie que la moyenne mobile est d'ordre p+f+1 avec p termes dans le passé et f termes dans le futur. La valeur i=0 correspond à l'instant t présent. Les moyennes D (respectivement T) conservent les polynômes de degré 2 (respectivement 3).

III.2. Les moyennes mobiles non centrées de Musgrave

III.2.1. La démarche initiale de Musgrave

Musgrave a cherché à résoudre ce problème d'estimation des données les plus récentes dans le cadre de la désaisonnalisation et, plus précisément, pour la mise au point du logiciel X11 qui, entre autres, utilise pour estimer la tendance de la série, des moyennes mobiles de Henderson et des moyennes composées de moyennes mobiles arithmétiques simples pour l'estimation des saisonnalités. L'idée de base de Musgrave est que les estimations des derniers points, faites grâce à la moyenne mobile non centrée, devraient être le moins possible révisées lorsque sera disponible l'information à la date T+1.

Pour cela, il pose les hypothèses suivantes :

la série peut se modéliser linéairement sous la forme : $X = a + bi + \varepsilon_i$ où a et b sont des constantes, et les ε_i sont des variables aléatoires non corrélées, de moyenne nulle et de variance σ^2.

On dispose d'une série de poids $w_1, w_2, \ldots\ldots\ldots, w_N$ de somme égale à 1 (c'est en fait par exemple notre moyenne mobile centrée de Henderson) et on cherche une série de poids $u_1, u_2, \ldots\ldots, u_m$, avec m < N, de somme égale à 1.

Cette nouvelle moyenne mobile doit en outre minimiser les révisions des estimations, c'est-à-dire, par exemple, minimiser le critère : $E\left(\sum_{i=1}^{i=m} u_i \, X_i - \sum_{i=1}^{i=N} w_i \, X_i\right)^2$

Sous ces hypothèses, on montre (voir [2] pour le calcul détaillé) que les poids peuvent être calculés en fonction du rapport $D = \frac{b^2}{\sigma^2}$. Malheureusement, la valeur de D est inconnue mais Musgrave fait remarquer que le choix de l'ordre des moyennes mobiles de Henderson dans X11 se fait à partir de la valeur du rapport $M = \frac{\bar{I}}{\bar{C}}$ ou $\bar{I}$ désigne la moyenne des variations absolues au mois le mois dans la partie irrégulière de la série C et désigne la moyenne des variations absolues au mois le mois dans la tendance de la série, En supposant la normalité des ε_i, on montre alors que :

Tableau 4 :

Moyennes mobiles non centrées d'ordre 9 de Musgrave correspondant à une moyennes de Henderson

I	H4_4	M4_3	M4_2	M4_1	M4_0
-4	- 0,041	- 0,031	- 0,023	- 0,049	- 0,156
-3	- 0,010	- 0,004	- 0,000	- 0,011	- 0,034
-2	0,118	0,120	0,120	0,126	0,185
-1	0,267	0,264	0,259	0,282	0,424
0	0,331	0,324	0,315	0,354	0,580
1	0,267	0,255	0,242	0,298	0,000
2	0,118	0,103	0,086	0,000	0,000
3	- 0,010	- 0,030	0,000	0,000	-0,000
4	- 0,041	0,000	0,000	0,000	-0,000
HENDERSON	0,067	0,00	0,060	0,432	2,750
BONGARD	0,283	0,266	0,248	0,312	0,576
PÉRIODE	10,975	10,786	11,253	9,064	6,451

$$D = \frac{4}{\pi}\left(\frac{1}{M}\right)^2$$

ce qui permet, pour des valeurs particulières de M, de déduire les moyennes mobiles non centrées solutions. À titre d'exemple, le *tableau 4* donne ces moyennes mobiles pour une valeur de M égale à 1, valeur pour laquelle X11 choisit une moyenne mobile de HENDERSON centrée sur 9 termes pour la partie "centrale" de la série.

La notation Mp_f signifie que la moyenne mobile est d'ordre p+f+1 avec p termes dans le passé et f termes dans le futur, La valeur i=0 correspond à l'instant t présent.

Cette approche est aussi valable, selon [9] et [10], pour les moyennes mobiles composées utilisées dans X11 pour traiter les saisonnalités (moyennes 3x3, 3x5, 3x9 ...) ; il suffit dans cette démonstration, de remplacer le rapport $\frac{\bar{I}}{\bar{C}}$ par le rapport dit "ratio de saisonnalité mobile", $\frac{\bar{I}}{\bar{S}}$ où $\bar{I}$ désigne toujours la moyenne des variations absolues au mois le mois dans la partie irrégulière de la série $\bar{S}$ et désigne la moyenne des variations absolues au mois le mois dans la saisonnalité de la série.

