COMPRENDRE LA METHODE X11

Dominique LADIRAY[1], Benoît QUENNEVILLE[2]

Juillet 1999

---

[1] Dominique Ladiray est Administrateur de l'Institut National de la Statistique et des Études Économiques, 18 Boulevard Adolphe Pinard, 75014 Paris, France. Ce travail a été réalisé pendant un séjour au Centre de Recherche et d'Analyse des Séries Temporelles, à Statistique Canada.
[2] Benoît Quenneville est Méthodologiste au Centre de Recherche et d'Analyse des Séries Temporelles, 3G-RHC-BSMD, Statistique Canada, Ottawa, Ontario, Canada, K1A 0T6. Tél : 613-951-1605, Fax : 613-951-5711, courriel : quenne@statcan.ca.

RÉSUMÉ

En matière de désaisonnalisation, la méthode statistique la plus utilisée est sans aucun doute celle mise en œuvre dans le logiciel X11. Développé dans les années 50-60 au US Bureau of Census, ce programme a fait l'objet de nombreuses modifications et améliorations, en particulier les logiciels X11-ARIMA (1975, 1988) et le tout nouveau X12-ARIMA (dont une première version de test a été diffusée dès 1998). Si ces logiciels incorporent, à des degrés divers, des méthodes d'analyse paramétrique et en particulier les modèles ARIMA popularisés par Box & Jenkins, ils restent sur le fond très proches de la méthode initiale X11 et c'est à ce "noyau" que nous nous intéressons dans la suite.
Ce document a pour ambition d'expliquer la méthode de désaisonnalisation mise en œuvre dans les logiciels de la « famille X11 ». Après un bref historique de la désaisonnalisation, vous trouverez une présentation générale de la méthode X11 et des moyennes mobiles. Un exemple complet de désaisonnalisation sera ensuite présenté et vous pourrez suivre, dans le détail, l'ensemble des calculs faits. On se concentre ici sur la partie X11 des logiciels actuels c'est-à-dire sans référence à une modélisation ARIMA *a priori* de la série à désaisonnaliser. Les différences, en général minimes, entre X11-ARIMA88 et X12-ARIMA sur le fonctionnement de ce "noyau central" seront précisées le cas échéant.

SUMMARY

When it comes to seasonal adjustment, the most widely used statistical method is without a doubt that implemented in the X11 software. Developed at the US Bureau of the Census in the 1950s and 1960s, this computer program has undergone numerous modifications and improvements, relating especially to the X11-ARIMA software packages (1975, 1988) and the all-new X12-ARIMA (a first beta version of which is available since 1998). While these software packages integrate, to varying degrees, parametric analysis methods, and especially the ARIMA models popularized by Box & Jenkins, they remain in essence very close to the initial X11 method, and it is this "core" that will interest us here.
This work presents the seasonal adjustment methodology implemented in the X11-based softwares. After some historical notes, you will find a presentation of the X11 methodology and a complete description of the moving-averages used in the programs. Readers will also find a complete example of seasonal adjustment, and will have a detailed picture of all the calculations. Emphasis will be placed here on the X11 portion of the current software, i.e. without reference to *a priori* ARIMA modelling of the series to be seasonally adjusted. The generally minor differences between X11-ARIMA88 and X12-ARIMA concerning the operation of this "central core" will be specified where applicable.

TABLE DES MATIÈRES

En matière de désaisonnalisation, la méthode statistique la plus utilisée est sans aucun doute celle mise en œuvre dans le logiciel X11. Développé dans les années 50-60 au US Bureau of Census, ce programme a fait l'objet de nombreuses modifications et améliorations, en particulier les logiciels X11-ARIMA (1975, 1988) et le tout nouveau X12-ARIMA (dont une première version de test a été diffusée dès 1998). Si ces logiciels incorporent, à des degrés divers, des méthodes d'analyse paramétrique et en particulier les modèles ARIMA popularisés par Box & Jenkins, ils restent sur le fond très proches de la méthode initiale X11 et c'est à ce "noyau" que nous nous intéressons dans la suite.
Les détracteurs de X11 ont souvent mis en avant le côté "boîte noire" de ce logiciel. L'absence de modèle explicite, la multiplicité des options et des tableaux en sortie sont sans doute pour beaucoup dans cette appréciation. X11 est pourtant un logiciel reposant sur un principe itératif simple et assez facile à expliquer. Il est vrai cependant qu'il était difficile, pour ne pas dire impossible, à un utilisateur même très averti, de reconstruire et d'expliquer chaque tableau en sortie de X11 : des erreurs mineures de programmation et des imprécisions dans la documentation rendaient cette tâche insurmontable. Ces petites erreurs de programmation ont, pour la plupart, disparu des nouvelles versions du logiciel et nous avons décidé de faire ce travail jamais achevé jusqu'à présent : traiter à fond un exemple de désaisonnalisation par la méthode X11.
Pour ce faire, nous avons programmé la méthode X11 en langage SAS et en Mathematica, vérifiant ainsi chaque étape de la désaisonnalisation et validant pas à pas les résultats de X11-ARIMA88 et X12-ARIMA. Après quelques corrections d'erreurs repérées dans chacun des logiciels, tous ces programmes convergent.

Ce document a pour ambition d’expliquer la méthode de désaisonnalisation mise en œuvre dans les logiciels de la « famille X11 ». Après un bref historique de la désaisonnalisation, vous trouverez une présentation générale de la méthode X11. Le chapitre suivant sera consacré à l’étude des moyennes mobiles. Un exemple complet de désaisonnalisation sera ensuite présenté et vous pourrez suivre, dans le détail, l'ensemble des calculs faits. Les modèles de régression linéaires utilisés pour les effets de jours ouvrables et le procédé de détection et de correction des valeurs atypiques sont étudiés dans l’exemple. L’estimation de l’effet de Pâques fera l’objet d’un chapitre séparé dans la mesure où les modèles utilisés dans X11-ARIMA88 et X12-ARIMA sont sensiblement différents.
On se concentre ici sur la partie X11 des logiciels actuels c'est-à-dire sans référence à une modélisation ARIMA *a priori* de la série à désaisonnaliser. Les différences, en général minimes, entre X11-ARIMA88 et X12-ARIMA sur le fonctionnement de ce "noyau central" seront précisées le cas échéant.

Lorsque par la suite nous ferons référence à X11, c’est à la méthode de désaisonnalisation et non au logiciel X11 que nous ferons allusion.

Chapitre 1 BREF HISTORIQUE DE LA DÉSAISONNALISATION[3]

Il est aujourd'hui usuel de décomposer une série X_t observée, en plusieurs composantes elles-mêmes non observées, selon un modèle du genre : $X_t = T_t + C_t + S_t + I_t$, où T_t, C_t, S_t et I_t désignent respectivement la tendance, le cycle, la saisonnalité et la partie irrégulière. Cette idée est ancienne et c'est sans doute en astronomie qu'il faut en rechercher l'origine.
Au 17ème siècle, les mesures plus précises faites sur les mouvements des planètes semblèrent infirmer les lois de Kepler, et l'idée qu'elles donnaient une approximation de la position de la planète plutôt que sa position exacte fut peu à peu acceptée (Nerlove, Grether, Carvalho, 1979). La position observée était alors considérée comme la somme de la position "théorique" et d'une fluctuation irrégulière. Plus tard, on s'aperçut que les orbites des planètes se modifiaient insensiblement et la distinction fut faite entre mouvements séculaires et périodiques. Le modèle à composantes inobservables était né. L'explication de ces mouvements périodiques ou irréguliers a passionné nombre de mathématiciens en cette fin du 18ème siècle et début du 19ème siècle, parmi lesquels Euler, Lagrange et Laplace.
Cette idée de décomposition d'une série temporelle est apparue dès lors dans les travaux des économistes et certains d'entre eux n'hésitaient pas à reconnaître qu'elle leur venait directement de l'astronomie[4] ou de la météorologie[5]. Parallèlement, le développement des connaissances mathématiques va donner aux chercheurs les moyens de dépasser la simple visualisation graphique dans l'analyse des séries temporelles. Parmi les travaux importants en la matière, il faut citer évidemment ceux de Jean-Baptiste Fourier sur la décomposition d'une série en somme de fonctions trigonométriques[6], travaux qui donneront naissance à l'analyse harmonique et, plus tard, lorsque la notion de processus stochastique sera définie, à l'analyse spectrale. Car, à cette époque, en économie comme dans les autres sciences, l'heure est au déterminisme et à la recherche de lois exactes expliquant tous les phénomènes physiques, économiques, démographiques, biologiques etc.
L'objet de nombreuses études est alors la mise en évidence de "cycles" dont l'étude et l'analyse pourraient permettre d'expliquer et de prévoir les crises économiques (Armatte, 1992). Dans ces conditions, les composantes périodiques de court terme sont de peu d'intérêt et il convient de les éliminer :

"Every kind of periodic fluctuations, whether daily, weekly, quarterly, or yearly, must be detected and exhibited not only as a subject of study in itself, but because we must ascertain and eliminate such periodic variations before we can correctly exhibit those which are irregular or non-periodic and probably of more interest and importance." (Jevons, 1862)

---

[3] Cette partie s'inspire beaucoup de : Armatte (1992), Bell & Hillmer (1984), Hylleberg (1992), Nerlove, Grether & Carvalho, (1979)
[4] Nerlove, Grether, Carvalho (1979) citent plusieurs exemples dont ceux de Cournot (1838) et Jevons (1884). En 1801, l'astronome britannique William Herschel publie un travail mettant en relation les périodicités observées entre les taches solaires et celles du prix du blé.
[5] Difficile de ne pas citer les travaux du météorologue Buys-Ballot qui, en 1847, étudiait les variations périodiques de température en modélisant la "tendance" par un polynôme, la saisonnalité par des indicatrices et faisait implicitement appel à des techniques de régression linéaire pour estimer les paramètres.
[6] *Théorie analytique de la chaleur*, publié en 1807.

Les fin du $19^{\text{ème}}$ et début du $20^{\text{ème}}$ siècles abondent de publications sur la décomposition de séries économiques, de techniques d'estimation des composantes, d'éléments de définition des composantes[7]. Mais c'est sans doute à W.M. Persons, en 1919, que revient le mérite de proposer dans un même travail une méthode "complète" de décomposition incluant une tentative de "définition" et de formalisation des composantes inobservables, un modèle de composition et une méthode d'estimation. Pour ce statisticien, une série temporelle se décompose en quatre types de fluctuations aujourd'hui familières :

1. *A long-time tendency or secular trend ; in many series such as bank clearings or production commodities, this may be termed the growth element ;*
2. *A wavelike or cyclical movement superimposed upon the secular trend ; these curves appear to reach their crests during the periods of industrial prosperity and their troughs during periods of industrial depression, their rise and fall constituting the business cycle ;*
3. *A seasonal movement within the year with a characteristic shape for each series ;*
4. *Residual variation due to developments which affect individual series, or to momentous occurrences such as wars or national catastrophes, which affect a number of series simultaneously.*

Ces composantes sont ensuite combinées selon les schémas de composition, additif ou multiplicatif, bien connus :

Schéma additif : $X_t = T_t + C_t + S_t + I_t$

Schéma multiplicatif : $X_t = T_t * C_t * S_t * I_t$

ou encore : $X_t = T_t * (1 + C_t) * (1 + S_t) * (1 + I_t)$

La plupart des publications de l'époque acceptent ces modèles et "définitions" sans beaucoup de discussion, l'accent étant plutôt mis sur les techniques proprement dites d'ajustement saisonnier ou d'extraction de cycle. De même, d'autres concepts furent étudiés et acceptés : l'idée que la saisonnalité varie dans le temps ; la nécessité de prendre en compte simultanément toutes les composantes lorsqu'on estime la partie saisonnière ; l'impossibilité de décrire les tendances et les cycles par des formules mathématiques simples et explicites ; la nécessité de traiter les points aberrants...., toutes ces idées donnant naissance à des méthodes d'estimation différentes (Menderhausen, 1937).
Les travaux de l'époque étaient cependant essentiellement inspirés par deux grandes méthodes, dont on donnera une brève description pour le cas d'un modèle multiplicatif (Armatte, 1992).