III.2.2. Généralisation de cette approche

Nous allons montrer à présent que cette démarche peut se généraliser et se mettre sous la forme envisagée dans cette note, Pour cela, on peut envisager une modélisation polynômiale de degré d de la série brute :

$$X_I = a_0\, i + a_1 i + a_2 i^2 + \dots\dots\dots + a_d\ i^d + \varepsilon_i \text{ pour } i \in \{-p, -p+1, \dots\dots, f-1, f\}$$

Ce qui peut s'écrire matriciellement :

$$X = A\beta + \Sigma$$

avec

$$\underset{(p+f+d+1)}{A} = \begin{bmatrix} 1 & -p & (-p)^2 & \dots & (-p)^d \\ \dots & \dots & \dots & \dots & \dots \\ 1 & f & f^2 & \dots & f^d \end{bmatrix}$$

$$\underset{(d+1,1)}{\beta} = \begin{bmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_d \end{bmatrix} \qquad \underset{(p+f+1,1)}{\Sigma} = \begin{bmatrix} \varepsilon_{-p} \\ \varepsilon_{-p+1} \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_f \end{bmatrix}$$

Soient en outre w le vecteur des coefficients de la moyenne mobile d'ordre p+f+1 ayant servi à lisser la partie centrale de la série et u le vecteur des coefficients de la moyenne mobile non centrée recherchée. Comme le montre le tableau précédent, on peut toujours supposer que cette moyenne mobile est d'ordre p+f+1 si on ajoute une condition de nullité des derniers coefficients. Le critère de Musgrave s'écrit alors :

$$E\left(\sum_{i=-p}^{i=f} u_i\ X_i - \sum_{i=-p}^{i=f} w_i\ X_i\right)^2 = E\left[{}^t(u-w)\ X^t\ X\,(u-w)\right]$$

$$E\left[{}^t(u-w)\ \Omega\,(u-w)\right]$$

avec

$$\Omega = \left(X^t X\right) = E\ \left((A\beta+\Sigma)^t\ (A\beta+\Sigma)\right) = A\beta\,{}^t\beta\,{}^t A\ + \sigma^2\, 1$$

puisque $E\ (\Sigma) = 0$ *et* $\mathrm{Var}\,(\Sigma) = \sigma^2\, \mathrm{I}$

La matrice Ω est une matrice définie positive d'ordre p+f+1, comme il est immédiat de le vérifier.

Il suffit donc à présent d'inclure dans la matrice des contraintes C la nullité de certains coefficients u_i ainsi que le fait que la somme des coefficients est égale à 1. La matrice C, de plein rang, s'écrit alors :

$$C = \begin{bmatrix} 1 & 1 & \ldots & 1 & 1 & 1 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & 0 & 0 \\ 0 & 0 & \ldots & 0 & 1 & 0 \\ 0 & 0 & \ldots & 0 & 0 & 1 \end{bmatrix}$$

III.2.3. Quelques remarques

On peut alors imaginer la démarche suivante : une fois fixé le degré du polynôme auquel est supposée s'adapter la série brute, on estime par les moindres carrés ordinaires les coefficients du polynôme, soit la matrice β et la variance des perturbations, soit la quantité σ^2. La matrice Ω est alors déterminée et on peut calculer automatiquement les coefficients u_i de la moyenne mobile non symétrique.

Reprenons le cas traité par Musgrave, où w est une moyenne mobile de Henderson. Il est facile de constater que si on impose à u de conserver, non plus les constantes, mais les droites, alors u ne dépend plus de la série initiale. En effet, on a alors, à cause des contraintes sur les coefficients de u et w imposées par leurs propriétés de conservation de tendances :

$$\left(\sum_{i=-p}^{i=+f} u_i\ X_i - \sum_{i=-p}^{i=+f} w_i\ X_i\right)^2 = \left[\sum_{i=-p}^{i=+f} (u_i - w_i)\ (a + ib + \varepsilon_i)\right]^2$$

$$= \left[a \sum_{i=-p}^{i=+f} u_i - a \sum_{i=-p}^{i=+f} w_i + b \sum_{i=-p}^{i=+f} iu_i - b \sum_{i=-p}^{i=+f} iw_i \sum_{i=-p}^{i=+f} \varepsilon_i\ (u_i - w_i)\right]^2$$

$$= \left[\sum_{i=-p}^{i=+f} \varepsilon_i\ (u_i - w_i)\right]^2$$

$$= {}^t(u - w)\ \Sigma\ {}^t\Sigma\ (u - w)$$

puisque $\sum_{-p}^{f} u_i = \sum_{-p}^{f} w_i = 1$ et $\sum_{-p}^{f} iu_i = \sum_{-p}^{f} iw_i = 0$

Et, si on prend l'espérance de cette quantité, on trouve: $\sigma 2\ t\ (u\text{-}w)\ (u\text{-}w)$, La minimisation de ce critère ne dépend alors plus de σ^{2t} Par ailleurs, on retrouve ici une formulation du problème voisine de l'approche de Kendall et Stuart : on s'intéresse à un problème de régression locale (ici on ajuste à une droite) en imposant en outre à la moyenne mobile d'une part d'avoir certains coefficients nuls et, d'autre part, d'être "proche" de la moyenne mobile w initiale.

Il est difficile de comprendre pourquoi, si ce n'est pour des raisons de calcul, Musgrave s'est cantonné à des hypothèses si "restrictives" qui l'amènent à construire des moyennes mobiles dépendant de façon un peu curieuse de la série : le critère M explicité plus haut est en effet calculé sur l'ensemble de la série et, si de ce fait il peut être utilisé pour choisir l'ordre de la moyenne mobile lissant la partie centrale de la série, il est par nature beaucoup moins robuste pour les données les plus récentes.

Dans le cas particulier donc où la moyenne "centrale" w est une moyenne mobile de Henderson, il nous parait donc plus efficace de chercher une moyenne mobile non centrée dont certains coefficients sont eventuellement nuls, conservant les polynômes de degré 2 (comme la moyenne w, ce qui équivaut à supposer que la série est localement assimilable à ce type de fonction) et "proche" de w au sens où elle minimise le critère ${}^t(u-w)\,(u-w)$.

III.3. En résumé

À partir d'une moyenne mobile MM donnée, d'ordre p+f+1, le problème des extrémités de série se ramène donc à celui de la génération de moyennes mobiles non centrées permettant d'estimer les points manquants. Deux cas se présentent en pratique :

On ne connaît rien sur la moyenne mobile MM, c'est-à-dire rien d'autre que ses coefficients et en tout cas pas ses propriétés en termes de conservation de tendance, de lissage, d'élimination de saisonnalités, de réduction de variance. Dans ce cas, il parait assez naturel de chercher des moyennes U_k, (k=1..... f si on s'intéresse aux instants les plus récents), les plus proches possibles de MM au sens où elles minimisent le critère ${}^t(u-w)\,(u-w)$, w désignant ici le vecteur des coefficients de MM, et ayant k coefficients nuls, Evidemment, il est toujours possible d'imposer à ces moyennes mobiles d'autres propriétés, par exemple celles de la moyenne "initiale" MM que l'on pourrait retrouver à partir de ses coefficients (vérifier que la somme des coefficients est égale à 1....).

Le second cas est plus simple dans la mesure où on suppose que l'on connaît, sinon le type de la moyenne mobile, au moins le pourquoi de son utilisation. On peut alors en connaissance de cause choisir les contraintes et éventuellement le critère.

Le *tableau 5* montre le résultat, à partir d'une moyenne mobile de Henderson centrée sur 9 termes, de ces possibilités, en se restreignant aux moyennes mobiles permettant d'estimer le premier point manquant :

H4_4 désigne cette moyenne mobile conservant les polynômes de degré 2.

Tableau 5 : Moyennes mobiles non centrées d'ordre 9 associées à une moyenne de Henderson.

I	H4_4	M1_4_3	M2_4_3	M3_4_3	D5_3	H4_3
-5					- 0,051	
-4	- 0,041	- 0,046	- 0,031	- 0,056	- 0,049	- 0,063
-3	- 0,010	- 0,015	- 0,004	- 0,008	0,048	- 0,021
-2	0,118	0,113	0,120	0,131	0,192	0,147
-1	0,267	0,261	0,264	0,282	0,292	0,315
0	0,331	0,326	0,324	0,342	0,289	0,350
1	0,267	0,261	0,255	0,266	0,195	0,231
2	0,118	0,113	0,102	0,099	0,077	0,063
3	- 0,010	- 0,015	- 0,030	- 0,056	0,007	- 0,021
4	- 0,041	0,000	0,000	0,000		0,000
HENDERSON	0,067	0,090	0,101	0,173	0,042	0,112
BONGARD	0,283	0,271	0,266	0,300	0,257	0,305
PÉRIODE	10,975	10,853	10,785	10,135	11,860	10,154

M1_4_3 désigne la moyenne "la plus proche" de H4_4 au sens du critère précédent, n'utilisant que 3 points dans le "futur", et conservant les constantes.