- Dans les années 1910, c'est la méthode des "link relatives", mise au point par Persons, qui a la faveur des économistes statisticiens. Son principe consiste à calculer pour chaque valeur mensuelle de la série le rapport x_t / x_{t-1} de cette valeur à celle du mois précédent, à faire une table de fréquence des valeurs de ces rapports pour les 12 mois, puis à déterminer les médianes de ces douze séries. Ces médianes sont ensuite chaînées par multiplication en prenant une base 100 pour janvier : $S_1 = 100$, $S_i = M_i S_{i-1}$. Ces coefficients saisonniers étant finalement corrigés par un facteur $(S_{13} / S_1)^{1/12}$ pour que $S_1 = S_{13} = 100$.

[7] Par exemple, en 1905, Lucien March distingue "*des changements annuels, des changements polyannuels (décennaux par exemple), des changements séculaires, sans parler des périodes plus courtes qu'une année*" (cité par Yule(1921) et Bell, Hillmer (1984)). On peut aussi citer les travaux précurseurs des sœurs Maballée (1906) cherchant à isoler les points de retournement d'une série à l'aide du corrélogramme.

- La seconde façon de déterminer la tendance est la méthode de la moyenne mobile, méthode appliquée dès 1922 par la Federal Reserve Board[8] et popularisée plus tard par Macaulay (1931). Elle s'appuie sur le calcul d'une moyenne mobile centrée d'ordre 12 pour obtenir une estimation de la tendance. Le rapport entre les données originales et cette estimation fournit une première estimation de la composante saisonnière. Pour en éliminer l'irrégulier, on calcule ensuite les médianes (ou moyennes) de la composante pour chaque mois. Puis on ajuste ces nouveaux indices pour que leur somme fasse 1 et obtenir ainsi les indices saisonniers définitifs.

Bien que très populaires, ces méthodes ont fait l'objet de nombreuses critiques sur le plan théorique. Ainsi, Slutsky (1927) et Yule (1927) montrèrent que l'utilisation de moyennes mobiles pouvait introduire des cycles artificiels dans les données, Fisher (1937) déplorait que l'on applique des méthodes "empiriques" *ad hoc* alors qu'il existait des outils mathématiques adéquats. L'événement le plus important de cette période est sans aucun doute l'apparition des modèles autorégressifs (Yule, 1927) et moyennes mobiles (Slutsky, 1927) pour l'analyse des séries temporelles, soit en d'autres termes, le moyen pour les économistes statisticiens de sortir du cadre déterministe traditionnel en utilisant les premiers processus stochastiques. Mais il faudra attendre de nombreuses années pour que ces méthodes connaissent un certain succès en désaisonnalisation.

Dans les années 30, les méthodes de désaisonnalisation basées sur des techniques de régression ont connu des fortunes diverses. Celles-ci s'appuyaient en général sur une décomposition additive de la série initiale ou d'une transformation simple de cette série, sur une modélisation de la série initiale et de chacune des composantes par des fonctions paramétriques simples et sur l'estimation des paramètres par des méthodes de type "moindres carrés ordinaires". La difficulté de trouver une bonne spécification du modèle, et en particulier les fortes hypothèses nécessaires sur les composantes inobservables, expliquent sans aucun doute le rejet, momentané, de ces méthodes (Bell, Hillmer 1984).

Le développement de l'informatique, après la seconde guerre mondiale, a fortement contribué à la diffusion et à l'amélioration des méthodes de désaisonnalisation. Ainsi, en 1954, Julius Shiskin met au point une méthode ("Method I") sur l'ordinateur Univac du US Bureau of Census. Cette technique de désaisonnalisation sera suivie par onze versions eXpérimentales d'une "Method II" (X0, X1, etc) pour finalement aboutir au logiciel ***X11*** en 1965 (Shiskin, Young, Musgrave 1967). Inspirées directement des lissages par moyennes mobiles et des travaux de Macaulay (1931), ces diverses versions constituaient les premières méthodes automatiques d'ajustement saisonnier et ***X11*** devint rapidement un standard utilisé dans le monde entier. Ces nouvelles possibilités de calcul facilitèrent l'utilisation des techniques paramétriques de régression pour l'estimation et la correction des effets de calendrier (jours ouvrables, jours fériés, vacances ....). Un traitement automatique de ces effets, basé sur les travaux de Young (1965), fut d'ailleurs intégré dans le logiciel ***X11***.

Parallèlement, la modélisation paramétrique des séries temporelles et l'analyse spectrale ont fait des progrès importants, essentiellement grâce au développement de la théorie des processus stochastiques, progrès dont les outils de désaisonnalisation ont bénéficié peu à peu.

L'analyse harmonique, employée très tôt pour résoudre des problèmes de décomposition de série, se plaçait dans un cadre résolument déterministe et était utilisée pour mettre en évidence des périodicités exactes alors qu'on était bien conscient que les cycles pouvaient n'être pas rigoureusement périodiques ou que les saisonnalités pouvaient évoluer. Là encore, il fallut

[8]C'est au physicien anglais Poynting qu'on attribue souvent la première utilisation en 1884 d'une moyenne mobile pour éliminer la tendance et isoler les fluctuations de la série qu'il peut ainsi traiter par l'analyse harmonique.

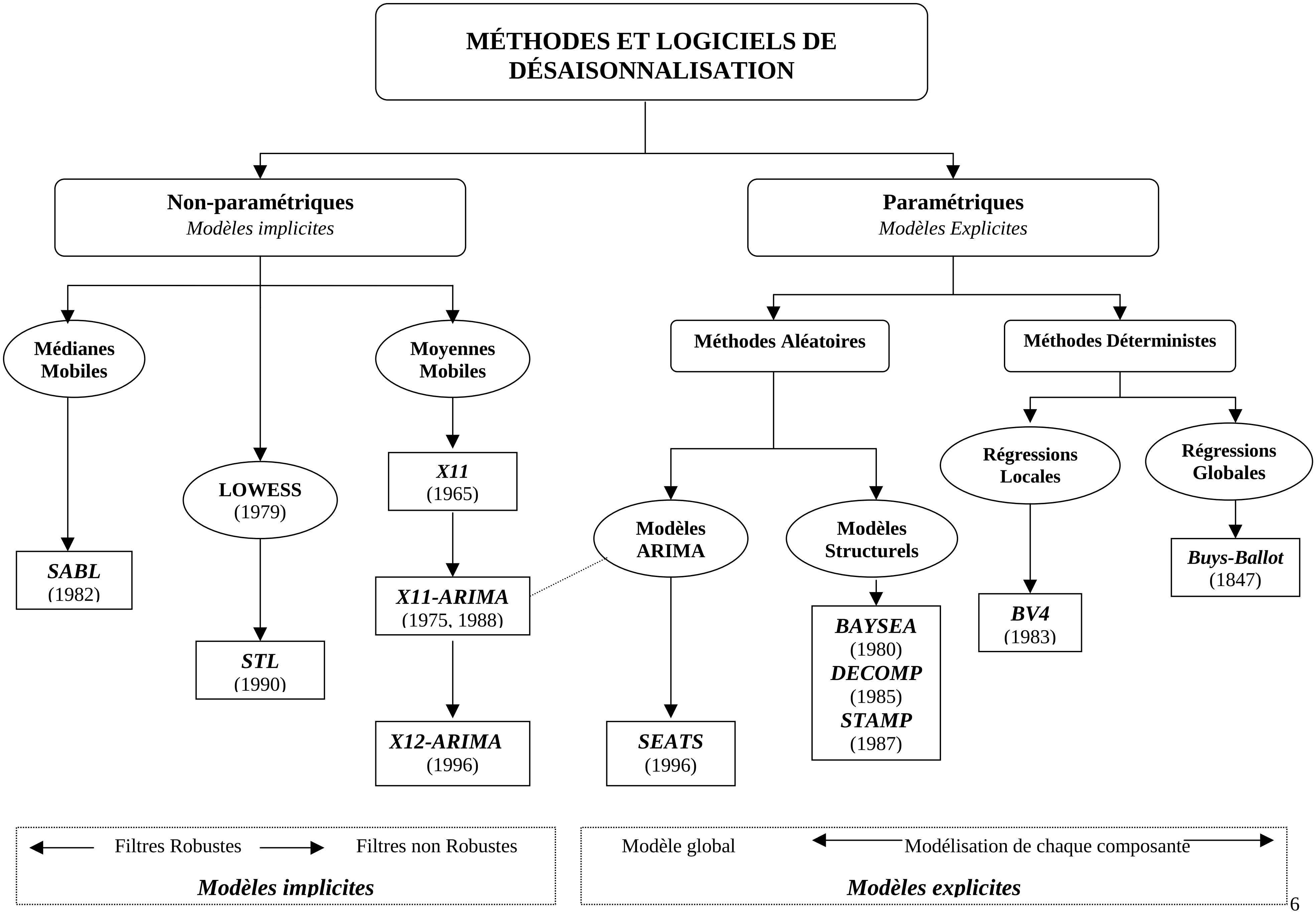
attendre les ordinateurs des années 60 pour profiter des progrès de la théorie et mieux utiliser l'analyse spectrale : estimations améliorées de la densité spectrale (Bartlett 1950, Tukey 1950), étude des processus non-stationnaires (Priestley 1965), transformée de Fourier rapide (Cooley, Tukey 1965) etc.
La popularisation des modèles ARIMA à partir de la publication de Box et Jenkins (1970) a permis de faire progresser les outils de désaisonnalisation dans deux directions. D'une part, elle a constitué un développement important de ***X11*** qui a évolué vers ***X11-ARIMA*** en 1975 (Dagum, 1980). Dans cette nouvelle version, les modèles ARIMA sont utilisés pour prolonger la série initiale afin de limiter les révisions des estimations lorsque l'on dispose d'un point supplémentaire. D'autre part, la modélisation ARIMA a aussi été introduite dans les méthodes de désaisonnalisation fondées sur la théorie de l'extraction du signal. Les exemples de travaux utilisant la modélisation ARIMA et l'analyse spectrale à des fins de désaisonnalisation sont nombreux (voir Bell, Hillmer 1984).
Toutes ces avancées, techniques et théoriques, font qu'aujourd'hui les méthodes de décomposition fondées sur des modèles se développent et se popularisent.