M2_4_3 est calculée de la même façon en rajoutant la contrainte de conservation des droites.

M3_4_3 est calculée de la même façon en rajoutant la contrainte de conservation des polynômes de degré 2.

D5_3 est la moyenne mobile de Henderson non centrée sur 9 termes, conservant les polynômes de degré 2.

Enfin H4_3 est la moyenne mobile de Henderson non centrée avec un coefficient nul.

IV- Un exemple simple d'application

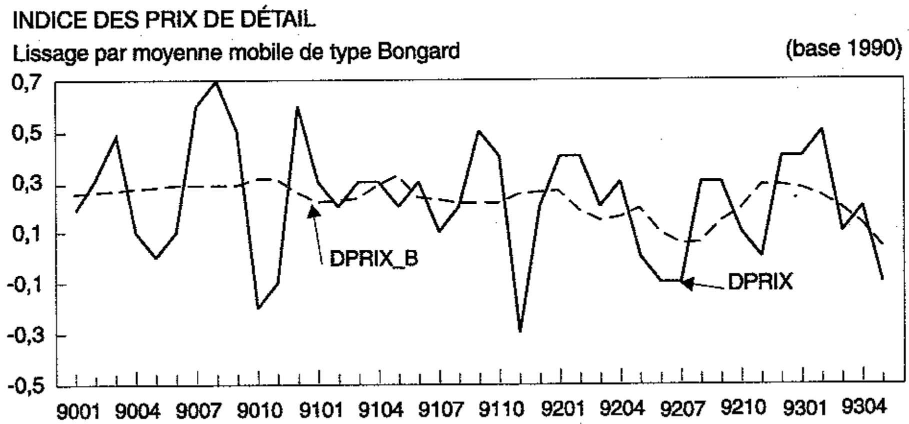
L'indice des prix de détail calculé mensuellement par l'Insee a récemment changé de méthodologie, Ainsi, les phénomènes de soldes, peu ou pas pris en compte par le passé, sont inclus aujourd'hui dans cet indice; les évolutions des prix des produits frais sont désormais prises en compte chaque mois, alors qu'elles étaient lissées auparavant. Ces phénomènes introduisent d'une part une saisonnalité dans l'indice des prix et, d'autre part, une discontinuité dans la série.

Une rétropolation a été possible jusqu'au début de 1990 mais on ne dispose à l'heure actuelle que de 3 ans et demi d'observations. Or, la plupart des logiciels de désaisonnalisation (dont X11 et X11-ARIMA) nécessitent au moins 5 années pleines d'observations, Il faut donc effectuer une désaisonnalisation "à la main".

Ceci implique donc d'estimer une tendance "lisse", ce qui permet ensuite d'estimer des coefficients saisonniers selon une méthode assez naturelle :
On part d'un schéma additif :$X_t = T_t + S_t + \varepsilon_t$. On estime T_t à l'aide d'une moyenne mobile appropriée, Par différence, on obtient une première estimation des coefficients saisonniers. Comme pour chaque mois on obtient ainsi plusieurs estimation d'un même coefficient saisonnier, ces valeurs sont synthétisées en une seule, la médiane par exemple, ce qui fournit une seconde estimation des coefficients saisonniers. Enfin, cette estimation est corrigée afin que les coefficients soient de moyenne nulle sur une année,

Les trois graphiques suivants montrent l'importance du choix du critère pour le calcul de la moyenne mobile et donc l'estimation de la tendance.

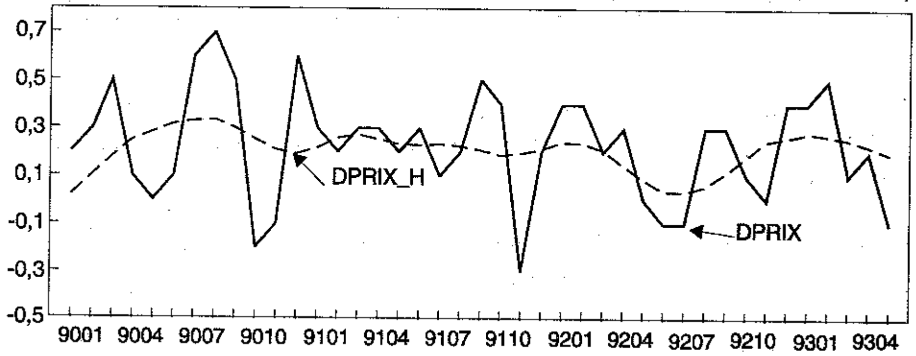
Le premier graphique présente le résultat du lissage par une moyenne mobile sur 13 termes, conservant les polynômes de degré inférieur ou égal à 2 et minimisant le critère de Bongard, résultat noté DPRIX_B.



La méthode décrite ci-dessus permet de ne pas perdre d'information aux extrémités de la série. Le lissage est correct à ces extrémités mais semble un peu "chahuté" au milieu.

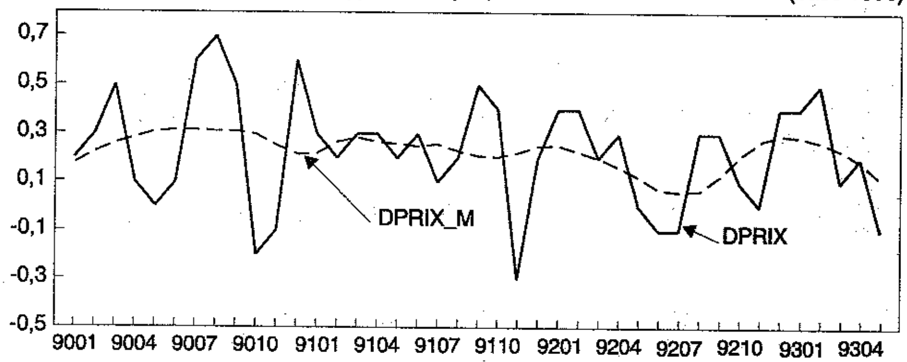
Le second graphique présente le résultat du lissage par une moyenne mobile sur 13 termes, conservant les polynômes de degré inférieur ou égal à 2 et minimisant le critère de Henderson, résultat noté DPRIX_H. Le lissage est correct au milieu mais un peu "brusque" aux extrémités.

INDICE DES PRIX DE DÉTAIL
Lissage par moyenne mobile de type Henderson (base 1990)



Le graphique 3 présente enfin le résultat du lissage par une moyenne mobile sur 13 termes, conservant les polynômes de degré inférieur ou égal à 2 et minimisant un critère mixte (mélange de Henderson et de Bongard) dans des proportions 0,5 ; 0,5 résultat noté DPRIX_M. Le lissage parait plus régulier et c'est celui qui a été retenu pour la désaisonnalisation.

INDICE DES PRIX DE DÉTAIL
Lissage par moyenne mobile de type mixte (0.5) (base 1990)



BIBLIOGRAPHIE

[1] BONGARD, J. (1962) : "Quelques remarques sur les moyennes mobiles", dans "L'élimination des variations saisonnières à l'aide de calculatrices électroniques", OCDE.

[2] DOHERTY, M. (1990) : "Note on Musgrave's Method", *note interne*, Department of Statistics, Wellington, New Zealand.

[3] GOURIEROUX, C. LEGALLO, F. (1981) : "Construction de moyennes mobiles par minimisation sous contraintes d'une forme quadratique des coefficients", *Annales de l'Insee*, n°42.

[4] GOURIEROUX, C. MONFORT, A. (1990) : "Séries temporelles et modèles dynamiques", Economica,

[5] GRAY, A., THOMSON, P. (1992) : "Design of Filters for Seasonal Adjustment", pre-print, Department of Statistics and Victoria University, Wellington,New Zealand.

[6] KENDALL, G., STUART, A. (1968) : "The Advanced Theory of Statistics", Griffin, vol,3, chapitre 4.

[7] LAROQUE, G. (1977) : "Analyse d'une méthode de désaisonnalisation : le programme X11 du US bureau of Census, version trimestrielle, *Annales de l'Insee*, n°28, pp 105-127.

[8] MACAULAY, F. (1931): "Smoothing of Time Series", National Bureau of Economic Research.

[9] MUSGRAVE, J. (1964) : "A Set of End Weights to End all End Weights", note interne, US Bureau of Census, Washington.

[10] MUSGRAVE, J. (1964) : "Alternative Sets of Weights for Proposed X-11 Seasonal Factor Curve Moving Averages", *note interne*, US Bureau of Census, Washington.