Aujourd'hui, les deux grandes philosophies de la désaisonnalisation, à savoir l'approche empirique et l'approche par modélisation, inspirent diverses méthodes, dont certaines mêlent les deux philosophies. Parmi les logiciels actuellement les plus utilisés, on peut citer ***BV4*** (Technische Universität Berlin, Deutsche Institut für Wirtschaftsforschung), ***SABL*** et ***STL*** (Bell Laboratories), ***X11-ARIMA88*** (Statistique Canada), ***STAMP*** (London School of Economics and Political Science), ***BAYSEA*** et ***DECOMP*** (Institut de Statistique Mathématique, Japon). D'autres, mis au point plus récemment, devraient connaître un succès certain : ***X12-ARIMA*** (US Bureau of Census) et ***TRAMO-SEATS*** (Banco de Espana).
Les principales critiques que l'on peut faire à chacune des deux approches sont difficiles à éviter. Ainsi, reproche-t-on aux méthodes empiriques de ne pas s'appuyer sur des modèles explicites, ce qui rend particulièrement difficile, voire impossible, la connaissance des propriétés statistiques des estimateurs utilisés ; les méthodes fondées sur les modèles sont satisfaisantes sur ce plan mais on s'interroge sur la pertinence de la modélisation et la robustesse des méthodes d'estimation dans le cas de séries très perturbées, on invoque les difficultés de modéliser *a priori* des composantes sur lesquelles on sait peu de choses et la relative faiblesse de la théorie statistique pour les séries non stationnaires.
C'est pourquoi les améliorations apportées par les recherches récentes ne concernent pas le principe même des méthodes existantes mais visent plutôt à corriger certains de leurs défauts. Les principales préoccupations sont tournées d'une part vers les problèmes liés à l'estimation des composantes en début et fin de série, et d'autre part vers l'élimination des divers effets perturbateurs qui influencent les résultats de la désaisonnalisation (points atypiques, changements de régime, effets de calendrier...).

Schématiquement, comme le graphique qui suit le synthétise, les méthodes de désaisonnalisation peuvent être classées en deux grandes catégories : les méthodes non paramétriques et les méthodes paramétriques.

Les méthodes non paramétriques, dites « empiriques », permettent de décomposer la série en composantes inobservables par une procédure, souvent itérative, basée sur des lissages successifs. On peut résumer l'ensemble des lisseurs utilisés dans ces méthodes sous le nom de "régressions locales". Les régressions locales consistent à ajuster des polynômes, en général par les moindres carrés, pondérés ou non, sur des intervalles glissants (se décalant d'un point à chaque fois). Au centre de l'intervalle, la donnée lissée est la valeur, à cette date, du polynôme ajusté (la donnée lissée à la date suivante est obtenue par ajustement d'un polynôme sur l'intervalle suivant). On peut montrer que les régressions locales reviennent à appliquer des moyennes mobiles particulières lorsque les observations sont régulièrement espacées. Les méthodes se distinguent essentiellement par leur degré de robustesse : dans un premier groupe, on trouve ***STL*** (Cleveland et *alii*, 1990), méthode fondée sur le "lowess", une technique de lissage robuste par régressions locales (Cleveland, 1979) ; dans un second groupe figure la célèbre méthode de désaisonnalisation ***X11***.

Les méthodes paramétriques peuvent elles aussi se diviser en deux grands ensembles : les méthodes par régression (inspirées de Buys-Ballot) qui posent pour chaque composante, excepté l'irrégulier, une fonction déterministe du temps ; les méthodes fondées sur des modèles stochastiques (non déterministes) : il s'agit principalement de modèles ARIMA utilisés pour modéliser les composantes inobservables. Parmi celles-ci on distingue encore deux groupes : celles qui estiment les modèles des composantes à partir du modèle ARIMA de la série initiale (Burman 1980, Hillmer et Tiao 1982), ***SEATS*** (Gomez, Maravall 1997) est la plus récente, et celles qui les modélisent et les estiment directement (Akaike 1980, Kitagawa et Gersch 1984) comme par exemple la méthode ***STAMP*** (Harvey et *alii*, 1995).



MÉTHODES ET LOGICIELS DE DÉSAISONNALISATION
Non-paramétriques
Modèles implicites
Paramétriques
Modèles Explicites
Médianes Mobiles
Moyennes Mobiles
Méthodes Aléatoires
Méthodes Déterministes
LOWESS (1979)
X11 (1965)
Régressions Locales
Régressions Globales
Modèles ARIMA
Modèles Structurels
SABL (1982)
X11-ARIMA (1975, 1988)
Buys-Ballot (1847)
BAYSEA (1980)
DECOMP (1985)
STAMP (1987)
BV4 (1983)
STL (1990)
X12-ARIMA (1996)
SEATS (1996)
Filtres Robustes
Filtres non Robustes
Modèles implicites
Modèle global
Modélisation de chaque composante
Modèles explicites

Chapitre 2 LA PHILOSOPHIE DE LA MÉTHODE X11

La méthode X11 repose sur un principe itératif d'estimation des différentes composantes, cette estimation étant faite à chaque étape grâce à des moyennes mobiles adéquates.

2.1 Composantes et schémas de composition.

La méthode X11 permet de décomposer et de désaisonnaliser des séries mensuelles et trimestrielles. Les composantes qui peuvent apparaître à un moment ou à un autre de la décomposition sont :

1. La tendance de la série qui représente l'évolution de long terme de la série ;
2. Le cycle, mouvement lisse, quasi périodique, autour de la tendance et met en évidence une succession de phases de croissance et de récession.

X11 ne sépare pas ces deux composantes : les séries étudiées sont en général trop courtes pour que l'estimation des deux composantes puisse se faire aisément. On parlera donc dans la suite de composante tendance-cycle, notée TC_t .

3. La composante saisonnière, notée S_t, représentant des fluctuations infra annuelles, mensuelles ou trimestrielles, qui se répètent plus ou moins régulièrement d'année en année ;
4. Une composante dite de « jours ouvrables », notée D_t, qui mesure l'impact sur la série de la composition journalière du mois ou du trimestre ;
5. Une composante mesurant l'effet de la fête de Pâques, notée E_t ;
6. Et enfin, la composante irrégulière, notée I_t, regroupant toutes les autres fluctuations plus ou moins erratiques non prises en compte dans les composantes précédentes.

Remarquons que ces définitions sont de fait qualitatives et peu précises. Elles restent d'ailleurs aujourd'hui l'objet de controverses et d'interprétations diverses. A titre d'exemple, voici deux citations d'éminents statisticiens qui, à l'évidence, n'ont pas le même objectif :

Sir KENDALL : *« ...The essential idea of trend is that it shall be smooth. »*[9]
Andrew HARVEY : « *There is no fundamental reason, though, why a trend should be smooth* »[10]

Dans la méthode X11, les composantes sont en fait définies de façon implicite par les outils qui servent à les estimer.

La méthode X11 considère trois schémas de composition possibles :

1. Le modèle additif : $X_t = TC_t + S_t + D_t + E_t + I_t$
2. Le modèle multiplicatif : $X_t = TC_t * S_t * D_t * E_t * I_t$
3. Le modèle log-additif : $X_t = Log(TC_t) + Log(S_t) + Log(D_t) + Log(E_t) + Log(I_t)$

X12-ARIMA propose en outre un modèle « pseudo-additif » : $X_t = TC_t(S_t + D_t + E_t + I_t - 1)$

[9] Dans son ouvrage « Time Series », 1973, p 29
[10] Dans son ouvrage « Forecasting, Structural Time Series Models and the Kalman Filter », 1989, p 284

2.2 Moyennes mobiles

Les ***moyennes mobiles***, qui constituent l'outil de base de la méthode de désaisonnalisation X11, sont utilisées pour estimer les principales composantes de la série, tendance et saisonnalité. Ce sont avant tout des outils de **lissage** conçus pour éliminer une composante indésirable de la série. Prenons l'exemple simple d'une série constituée d'une tendance et d'une composante irrégulière : si la tendance est lisse, alors les valeurs de la série autour de la date t doivent contenir de l'information sur la valeur de cette tendance à l'instant t et il doit être possible d'utiliser une moyenne de ces valeurs comme estimation.

Une moyenne mobile de coefficients $\{\theta_i\}$ se définit donc comme $M(X_t) = \hat{T}_t = \sum_{i=-p}^{+f} \theta_i X_{t+i}$ [11] et tout le problème est alors de trouver le « bon » ensemble de coefficients $\{\theta_i\}$. Les possibilités de calcul très limitées à la fin du siècle dernier ont conduit les statisticiens à chercher des coefficients de pondération indépendants des valeurs de la série selon des méthodes qui seront étudiées en détail dans un prochain chapitre.

2.3 Un algorithme simple de désaisonnalisation

Soit une série brute mensuelle X_t que nous supposerons ici décomposée en tendance-cycle, saisonnalité et partie irrégulière selon un schéma additif : $X_t = TC_t + S_t + I_t$. On peut imaginer un algorithme simple de désaisonnalisation en quatre étapes :

1. ***Estimation de la Tendance-Cycle par moyenne mobile :***

$$TC_t^{(1)} = M_0(X_t)$$

La moyenne mobile utilisée ici devra donc reproduire au mieux la composante tendance-cycle tout en éliminant la composante saisonnière et en réduisant la composante irrégulière au maximum.

2. ***Estimation de la composante Saisonnier-Irrégulier :***

$$(S_t + I_t)^{(1)} = X_t - TC_t^{(1)}$$

3. ***Estimation de la composante Saisonnière par moyenne mobile sur chaque mois :***

$$S_t^{(1)} = M_1\left[(S_t + I_t)^{(1)}\right] \qquad \text{et donc aussi } I_t^{(1)} = (S_t + I_t)^{(1)} - S_t^{(1)}$$

Il s'agit donc ici de lisser les valeurs de la composante saisonnier-irrégulier de chaque mois pour extraire l'évolution du coefficient saisonnier du mois concerné. La moyenne mobile utilisée ici devra reproduire au mieux la composante saisonnière de chaque mois en réduisant au maximum la composante irrégulière.
On peut imposer ici une contrainte de normalisation des coefficients qui leur imposera par exemple d'être de somme nulle.

4. ***Estimation de la série corrigée des variations saisonnières :***

$$Xsa_t^{(1)} = (TC_t + I_t)^{(1)} = X_t - S_t^{(1)}$$

[11] 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.

Toute la difficulté réside donc dans le choix des moyennes mobiles utilisées aux étapes 1 et 3.

2.4 L'algorithme de base de la méthode X11.

C'est cet algorithme simple que la méthode X11 met finalement en œuvre, en utilisant des moyennes mobiles judicieusement choisies et en affinant peu à peu, par itération de l'algorithme, les estimations des composantes.
On peut ainsi définir l'algorithme de base de la méthode X11 qui correspond en fait à utiliser deux fois l'algorithme simple précédent en changeant à chaque fois les moyennes mobiles.

1. Estimation de la Tendance-Cycle par moyenne mobile 2x12 :

$$TC_t^{(1)} = M_{2x12}(X_t)$$

La moyenne mobile utilisée ici est une moyenne mobile dite « 2x12 », de coefficients $\frac{1}{24}\{1,2,2,2,2,2,2,2,2,2,2,2,1\}$, qui conserve les tendances linéaires, élimine les saisonnalités constantes d'ordre 12 et minimise la variance de la partie irrégulière.

2. Estimation de la composante Saisonnier-Irrégulier :

$$(S_t + I_t)^{(1)} = X_t - TC_t^{(1)}$$

3. Estimation de la composante Saisonnière par moyenne mobile 3x3 sur chaque mois :

$$S_t^{(1)} = M_{3x3}\left[(S_t + I_t)^{(1)}\right]$$

La moyenne mobile utilisée ici est une moyenne mobile sur 5 termes, dite « 3x3 », de coefficients $\frac{1}{9}\{1,2,3,2,1\}$, et qui conserve les tendances linéaires. Les coefficients sont ensuite normalisés de telle sorte que leur somme sur toute période de 12 mois soit approximativement nulle.

$$Snorm_t^{(1)} = S_t^{(1)} - M_{2x12}\left(S_t^{(1)}\right)$$

4. Estimation de la série corrigée des variations saisonnières :

$$Xsa_t^{(1)} = (TC_t + I_t)_t^{(1)} = X_t - Snorm_t^{(1)}$$

Cette première estimation de la série corrigée des variations saisonnières doit, par construction, contenir moins de saisonnalité. La méthode X11 remet en œuvre notre algorithme simple en changeant les moyennes mobiles pour tenir compte de cette propriété.

5. Estimation de la Tendance-Cycle par moyenne mobile de Henderson sur 13 termes :

$$TC_t^{(2)} = H_{13}\left(Xsa_t^{(1)}\right)$$

Les moyennes mobiles de Henderson, si elles n'ont pas de propriétés spéciales en terme d'élimination de la saisonnalité (il n'y en a pas ou peu à ce stade), ont un très bon pouvoir de lissage et conservent les tendances localement polynômiales de degré 2.

6. Estimation de la composante Saisonnier-Irrégulier :

$$(S_t + I_t)^{(2)} = X_t - TC_t^{(2)}$$

7. Estimation de la composante Saisonnière par moyenne mobile 3x5 sur chaque mois :

$$S_t^{(2)} = MM_{3x5}\left[(S_t + I_t)^{(2)}\right]$$

La moyenne mobile utilisée ici est une moyenne mobile sur 7 termes, dite « 3x5 », de coefficients $\frac{1}{15}\{1,2,3,3,3,2,1\}$, et qui conserve les tendances linéaires. Les coefficients sont ensuite normalisés de telle sorte que leur somme sur toute période de 12 mois soit approximativement nulle.

$$Snorm_t^{(2)} = S_t^{(2)} - M_{2x12}\left(S_t^{(2)}\right)$$

8. Estimation de la série corrigée des variations saisonnières :

$$Xsa_t^{(2)} = (TC_t + I_t)^{(2)} = X_t - Snorm_t^{(2)}$$

ALGORITHME DE BASE DE X11

Série brute X_t mensuelle : $X_t = TC_t + S_t + I_t$

1. **Estimation de la Tendance-Cycle par moyenne mobile 2x12 :**
   $TC_t^{(1)} = M_{2x12}(X_t)$
2. **Estimation de la composante Saisonnier-Irrégulier :**
   $(S_t + I_t)^{(1)} = X_t - TC_t^{(1)}$
3. **Estimation de la composante Saisonnière par moyenne mobile 3x3 sur chaque mois**
   $S_t^{(1)} = M_{3x3}\left[(S_t + I_t)^{(1)}\right]$ et normalisation $Snorm_t^{(1)} = S_t^{(1)} - M_{2x12}\left(S_t^{(1)}\right)$
4. **Estimation de la série corrigée des variations saisonnières**
   $Xsa_t^{(1)} = (TC_t + I_t)^{(1)} = X_t - Snorm_t^{(1)}$

5. **Estimation de la Tendance-Cycle par une moyenne de Henderson (sur 13 termes)**
   $TC_t^{(2)} = H_{13}\left(Xsa_t^{(1)}\right)$
6. **Estimation de la composante Saisonnier-Irrégulier**
   $(S_t + I_t)^{(2)} = X_t - TC_t^{(2)}$
7. **Estimation de la composante Saisonnière par moyenne mobile 3x5 sur chaque mois**
   $S_t^{(2)} = M_{3x5}\left[(S_t + I_t)^{(2)}\right]$ et normalisation $Snorm_t^{(2)} = S_t^{(2)} - M_{2x12}\left(S_t^{(2)}\right)$
8. **Estimation de la série corrigée des variations saisonnières**
   $Xsa_t^{(2)} = (TC_t + I_t)^{(2)} = X_t - Snorm_t^{(2)}$

2.5 Points aberrants et effets de calendrier.

Comme tout opérateur linéaire, les moyennes mobiles réagissent mal à la présence de valeurs aberrantes. La méthode X11 incorpore donc un ***outil de détection et de correction des points atypiques*** utilisé pour nettoyer la série préalablement à la désaisonnalisation.
Par ailleurs d'autres effets que la saisonnalité peuvent expliquer des variations constatées dans la série ; les plus usuels sont des effets liés au calendrier : effet de jours ouvrables, effet de Pâques etc. Ces composantes sont estimées par des ***modèles de régression linéaire,*** à partir de la composante irrégulière[12].
L'algorithme de base de X11 nous permet d'obtenir 3 estimations différentes de la composante irrégulière :

- À l'étape 3, en enlevant l'estimation de la composante saisonnière de l'estimation de la composante saisonnier-irrégulier obtenue à l'étape 2 :
  $I_t^{(1)} = (S_t + I_t)^{(1)} - Snorm_t^{(1)}$.
  X11 va utiliser cette estimation pour détecter et corriger les points aberrants et obtenir une meilleure estimation de la composante saisonnière.
- À l'étape 7, en enlevant l'estimation de la composante saisonnière de l'estimation de la composante saisonnier-irrégulier obtenue à l'étape 6 :
  $I_t^{(2)} = (S_t + I_t)^{(2)} - Snorm_t^{(2)}$
  X11 va utiliser cette estimation à nouveau pour détecter et corriger les points aberrants et obtenir une estimation plus fiable de la composante saisonnière.
- À l'étape 8, en enlevant à l'estimation de la série corrigée des variations saisonnières l'estimation de la composante tendance-cycle obtenue à l'étape 5 :
  $I_t^{(3)} = Xsa_t^{(2)} - TC_t^{(2)}$
  X11 va utiliser cette estimation pour évaluer, par régression linéaire, la composante pour jours ouvrables et détecter et corriger les points aberrants[13].

2.6 Le principe itératif de X11.

Pour évaluer les différentes composantes de la série, en tenant compte de la présence éventuelle de points aberrants, X11 va procéder de façon itérative : estimation des composantes, recherche des effets gênants dans la composante irrégulière, estimation des composantes sur une série corrigée, recherche des effets gênants dans la composante irrégulière etc.

Le programme X11 présente 4 étapes de traitement notées A, B, C, D plus 3 étapes notées E, F, G qui proposent des statistiques et des graphiques et qui ne font pas partie de la décomposition à proprement parlé.

2.6.1 Étape A : ajustements préalables

Cette étape, qui n'est pas obligatoire, permet à l'utilisateur d'effectuer une correction *a priori* de la série en introduisant des facteurs d'ajustement. Il peut ainsi :

[12] X12 Arima possède un module dit « Reg Arima » qui permet d'estimer directement ces effets sur la série brute, avant de procéder à la désaisonnalisation. Ce module ne sera pas étudié ici.
[13] Ces diverses méthodes seront présentées dans un prochain chapitre.

- introduire des coefficients mensuels (ou trimestriels) d'ajustement qui lui permettront de corriger l'effet de certains jours fériés, de modifier le niveau de la série (effet d'une grève …)
- dans le cas mensuel seulement, introduire 7 coefficients de pondération, à raison d'un par jour pour prendre en compte les variations de la série imputables à la composition des mois en « jours ouvrables ».

A partir de ces données, le programme calcule des coefficients de correction qui sont appliqués à la série brute. La série ainsi corrigée, tableau B1 des sorties imprimées, passe alors à l'étape B.

2.6.2 Étape B : Première correction automatique de la série

Cette étape consiste fondamentalement en une première estimation des points aberrants et, si on le demande, des effets liés aux jours ouvrables. Cette estimation se fait par application de l'algorithme de base détaillé ci-dessus.
Ces traitements aboutissent aux tableaux B19, évaluation des effets de jours ouvrables, et B20, valeurs de correction des points jugés atypiques, qui servent à corriger la série brute et qui conduisent à la série du tableau C1.

2.6.3 Étape C : Seconde correction automatique de la série

Cette étape, en appliquant toujours notre algorithme de base, aboutit à une estimation plus précise des effets pour jours ouvrables (tableau C19) et des valeurs de correction des points atypiques (tableau C20).
La série, enfin « nettoyée », figure dans le tableau D1 des sorties imprimées.

2.6.4 Étape D : Désaisonnalisation

Cette étape, qui applique pour la dernière fois notre algorithme de base, est celle de la désaisonnalisation proprement dite puisqu'elle aboutit aux estimations finales :

- de la composante saisonnière (tableau D10),
- de la série corrigée des variations saisonnières (tableau D11),
- de la composante tendance-cycle (tableau D12),
- de la composante aléatoire (tableau D13).

2.6.5 Étapes E, F et G

Les étapes E et F proposent des statistiques qui permettent de juger de la qualité de la désaisonnalisation.
La partie G propose des graphiques en mode caractère. Elle peut être oubliée dans la mesure où elle peut être aujourd'hui avantageusement remplacée par les logiciels graphiques usuels de bureautique.

2.6.6 Résumé

SCHÉMA SIMPLIFIÉ DU FONCTIONNEMENT DE X11

Étape A : Correction « manuelle »

- pour aléas importants
- pour jours ouvrables

↓

Étape B : Première correction automatique de la série

Estimation de la composante aléatoire

Détection et correction automatique des points « aberrants »

Correction des effets de jours ouvrables

↓

Etape C : Seconde correction automatique de la série

Estimation de la composante aléatoire

Détection et correction automatique des points « aberrants »

Correction des effets de jours ouvrables

↓

Étape D : Désaisonnalisation

1 - Calcul de la série désaisonnalisée provisoire (tableaux D1 à D6)

2 - Lissage de la série désaisonnalisée par une moyenne mobile de HENDERSON et nouvelle estimation des coefficients saisonniers (tableaux D7 à D10)

3 - Calcul de la série désaisonnalisée définitive (tableau D11) et extraction des composantes tendance-cycle (tableau D12) et aléatoire (tableau D13).

↓

Statistiques et graphiques

Étape E	**Étape F**	**Étape G**

2.7 De X11 à X11-ARIMA et X12-ARIMA.

L'utilisation de moyennes mobiles, comme nous le verrons dans le chapitre qui leur est consacré, pose des problèmes en début et fin de série, notamment en ce qui concerne la stabilité des estimations. Ainsi, lorsque l'on dispose d'un point supplémentaire et que l'on désaisonnalise à nouveau la série avec le logiciel X11, il n'est pas rare de constater des variations sensibles des estimations pour les dates les plus récentes.
Estella B. DAGUM, dès 1975, a proposé de remédier en grande partie à ces problèmes en utilisant les modèles ARIMA popularisés quelques années plus tôt par les travaux de BOX et JENKINS (1970). Elle a ainsi montré qu'on diminuait sensiblement les révisions en ajustant un modèle ARIMA à la série, en prévoyant les valeurs futures de la série grâce à ce modèle et en appliquant la procédure de désaisonnalisation X11 à cette série ainsi prolongée. C'est cette idée qui est à la base du logiciel X11-ARIMA (Dagum, 1988)[14].