[11] WALLIS, K. (1982) : "Seasonal adjustment and revision of current data : linear filters for the X11 Method,". *Journal of the Royal Statistical Society* A, 145, pp 74-85 .

ANNEXE 1

Contraintes sur les coefficients induites par des propriétés d'annulation de saisonnalités variant avec le temps

On cherche ici quelles conditions doivent vérifier les coefficients d'une moyenne mobile MM d'ordre p+f+1 pour que celle ci élimine les fonctions périodiques de période b variant linéairement avec le temps. Si (u_t) est une série périodique de période b, on supposera en outre qu'elle est de moyenne nulle sur toute période de b instants consécutifs. Le problème peut dans un premier temps s'écrire :

soit une série (X_t) telle que :

$X_t = t\, u_t$

(u_t) fonction périodique de période b : $\forall t \;\; u_{t+b} = u_t$

$$\sum_{k=1}^{k=b} u_{t+k} = 0 \;\; \forall t$$

on cherche une moyenne mobile M, de coefficients $\{\phi\ i\}$ telle que :

$$MX_t = \sum_{k=-p}^{k=+f} \theta_k X_{t+k} = 0$$

La difficulté dans le calcul provient du fait que p+f+1 n'est pas forcément un multiple de la période b. On notera :

$p+f+1 = nb + \Upsilon$, où n et Υ sont les quotient et reste de la division euclidienne de p+f+1 par b,

On a :

$$MX_t = \sum_{k=-p}^{k=f} \theta_k\ (t+k)\ u_{t+k} = \sum_{k=-p}^{k=f-nb} \theta_k\ (t+k)\ u_{t+k} + \sum_{k=f-nb+1}^{k=+f} \theta_k\ (t+k)\ u_{t+k}$$

$$MX_t = \sum_{k=np-p}^{k'=f} \theta_{k'} - nb\,(t+k'-nb)\; u_{t+k'} = \sum_{k=f=nb+1}^{k=f} \theta_k\;(t+k)\; u_{t+k}$$

Ceci en posant k' = k + nb et en utilisant la périodicité de u_t : $u_{t+k'} = u_{t+k}$

La première somme correspond aux coefficients θ_i qui "débordent" dans le passé un multiple n de la période b (si $\lambda = 0$, cette première somme ne figure pas dans l'expression de MX_t), La deuxième somme correspond aux coefficients θ_i situés dans l'intervalle de longueur nb, à partir de θ_f, en allant vers le passé :

$$\sum_{k=f-nb+1}^{k=f} \theta_k\;(t+k)\; u_{t+k} = \sum_{j-1}^{j=n}\;\sum_{k=f-jb+1}^{k=f(j-1)b} \theta_k\;(t+k)\; u_{t+k}$$

Si on pose: k' = k + (j-1)b, on obtient :

$$\sum_{j=1}^{j=n}\;\sum_{k'=f-b+1}^{k'=f} \theta_{k'-(j-1)b}\;(t+k'-(j-1)\,b\,)\; u_{t+k}$$

puis, avec j'=j-1 et k'=k :

$$\sum_{k=f-nb+1}^{k=f} \theta_k\;(t+k)\; u_{t+k} = \sum_{j=0}^{j=n-1}\;\sum_{k=f-b+1}^{k=f} \theta_{kjb}\;(t+k-jb)\; u_{t+k}$$

$$= \sum_{k=f-b+1}^{k=f}\;\sum_{j=0}^{j=n} \theta_{kjb}\;(t+k-jb)\; u_{t+k}$$

et comme :

$$u_{(t+f)} = -\sum_{k=f-b+1}^{k=f-1} u_{t+k}$$

on en déduit :

$$\sum_{k=f-nb+1}^{k=f} \theta_k \; (t+k)\, u_{t+k} = \sum_{kf-b+1}^{k=f} \; \sum_{j=0}^{j=n-1} \theta_{k-jb} \; (t+k-jb)\; u_{t+k}$$

$$= \sum_{k=f-b+1}^{k=f-1} \; \sum_{j=0}^{j=n-1} \theta_{k-jb} \; (t+k-jb)\; u_{t+k} - \sum_{kf-b+1}^{k=f-1} \; \sum_{j=0}^{j=n-1} \theta_{f-jb} \; (t+f-jb)\; u_{t+k}$$

$$= \sum_{k=f-b+1}^{k=f} \left\{ \sum_{j=0}^{j=n-1} \theta_{k-jb} \; (t+k-jb)\; \theta_{f-jb} \; (t+f-jb) \right\} u_{t+k} \qquad \text{(a)}$$