Malheureusement, l'estimation de modèles ARIMA est rendue délicate par la présence de points aberrants, de rupture de niveau, d'effets de calendrier …. X11-ARIMA repose alors sur le schéma suivant :

1. **Première désaisonnalisation par la méthode X11.**
   Cette étape permet d'estimer les valeurs atypiques, les effets de jours ouvrables comme nous l'avons vu mais aussi les effets de Pâques en utilisant l'estimation de la composante irrégulière du tableau D13.
2. **Modélisation ARIMA de la série corrigée de tous ces effets**
3. **Seconde désaisonnalisation par la méthode X11.**

X12-ARIMA repose sur le même principe mais propose en outre un module très complet, appelé Reg-ARIMA, permettant de corriger la série initiale de toute sorte d'effets indésirables. L'estimation de ces effets se fait grâce à l'utilisation de modèles de régression à erreurs ARIMA (Findlay et *alii*, 1998)

[14] Cette idée avait déjà été implicitement exprimée par Frederick Macaulay (1931) :
" However, graduation of the ends of almost any series is necessarily extremely hypothetical unless facts outside the range covered by the graduation are used in obtaining the graduation …..
Though mathematically inelegant, the most desirable procedure in a majority of the cases of graduation is to graduate not only the actual data, but extrapolated data which sometimes may be extremely crude estimates ".
Rendons hommage à Estella B. DAGUM d'avoir réussi à la mettre en œuvre et à l'imposer.

Chapitre 3 LES MOYENNES MOBILES

La méthode de désaisonnalisation X11 utilise des ***moyennes mobiles*** pour estimer les principales composantes de la série, tendance et saisonnalité. Ces filtres, qui n'impliquent pas *a priori* l'utilisation de concepts ou de modèles sophistiqués, sont très simples de principe et se révèlent d'application particulièrement souple : il est possible de construire une moyenne mobile possédant de bonnes propriétés en termes de conservation de tendance, d'élimination de la saisonnalité, de réduction du bruit etc.

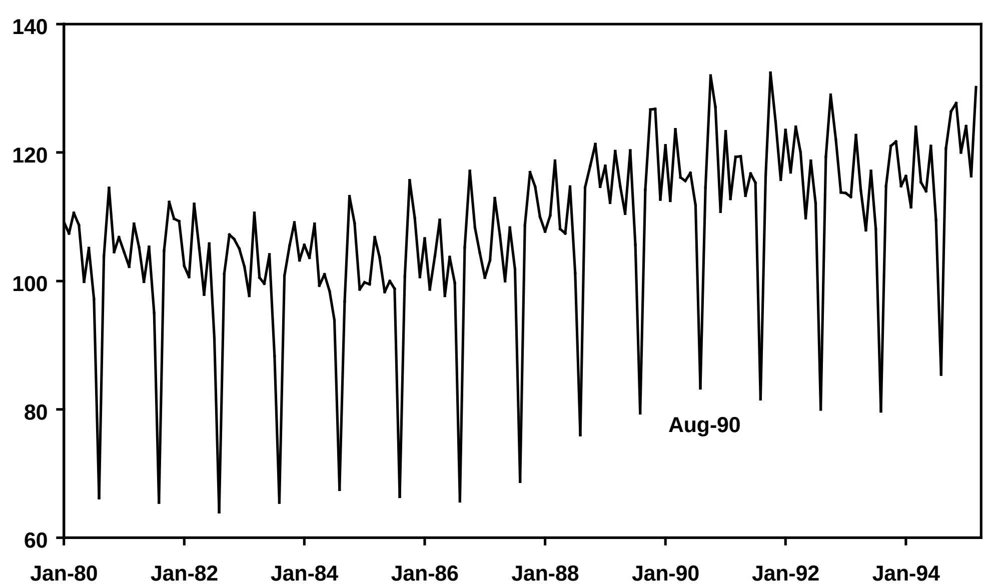
Dans ce chapitre nous allons étudier leurs propriétés et les principes qui ont guidé la construction des moyennes utilisées dans X11.

3.1 Quelques définitions et un peu de théorie

Une série temporelle peut être considérée de deux points de vue : celui des temps et celui des fréquences.

- Dans le *domaine des temps*, on regarde la série $\{X_t\}$ comme une succession de T valeurs observées aux instants *t*, *t* variant de 1 à T. C'est de cette façon que l'on aborde généralement une série temporelle et il est facile de représenter graphiquement, comme dans la figure 1, son évolution au cours du temps. On note que cette série est caractérisée par une forte saisonnalité traduisant la chute de l'activité industrielle au mois d'août.

Figure 1 : Évolution mensuelle de l'indice de la production industrielle française



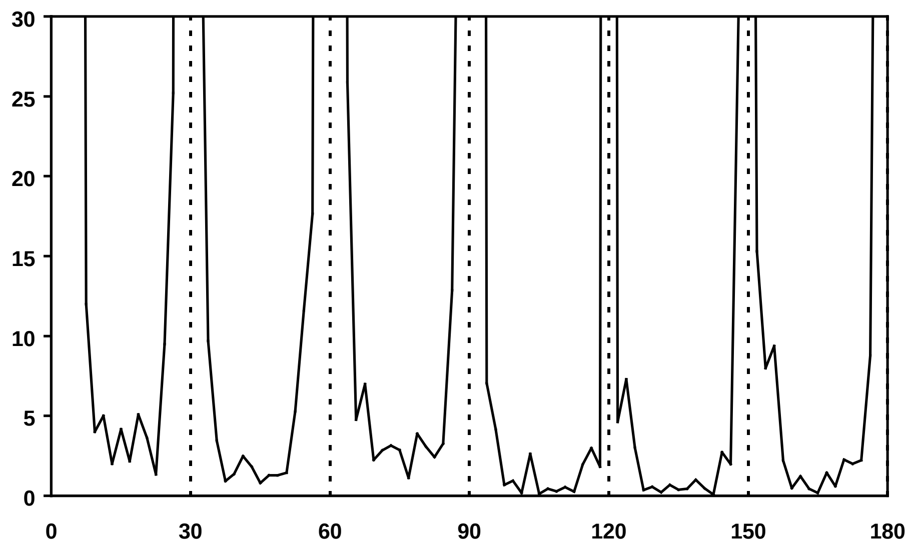
Les modélisations de la série, ou de ses composantes, mettant en relation la valeur à l'instant *t* et celles des instants passés sont particulièrement aisées à formaliser. C'est le cas par exemple de la modélisation de la série par un modèle ARIMA saisonnier, de l'expression d'une tendance linéaire, exponentielle ou encore localement polynomiale, ou de la modélisation de la composante irrégulière par un bruit blanc.

- Dans le *domaine des fréquences* au contraire, on part de l'expression de la série $\{X_t\}$ comme somme de fonctions sinusoïdales[15]. On mesure alors pour chaque fréquence, l'importance qu'elle a dans la composition de la série : le graphique qui associe à chaque fréquence son importance dans la série s'appelle le *spectre* de la série. Ainsi, le spectre de l'indice de la production industrielle française est représenté dans la figure 2.

  Comme on peut le voir, ce spectre laisse apparaître une forte contribution (on dit un *pic spectral*) des fréquences multiples de $\pi/6$ (30°, 60°, 90°….). La période associée à cette fréquence est $\omega = \frac{2\pi}{f} = 2\pi \Big/ \left[\frac{\pi}{6}\right] = 12$ et nous retrouvons la saisonnalité mensuelle observée sur le graphique précédent.

  Les basses fréquences correspondent par nature à des composantes évoluant lentement, tendance et cycle par exemple, et les hautes fréquences à des composantes évoluant plus vite comme la composante irrégulière.

Figure 2 : Spectre de l'indice de la production industrielle française



Ces deux approches s'avèrent souvent complémentaires et par la suite, nous utiliserons soit l'une, soit l'autre pour montrer les qualités et les défauts des filtres moyennes mobiles.

[15] Dans sa *Théorie Analytique de la Chaleur,* publiée en 1807, Jean-Baptiste FOURIER a établi que toute fonction mathématique pouvait être décomposée en une somme de fonctions sinus et cosinus. Ce théorème a donné naissance tout d'abord à *l'analyse harmonique* puis, lorsqu'il a été généralisé, à *l'analyse spectrale*.

3.1.1 Définitions et exemple

On appelle **moyenne mobile de coefficients** $\{\theta_i\}$ l'opérateur noté $M\{\theta_i\}$**,** ou plus simplement M, défini par: $M(X_t) = \sum_{k=-p}^{+f} \theta_k X_{t+k}$.

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.**
- Lorsque *p* est égal à *f*, c'est à dire qu'on utilise autant de points dans le passé et dans le futur, la moyenne mobile est dite **centrée**.
- Si en outre, on a $\theta_{-k} = \theta_k$ pour tout *k*, la moyenne mobile M est dite **symétrique**. Dans ce cas, lorsqu'il s'agira de lister les coefficients de la moyenne mobile, il suffira de préciser l'ordre de la moyenne mobile et les *k*+1 premiers coefficients (Kendall, 1973) :
  $\frac{1}{24}\{1,2,2,2,2,2,2,2,2,2,2,2,1\}$ s'écrit plus simplement $[13]; \frac{1}{24}\{1,2,2,2,2,2,2\}$

De façon générale, avec une moyenne mobile d'ordre *p*+*f*+1, calculée pour un instant *t* avec *p* points dans le passé et *f* points dans le futur, il sera impossible de lisser les *p* premières valeurs et les *f* dernières valeurs de la série.
Dans la méthode X11, les moyennes mobiles symétriques jouent un grand rôle ; pour éviter la perte d'information aux extrémités de la série, elles sont complétées par des moyennes mobiles asymétriques *ad hoc*.

3.1.2 Fonctions de gain et déphasage

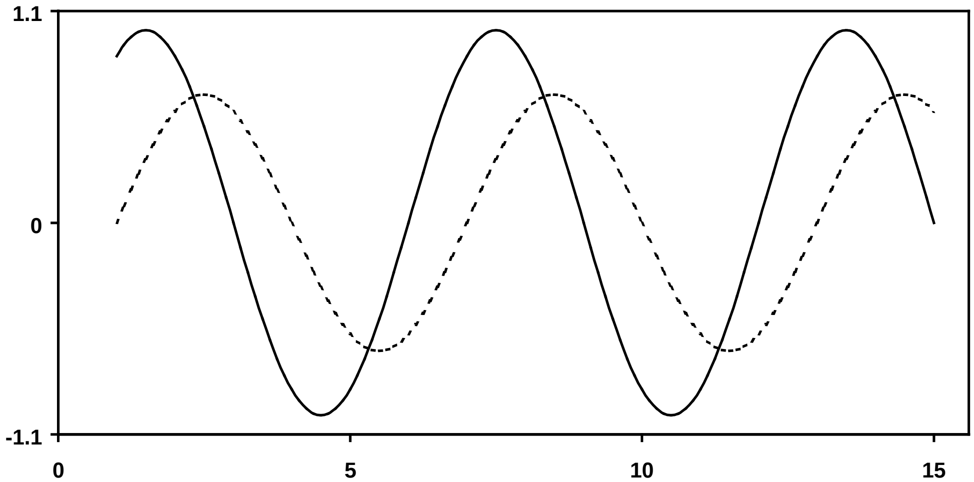
Considérons la série $X_t = \sin\left[\frac{\pi}{3}t\right]$ et transformons la par la moyenne mobile asymétrique définie par $M(X_t) = \frac{1}{3}\left[X_{t-2} + X_{t-1} + X_t\right]$ qui remplace la valeur à l'instant *t* par la moyenne simple des valeurs de l'instant présent et des deux instants précédents.

La figure 3 traduit le résultat du lissage et met en évidence deux phénomènes :

- Tout d'abord une réduction de l'amplitude de la série, qui répond bien à notre objectif de lissage;
- Mais aussi un décalage dans le temps, on dit un *déphasage* : les deux séries ne présentent pas des points de retournement aux mêmes dates.

Ce phénomène de déphasage est désagréable dans la mesure où il transforme les évolutions même de la série. On peut néanmoins démontrer que les moyennes mobiles symétriques n'induisent pas de déphasage (voir par exemple Koopmans, 1974)

Figure 3 :Lissage de la série $X_t = \sin\left[\frac{\pi}{3}t\right]$ par la moyenne mobile $\frac{1}{3}\left[X_{t-2} + X_{t-1} + X_t\right]$



De façon plus générale, soit une série $X_t = R\sin(\omega t + \varphi)$ de fréquence ω (ou de période $2\pi/\omega$), d'amplitude R et de phase φ. La transformée de $\{X_t\}$ par une moyenne mobile quelconque sera aussi une sinusoïde d'amplitude modifiée et présentant un déphasage par rapport à la série originale :

$$M(X_t) = M\left[\sin(\omega t + \varphi)\right] = G(\omega)\sin\left[\omega t + \varphi + \Gamma(\omega)\right]$$

- La fonction qui à ω associe $|G(\omega)|$ s'appelle la ***fonction de gain*** de la moyenne mobile.
- La fonction qui à ω associe $\Gamma(\omega)$ s'appelle la ***fonction de déphasage*** de la moyenne mobile. On représente parfois $\Gamma(\omega)/\omega$ ce qui permet de mesurer le déphasage en nombre de périodes.

Dans le cas de la moyenne mobile asymétrique sur 3 termes ci-dessus on a :

$$M(X_t) = \frac{1}{3}(X_{t-2} + X_{t-1} + X_t) = \frac{1}{3}R\left[\sin(\omega t - 2\omega + \varphi) + \sin(\omega t - \omega + \varphi) + \sin(\omega t + \varphi)\right]$$

$$= \frac{1}{3}R(1 + 2\cos\omega)\sin(\omega t + \varphi - \omega)$$

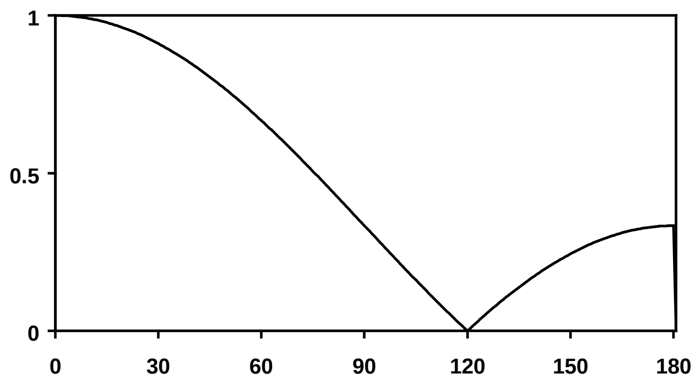
et donc :

- $G(\omega) = \dfrac{1 + 2\cos\omega}{3}$
- $\Gamma(\omega) = -\omega$ soit encore $\Gamma(\omega)/\omega = -1$

La fonction de gain, représentée à la figure 4, montre que la moyenne mobile annule les fréquences $2\pi/3$. Elle serait bien adaptée à des enquêtes ayant lieu tous les 4 mois (donc de période 3) puisqu'elle en éliminerait ainsi la saisonnalité tout en conservant les évolutions de fond correspondant à des basses fréquences. Par contre, cette moyenne introduit un déphasage systématique de une période qui conduirait à prendre trop tard conscience de possibles retournements de tendance.

La fonction de gain permet donc de voir essentiellement les fréquences éliminées et préservées par la moyenne mobile.
La fonction de déphasage montre les décalages introduits par lutilisation de moyennes mobiles asymétriques. Dans la mesure où la méthode X11 met l'accent sur des moyennes mobiles symétriques, nous délaisserons cette fonction dans la suite de ce document.

Figure 4 :Fonction de gain de la moyenne mobile $\frac{1}{3}\left[X_{t-2}+X_{t-1}+X_t\right]$



En matière de lissage, le filtre « idéal » serait celui qui laisserait inchangées les basses fréquences, c'est à dire par exemple les fonctions périodiques de période supérieure à l'année (tendance et cycle), et éliminerait au contraire toutes les hautes fréquences correspondant à des périodicités inférieures ou égales à l'année (saisonnalité et irrégulier). La fonction de gain « idéale » de ce filtre, dit ***filtre « passe-bas »***, serait donc de la forme :

$$G(\omega)=\begin{cases}1 & \text{pour} \quad \omega\leq\omega_0\\ 0 & \text{pour} \quad \omega>\omega_0\end{cases}$$

3.1.3 Conservation de tendance

Le déphasage introduit par la moyenne mobile $\frac{1}{3}\left[X_{t-2}+X_{t-1}+X_t\right]$ peut aussi se voir en appliquant cette moyenne asymétrique à une simple droite $X_t=at+b$. En effet, on a :

$$M(X_t)=\frac{1}{3}(X_{t-2}+X_{t-1}+X_t)=\frac{1}{3}\left[a(t-2)+b+a(t-1)+b+at+b\right]=a(t-1)+b=X_{t-1}$$

Il serait pourtant souhaitable qu'une moyenne mobile respecte certaines tendances simples, en particulier polynomiales.

- Or, pour qu'une moyenne mobile quelconque respecte les séries constantes $X_t=a$, il faut que : $M(X_t)=\sum_{k=-p}^{+f}\theta_k X_{t+k}=\sum_{k=-p}^{+f}\theta_k a=a\sum_{k=-p}^{+f}\theta_k=a$

et donc que la somme des coefficients de la moyenne mobile $\sum_{k=-p}^{+f} \theta_k$ soit égale à 1.

- pour qu'une moyenne mobile quelconque conserve les droites, il faut que pour tout t : $M(X_t) = \sum_{k=-p}^{+f} \theta_k X_{t+k} = \sum_{k=-p}^{+f} \theta_k [a(t+k)+b] = at \sum_{k=-p}^{+f} \theta_k + a \sum_{k=-p}^{+f} k\theta_k + b \sum_{k=-p}^{+f} \theta_k = at + b$,

  ce qui entraîne que $\sum_{k=-p}^{+f} \theta_k = 1$ et $\sum_{k=-p}^{+f} k\theta_k = 0$

- De façon générale, on démontrerait de même que pour qu'une moyenne mobile conserve un polynôme de degré d, il faut et il suffit que ses coefficients vérifient :

  $\sum_{k=-p}^{+f} \theta_k = 1$ et $\sum_{k=-p}^{+f} k^j \theta_k = 0 \quad \mathrm{j} = 1, \ldots\ldots, d$

Ainsi, pour la moyenne mobile asymétrique sur 3 termes définie ci-dessus, on a :

$$\sum_{k=-2}^{0} \theta_k = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1$$

$$\sum_{k=-2}^{0} k\theta_k = -2 * \frac{1}{3} - 1 * \frac{1}{3} + 0 * \frac{1}{3} = -1$$

et cette moyenne, si elle conserve les constantes, ne conserve pas les droites.
Par contre il est facile de vérifier que les moyennes mobiles symétriques suivantes conservent les droites :

$$M(X_t) = \frac{1}{3}(X_{t-1} + X_t + X_{t+1}), \quad M(X_t) = \frac{1}{4}\left(X_{t-2} + 2X_{t-1} + 2X_t + 2X_{t+1} + X_{t+2}\right)$$

3.1.4 Élimination de saisonnalité

Comme nous l'avons vu lors de la définition de la fonction de gain (paragraphe 3.1.2), les moyennes mobiles peuvent éliminer certaines fréquences et donc certaines composantes saisonnières. C'est d'ailleurs cet outil qui permet le plus facilement de repérer les fréquences éliminées par une moyenne mobile quelconque.

Les saisonnalités fixes peuvent être "modélisées" par des fonctions périodiques de période k (4 pour une série trimestrielle, 12 pour une série mensuelle ....). Si on suppose en outre que la somme des coefficients saisonniers est nulle sur une année, on montre (Gouriéroux, Monfort 1990) que ces fonctions engendrent alors un sous espace vectoriel de dimension k-1 dont il est facile d'exhiber une base. Ainsi, dans le cas trimestriel, on trouve le sous espace engendré par les vecteurs colonnes de la matrice :

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

L'annulation de telles séries introduit donc des contraintes sur les coefficients de la moyenne mobile qui s'expriment matriciellement : $C\Theta = \alpha$, où Θ est le vecteur colonne des coefficients de la moyenne mobile M, C la matrice de dimensions (k-1,p+f+1) dont les lignes sont, pour k=4, les vecteurs de base ci-dessus et α la matrice nulle de dimensions (k-1,1).

Par exemple, si nous considérons une moyenne mobile sur 4 termes, on devrait avoir :

$$C\Theta = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ soit encore } \begin{cases} \theta_1 = \theta_2 \\ \theta_1 = \theta_3 \\ \theta_1 = \theta_4 \end{cases} \text{ et donc } \theta_1 = \theta_2 = \theta_3 = \theta_4$$

et cette moyenne mobile élimine les saisonnalités de période 4. En conséquence, sa fonction de gain s'annulera à la fréquence $\pi/2$ (= $2\pi/4$).

De façon générale, une moyenne mobile simple d'ordre k (et donc de coefficients tous égaux à 1/k) annule les saisonnalités fixes de période k et sa fonction de gain s'annule donc en $2\pi/k$.

Par ailleurs, il est possible de traiter le cas de saisonnalités variant linéairement, ou même de façon polynomiale, avec le temps (Grun-Rehomme, Ladiray 1994) : là encore, l'élimination de ces saisonnalités se traduit par des contraintes linéaires sur les coefficients.

3.1.5 Réduction de la composante irrégulière

Après la tendance et la saisonnalité, il nous reste à voir l'effet d'une moyenne mobile sur la composante irrégulière. 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 : $\sigma^{*2} = \sigma^2 \sum_{k=-p}^{+f} \theta_k^2$

Diminuer la composante irrégulière, et donc sa variance, revient à diminuer la quantité : $\sum_{k=-p}^{+f} \theta_k^2$.