De même, pour la première somme :

$$\sum_{k=-p+nb}^{k=f} \theta_{k-nb} \; (t+k-nb)\; u_{t+k} = \sum_{k-p+nb}^{k=f-1} \theta_{k-nb} \; (t+k-nb)\; u_{t+k}$$

$$- \sum_{k=f-b+1}^{k=f-1} \theta_{f-nb}\, (t+f-nb)\; u_{t+k}$$

$$= \sum_{k=f+1-\gamma}^{k=f-1} \theta_{k-nb} \; (t+k-nb)\; u_{t+k} \qquad \text{(b)}$$

$$- \sum_{k=f-b+1}^{k=f-1} \theta_{j-nb} \; (t+k-nb)\; u_{t+k}$$

Si $\gamma = 1$, la première partie de la somme ne figure pas, Il faut ensuite additionner (a) et (b), ce qui revient dans (a) à faire varier l'indice j entre 0 et n, en faisant attention de ne pas dépasser l'ordre de la moyenne mobile (ce qui revient à annuler les termes associés aux coefficients θ_{k-nb} tels que k-nb < -p).

En conclusion :

1. Si $\lambda = 0$, on a simplement :

$$MX_t = \sum_{kf-b+1}^{k=f-1} \left\{ \sum_{j=0}^{j=n-1} \theta_{k-jb} \ (t+k-jb) \ \theta_{f-jb} \ (t+f-jb) \right\} u_{t+k}$$

et on obtient donc, pour chaque valeur de k, k = f - b +1, ... ,f - 1, deux équations de contraintes :

$$\sum_{j=o}^{j=n-1} \left(\theta_{k-jb} - \theta_{f-jb} \right) = 0 \text{ (coefficient de } t\text{)} \qquad (1)$$

$$\sum_{j=0}^{j=n-1} [(k-jb) \ \theta_{k-jb} - (f-jb) \ \theta_{f-jb}] = 0 \qquad \text{(constante)} \qquad (2)$$

2. Si , $\gamma \neq 1$ on a, pour chaque valeur de k, k=f-b+1, ..., f-1 :

$$\sum_{j=0}^{j=n} \theta_{k-jb} \, \varphi(k-jb) - \sum_{j=0}^{j=n} \theta_{f-jb} = 0 \quad \text{(coefficient de } t\text{)} \qquad (1)$$

$$\sum_{j=0}^{j=n} \theta_{k-jb}(k-jb) \, \varphi \, (k-jb) - \sum_{j=0}^{j=n} \theta_{f-jb} \, (f-jb) = 0 \quad \text{(constante) (2)}$$

où

$$\varphi \text{ (x)} \begin{cases} 0 \text{ si } x < -p \\ 1 \text{ si } x \leq -p \end{cases}$$

La généralisation à des polynômes de degré d se fait alors sans problème; on aboutit , dans le cas simple $\gamma = 0$, pour i=1,2, ... d à b-1 équations :

$$\sum_{j=0}^{j=n-1} [(k-jb)^i \theta_{k-jb} \, (f-jb)^i \theta_{f-jb}] = 0$$

On peut remarquer que les équations (1) assurent l'annulation des fonctions périodiques de période b. L'annulation du monôme d'ordre i dans le polynôme est assuré par l'équation "d'ordre i".

Régressions mobiles et minimisation sous contraintes du pouvoir de réduction

Il s'agit, dans cette annexe, de montrer que l'approche de Kendall et Stuart consistant à construire des moyennes mobiles à partir de régressions polynômiales locales est un cas particulier de l'approche de Bongard consistant à construire des moyennes mobiles par minimisation sous contraintes du pouvoir de réduction.

1 - Optique régressions polynômiales mobiles

L'idée de base repose sur la définition suivante : Une série (X_t) est assimilable à un polynôme de degré d, sur tout intervalle de temps de longueur p+f+1,si , ajustant sur chacun de ces intervalles (t-p, t-p+1, ... t ...... t+f) un polynôme de degré d à la série par la méthode des moindres carrés, la valeur ajustée est égale à la valeur vraie à l'instant t : $\hat{X}_t = X_t$.