3.1.6 Un exemple de construction de moyenne mobile

Cherchons par exemple une moyenne mobile centrée sur trois termes, de coefficients $\{\theta_{-1}, \theta_0, \theta_1\}$ qui réduit au maximum la composante irrégulière et qui conserve les droites. D'après ce qui précède, cela revient à résoudre le problème :

$$\begin{cases} \underset{\theta}{Min} \sum_{k=-1}^{+1} \theta_1^2 \\ \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \theta_{-1} \\ \theta_0 \\ \theta_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{cases} \Leftrightarrow \begin{cases} \underset{\theta}{Min} \sum_{k=-1}^{+1} \theta_k^2 \\ \theta_{-1} + \theta_0 + \theta_1 = 1 \\ -\theta_{-1} + \theta_1 = 0 \end{cases} \Leftrightarrow \begin{cases} \underset{\theta}{Min} \left[2\theta_1^2 + (1 - 2\theta_1)^2 \right] \\ \theta_0 = 1 - 2\theta_1 \\ \theta_{-1} = \theta_1 \end{cases}$$

La dérivée, par rapport à θ_1, de la fonction à minimiser est $12\theta_1 - 4$ qui s'annule donc pour $\theta_1 = 1/3$ et nous trouvons donc la moyenne mobile simple sur 3 termes de coefficients tous égaux à 1/3, moyenne qui, comme nous l'avons déjà vu, élimine les saisonnalités d'ordre 3.

3.2 Les moyennes mobiles symétriques utilisées dans X11

3.2.1 Les moyennes mobiles simples composées.

Une moyenne mobile dite *PxQ* s'obtient en composant une moyenne mobile simple d'ordre P, de coefficients tous égaux à 1/P, et une moyenne mobile simple d'ordre Q, de coefficients tous égaux à 1/Q. Concrètement, cela revient à appliquer successivement à notre série les deux moyennes mobiles simples.
Ainsi, la moyenne mobile 3x3 qui résulte de la double applications de la moyenne mobile arithmétique simple sur 3 termes est une moyenne mobile de coefficients $\left\{\frac{1}{9},\frac{2}{9},\frac{3}{9},\frac{2}{9},\frac{1}{9}\right\}$. De façon générale, une moyenne mobile *PxQ* est une moyenne mobile symétrique d'ordre p+q-1.
Lorsque Q est pair, par exemple égal à *2q*, on a une petite ambiguïté de définition dans la mesure où on peut choisir soit *q* points dans le passé et *q-1* points dans le futur, soit *q-1* points dans le passé et *q* points dans le futur. Le plus souvent, on résout le problème en utilisant une moyenne composée symétrique *2xQ* qui correspond à la moyenne des deux moyennes mobiles candidates.

3.2.1.1 Estimation de la tendance : moyennes 2x4 et 2x12

Lorsque X11 fait une première estimation de la tendance (tableaux B2, C2 et D2), il utilise des moyennes mobiles 2x4 dans le cas trimestriel et 2x12 dans le cas mensuel. A ce moment, la série qui doit être lissée est composée d'une tendance, d'une saisonnalité et d'une composante irrégulière. Dans le cas d'un schéma de composition additif, elle peut s'écrire :
$$X_t = TC_t + S_t + I_t$$

La moyenne 2x4

C'est une moyenne mobile d'ordre 5, de coefficients $\left\{\frac{1}{8},\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{8}\right\}$. La courbe des coefficients et la fonction de gain, présentées à la figure 5, permettent de mettre en évidence les propriétés de cette moyenne mobile :

- Elle élimine la fréquence $\pi/2$ correspondant à la période 4 et de ce fait convient bien aux séries trimestrielles ayant une saisonnalité constante.
- La somme de ses coefficients est égale à 1 et elle est symétrique : elle conserve donc les tendances linéaires.
- La somme des carrés de ses coefficients est égale à 0.250 et elle réduit donc de 75% la variance d'un bruit blanc.

Utiliser cette moyenne mobile revient donc à supposer que la tendance-cycle de notre série est linéaire, ou linéaire par morceaux, que la saisonnalité est constante, ou varie peu dans le temps, et que la composante irrégulière ne possède aucune structure et est de faible amplitude. Dans ce cas, on aura :
$$M_{2x4}(X_t) = M_{2x4}(TC_t + S_t + I_t) = M_{2x4}(TC_t) + M_{2x4}(S_t) + M_{2x4}(I_t) \approx TC_t + 0 + \varepsilon \approx TC_t$$

Par contre cette moyenne mobile restitue assez mal les basses fréquences associées aux périodes supérieures à l'année. Ainsi, les fonctions périodiques de 3 ans, qui correspondent dans le cas trimestriel à des fréquences de $\pi / 6$ (soit 30°), ne sont restituées qu'à 80% environ.

La moyenne 2x12

Cette moyenne repose sur les mêmes idées que celles exposées pour la 2x4 et est utilisée dans le cas mensuel. Ses coefficients sont : $\frac{1}{24}\{1,2,2,2,2,2,2,2,2,2,2,2,1\}$. Elle conserve donc les droites et, comme le montre sa fonction de gain, elle élimine les saisonnalités mensuelles (qui correspondent à une fréquence $\pi/6$). De plus, la somme des carrés de ses coefficients étant égale à 1/12, elle réduit de plus de 90% la variance d'un bruit blanc.
Mais là encore, les séries périodiques de période inférieure à l'année ne sont pas toutes très bien restituées. Ainsi, une fonction périodique de période 3 ans, ce qui correspond ici à une fréquence $\pi / 18$, ne sera restituée qu'à 80%.

Remarques

- Les moyennes mobiles 2x4 et 2x12 sont aussi utilisées dans la méthode X11 pour normaliser les coefficients saisonniers.
- X11-ARIMA88 et X12-ARIMA proposent aussi en option des moyennes mobiles centrées sur 8 et 24 termes dues à Pierre Cholette (1981). Ces moyennes ont été construites sur des critères assez différents de ceux présentés ici, et aussi plus complexes. Elles ne seront pas étudiées ici.

3.2.1.2 Estimation de la saisonnalité : moyennes 3x3, 3x5, 3x9[16]

Ces moyennes sont utilisées par X11 pour extraire la composante saisonnière à partir d'une estimation de la composante saisonnier-irrégulier. Elles apparaissent donc dans la construction des tableaux B4, B5, B9, B10, C5, C10, D5 et D10.
A ce moment, la série qui doit être lissée est composée d'une saisonnalité et d'une composante irrégulière. Dans le cas d'un schéma de composition additif, elle peut s'écrire : $SI_t = S_t + I_t$ et, contrairement au cas précédent, nous sommes en présence d'un strict problème de lissage.

Les coefficients des moyennes 3x3, 3x5 et 3x9 sont les suivants :
$M_{3x3} : \frac{1}{9}\{1,2,3,2,1\}$, $M_{3x5} : \frac{1}{15}\{1,2,3,3,3,2,1\}$ et $M_{3x9} : \frac{1}{27}\{1,2,3,3,3,3,3,3,3,2,1\}$.

Comme on peut le vérifier simplement, chacune de ces moyennes mobiles symétriques conserve les droites. Mais les fonctions de gain, présentées à la figure 5, sont assez différentes de la forme idéale associée à un filtre passe-bas. Ici, ce n'est pas gênant dans la mesure où la composante saisonnier-irrégulier n'est pas sensée présenter de composante cyclique de périodicité de l'ordre de 3 à 6 ans.

En fait appliquer par exemple le filtre 3x3 à chaque mois d'une estimation de la composante Saisonnier-Irrégulier revient à appliquer sur cette composante elle même une moyenne mobile sur 49 termes de coefficients :

[16] X12-ARIMA permet aussi d'utiliser une moyenne mobile 3x15.

$$\frac{1}{9}\{1,0,0,0,0,0,0,0,0,0,0,0,2,0,0,0,0,0,0,0,0,0,0,0,3,0,0,0,0,0,0,0,0,0,0,0,2,0,0,0,0,0,0,0,0,0,0,0,1\}$$

Il en est de même pour les filtres 3x5 et 3x9 auxquels sont associées des moyennes mobiles sur, respectivement, 73 et 121 termes ! Les fonctions de gains de ces nouvelles moyennes mobiles sont présentées à la figure 6. Comme on peut le constater, ces moyennes mobiles conservent la saisonnalité mensuelle puisqu'elles restituent exactement les fréquences multiples de $\frac{\pi}{6}$.

Tableau 1 :Coefficients des moyennes mobiles composées utilisées dans X11, pouvoir de réduction de variance et critère de Henderson.

$\sum \theta_i^2$ représente le pouvoir de réduction de variance (voir 3.1.5)

$\sum \left(\nabla^3 \theta_i\right)^2$ représente la valeur du critère de Henderson (voir 3.2.2)

I	2 x 4	2 x 12	3 x 3	3 x 5	3 x 9
-6		1/24			
-5		1/12			1/27
-4		1/12			2/27
-3		1/12		1/15	3/27
-2	1/8	1/12	1/9	2/15	3/27
-1	1/4	1/12	2/9	3/15	3/27
0	1/4	1/12	3/9	2/15	3/27
1	1/4	1/12	2/9	3/15	3/27
2	1/8	1/12	1/9	2/15	3/27
3		1/12		1/15	3/27
4		1/12			2/27
5		1/12			1/27
6		1/24			
$\sum \theta_i^2$	0.2188	0.0800	0.2346	0.1644	0.1001
$\sum \left(\nabla^3 \theta_i\right)^2$	0.1250	0.0139	0.1481	0.0356	0.0110

Figure 5 : Courbe des coefficients et fonctions de gain des moyennes mobiles composées utilisées dans X11.