Le polynôme de degré d s'écrit:

$$a_0 + a_1\, i + a_2\, i^2 + \dots\dots\dots\dots + a_d\, i^d \quad \text{pour } i \in \{-p, -p+1, \dots, f-1, f\}$$

Les coefficients sont estimés par la méthode des moindres carrés ordinaires, c'est-à-dire en minimisant la quantité :

$$\sum_{i=-p}^{i=+f} [\, X_{t+i} - a_0 - a_1\, i^2 - \dots\dots - a_d\, i^d \,]$$

La condition $\hat{X}_t = X_t$ se réduit alors à $\hat{X}_t = \hat{a}_0$ et, comme les estimateurs des moindres carrés sont des combinaisons linéaires des observations, $\hat{a}_0$ s'écrit sous la forme:

$$\hat{a}_0 = \sum_{i=-p}^{i=+f} \theta_i\, X_{t+i}$$

où les réels θ_i sont indépendants de t, Cette quantité peut être considérée comme la valeur au point t de la série transformée par la moyenne mobile d'ordre p+f+1 et de coefficients θ_i. Posons, pour n entier positif ou nul :

$$A_n = \sum_{i=-p}^{i=+f} i^n$$

Les estimateurs $\hat{a}_k$ des coefficients s'obtiennent en résolvant le système des équations normales :

$$\begin{cases} \hat{a}_0 A_0 + \hat{a}_1 A_1 ...... + \hat{a}_d A_d = \sum_{i=-p}^{i=+f} X_{t+i} \\ \hat{a}_0 A_1 + \hat{a}_1 A_2 ...... + \hat{a}_d A_{d+1} = \sum_{i=-p}^{i=+f} iX_{t+i} \\ \hat{a}_0 A_2 + \hat{a}_1 A_3 ..... + \hat{a}_d A_{d+2} = \sum_{i=-p}^{i=+f} i^2 X_{t+i} \\ .... \\ \hat{a}_0 A_d + \hat{a}_1 A_{d+1} .. + \hat{a}_d A_{2d} = \sum_{i=-p}^{i=+f} i^t X_{t+i} \end{cases}$$

Soit, en désignant par Ξ, la matrice colonne des seconds membres des équations, par Ω la matrice symétrique des A_n , par Λ la matrice colonne des coefficients à estimer et par ${}^t v$ la matrice (1,0,0, ,,,,,0), on a :

$$\Omega\,\Lambda = \Xi \text{ soit } \Lambda = \Omega^{-1}\Xi \text{ et donc } \hat{a}_0 = {}^t v \Omega^{-1} \Xi$$

En posant enfin ${}^t u_i = [1 \;\; i \;\; i^2 \;\; \ldots \;\; i^d\,]$, on obtient :

$$\theta_i = {}^t v\, \Omega^{-1}\, u^i$$

2 - Optique minimisation sous contraintes du pouvoir de réduction

Le problème à résoudre ici est :

$$\left\{\begin{array}{l} \displaystyle \min_{\theta} \sum_{i=-p}^{i=+f} \theta_i \\ \displaystyle \text{sous} \sum_{i=-p}^{i=+f} \theta_i = 1 \quad \text{et} \qquad k \notin \{1, 2, \ldots.., d\} \sum_{i=-p}^{i=+f} i^k \theta_l = 0 \end{array}\right.$$

Soit Λ le vecteur colonne des d+1 paramètres $\lambda_0, \lambda_1 \ldots\ldots \lambda_d$. Le lagrangien du problème s'écrit :

$$L(\theta, \Lambda) = \sum_{i=-p}^{i=+f} \theta_i^2 + \lambda_0 \left(\sum_{i=+p}^{i=+f} \theta_i - 1 \right) + \sum_{k=1}^{k=d} \lambda_k \left(\sum_{i=-p}^{i=+f} i^k \theta_i \right)$$

donc

$$\frac{\delta L}{\delta \theta_i} = 2\,\theta_i + \lambda_0 + \sum_{k=1}^{k=d} \lambda_k\, i^k = 0$$

En multipliant chacune de ces p+f+1 équations par i^k, avec k=0,1, ...,d, et en faisant la somme sur i, on obtient alors le système matriciel suivant :

$$\Omega\,\Lambda = \begin{bmatrix} -2 \\ 0 \\ 0 \\ . \\ . \\ . \\ 0 \end{bmatrix}$$

or

$$\theta_i = -\frac{1}{2} \sum_{k=0}^{k=d} \lambda_k \;\; i^k$$

et donc, avec les notations du paragraphe précédent : $\theta_i = {}^t u_i\, \Omega^{-1} v$, Ce qui, compte tenu de la symétrie de Ω, prouve l'équivalence des deux approches.