Courbe des coefficients | Fonction de Gain

2 x 4

2 x 12

3 x 3

3 x 5

3 x 9

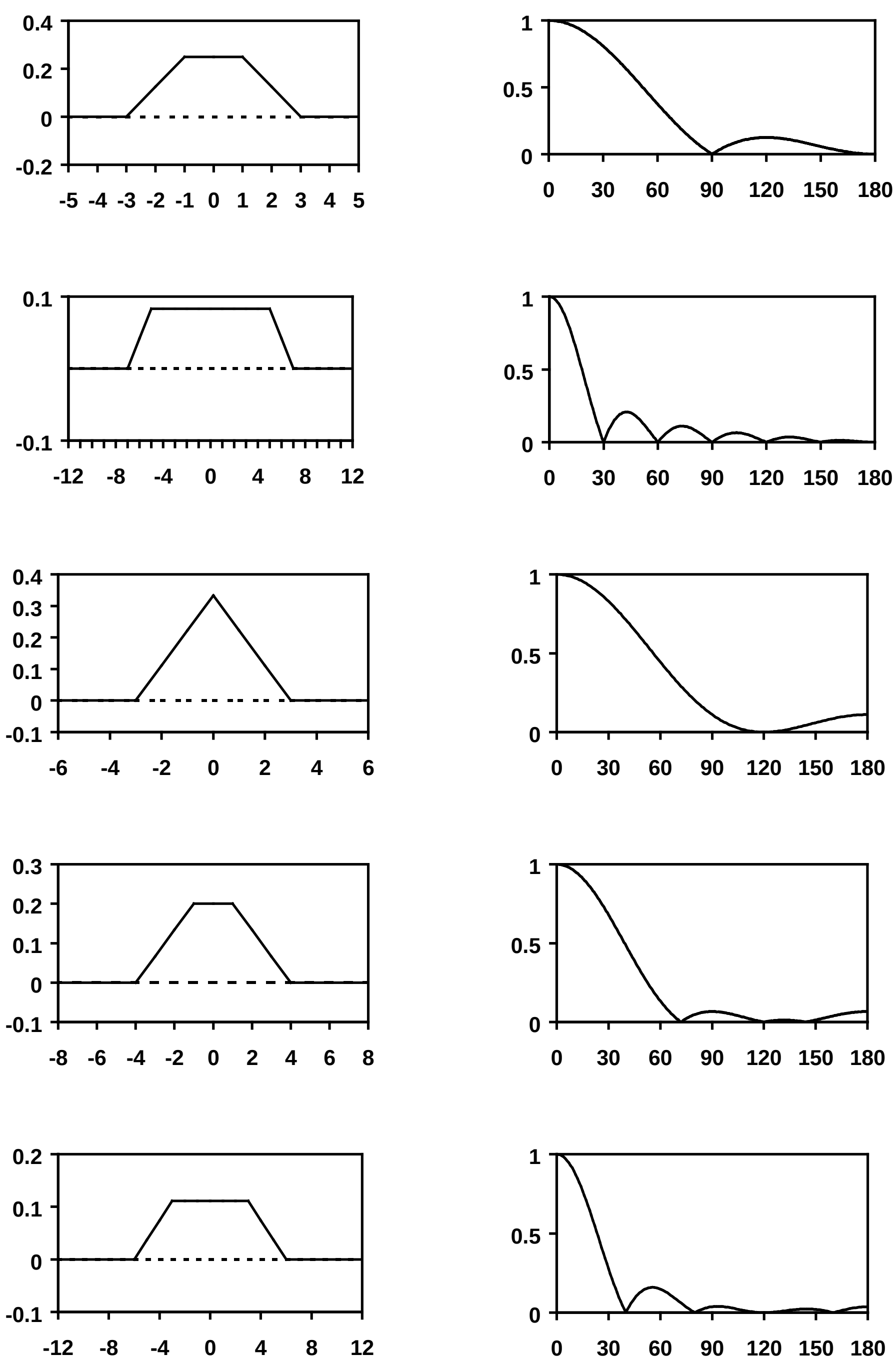
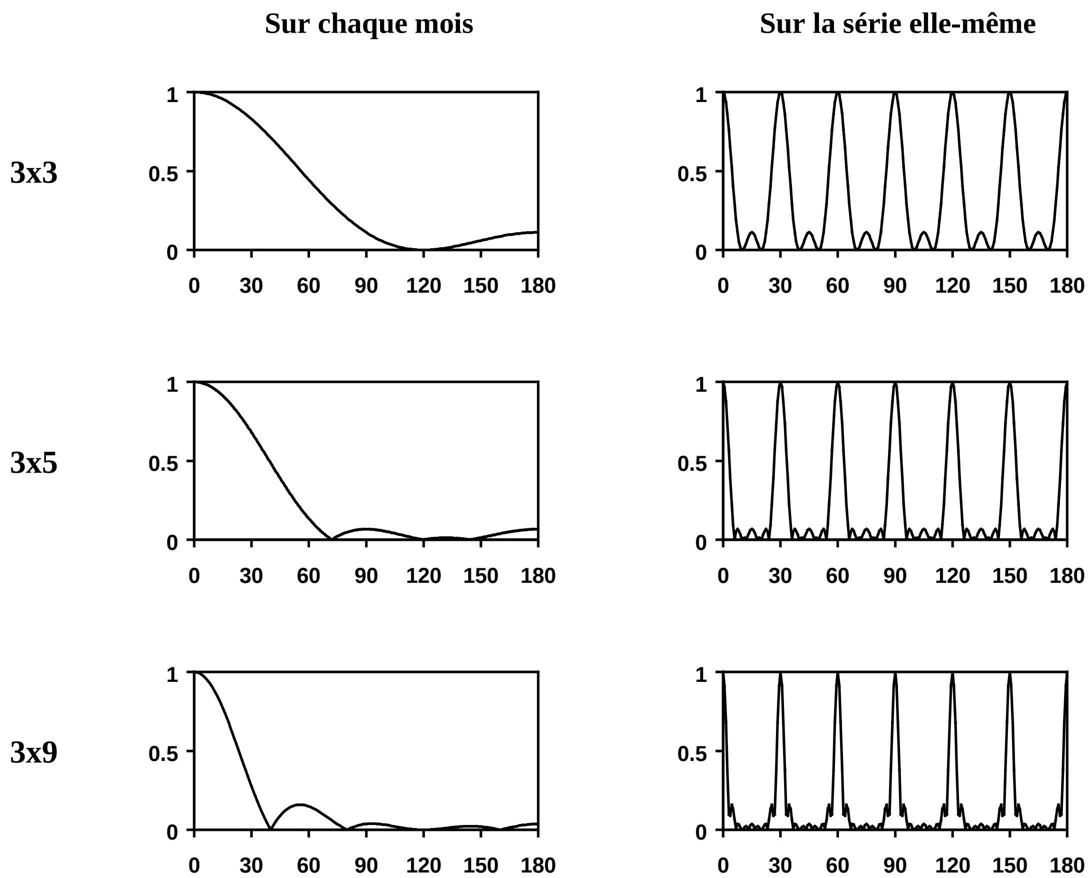


Figure 6 : Fonctions de gains des moyennes 3x3, 3x5 et 3x9



3.2.2 Les moyennes mobiles de Henderson.

Les moyennes mobiles de Henderson sont utilisées dans X11 pour extraire la tendance d'une estimation de la série corrigée des variations saisonnières (tableaux B7, C7, D7, D12). Dans le cas additif, on a donc un modèle du type : $Xsa_t = TC_t + I_t$.

Quel critère peut-on construire qui assurera une estimation lisse de la tendance-cycle ?

Considérons la série $X_t = \begin{cases} 1 & \text{si} \quad t = 0 \\ 0 & \text{si} \quad t \neq 0 \end{cases}$.

Sa transformée par une moyenne mobile M centrée, d'ordre 2p+1 et de coefficients $\{\theta_i\}$, est donnée par :

$$MX_t = \begin{cases} 0 & \text{si } t < -p \\ \theta_t & \text{si } -p \le t \le p \\ 0 & \text{si } t > p \end{cases}.$$

Cette transformée sera donc lisse si la courbe des coefficients de la moyenne mobile est lisse. Comme on peut le constater sur la figure précédente, les courbes des coefficients des moyennes mobiles simples composées ne sont pas « lisses » et présentent des ruptures marquées.

Henderson (1916, 1924) a proposé d'utiliser la quantité $H = \sum \left(\nabla^3 \theta_i\right)^2$, où ∇ représente l'opérateur différence première [17], pour mesurer la « souplesse » de la courbe des coefficients ». Cette quantité est nulle lorsque les coefficients $\{\theta_i\}$ se trouvent sur une parabole et, dans le cas général, elle mesure l'écart entre la forme parabolique et la forme de la fonction donnant les $\{\theta_i\}$.
Henderson a alors cherché des moyennes centrées, d'ordre 2p+1, qui conservent les polynômes d'ordre 2 et minimisent la quantité *H*.
Avec les notations du paragraphe 3.1.6, la moyenne mobile de Henderson d'ordre 2p+1 sera solution du programme de minimisation :

$$\begin{cases} \underset{\theta}{Min} \sum \left(\nabla^3 \theta_i\right)^2 \\ \sum_{i=-p}^{i=+p} \theta_i = 1, \quad \sum_{i=-p}^{i=+p} i\theta_i = 0 \quad \text{et} \quad \sum_{i=-p}^{i=+p} i^2\theta_i = 0 \end{cases}$$

Les coefficients de ces moyennes mobiles peuvent se calculer aussi explicitement et on a, pour une moyenne d'ordre 2p+1, en posant $n = p + 2$:

$$\theta_i = \frac{315\left[(n-1)^2 - i^2\right]\left[n^2 - i^2\right]\left[(n+1)^2 - i^2\right]\left[3n^2 - 16 - 11i^2\right]}{8n(n^2-1)(4n^2-1)(4n^2-9)(4n^2-25)}$$

Cette formule permet donc de calculer, sous forme rationnelle, les coefficients des moyennes mobiles de Henderson utilisées dans X11. On a donc, en ne présentant, pour cause de symétrie, que les coefficients nécessaires :

5 termes [5] ; $\frac{1}{286}\{-21,84,160\}$

7 termes [7] ; $\frac{1}{715}\{-42,42,210,295\}$

9 termes [9] ; $\frac{1}{2431}\{-99,-24,288,648,805\}$

13 termes [13] ; $\frac{1}{16796}\{-325,-468,0,1100,2475,3600,4032\}$

Et, pour la 23 termes

[23] ; $\frac{1}{4032015}\{-17250,-44022,-63250,-58575,-19950,54150,156978,275400,392700,491700,557700,580853\}$

Le tableau 2 donne, sous forme décimale, les coefficients des moyennes mobiles de Henderson utilisées dans X11 ainsi que les valeurs des critères, de réduction de variance et de lissage, associés.
X12-ARIMA permet d'utiliser n'importe quelle moyenne mobile de Henderson d'ordre impair.

Les courbes des coefficients, présentées dans la figure 7, sont lisses et les fonctions de gain de ces moyennes plus proches de la forme idéale « passe-bas » que celles des moyennes mobiles composées vues précédemment.

---

[17] $\nabla X_t = X_t - X_{t-1}$

Tableau 2 : Coefficients des moyennes mobiles de Henderson utilisées dans X11, pouvoir de réduction de variance et critère de Henderson.

$\sum \theta_i^2$ représente le pouvoir de réduction de variance (voir 3.1.5)

$\sum \left(\nabla^3 \theta_i\right)^2$ représente la valeur du critère de Henderson (voir 3.2.2)

i	5 termes	7 termes	9 termes	13 termes	23 termes
-11	.	.	.	.	-0.00428
-10	.	.	.	.	-0.01092
-9	.	.	.	.	-0.01569
-8	.	.	.	.	-0.01453
-7	.	.	.	.	-0.00495
-6	.	.	.	-0.01935	0.01343
-5	.	.	.	-0.02786	0.03893
-4	.	.	-0.04072	0.00000	0.06830
-3	.	-0.05874	-0.00987	0.06549	0.09740
-2	-0.07343	0.05874	0.11847	0.14736	0.12195
-1	0.29371	0.29371	0.26656	0.21434	0.13832
0	0.55944	0.41259	0.33114	0.24006	0.14406
1	0.29371	0.29371	0.26656	0.21434	0.13832
2	-0.07343	0.05874	0.11847	0.14736	0.12195
3	.	-0.05874	-0.00987	0.06549	0.09740
4	.	.	-0.04072	0.00000	0.06830
5	.	.	.	-0.02786	0.03893
6	.	.	.	-0.01935	0.01343
7	.	.	.	.	-0.00495
8	.	.	.	.	-0.01453
9	.	.	.	.	-0.01569
10	.	.	.	.	-0.01092
11	.	.	.	.	-0.00428
$\sum \theta_i^2$	0.4963	0.3566	0.2833	0.2038	0.1217
$\sum \left(\nabla^3 \theta_i\right)^2$	1.4965	0.2629	0.0675	0.0083	0.0003

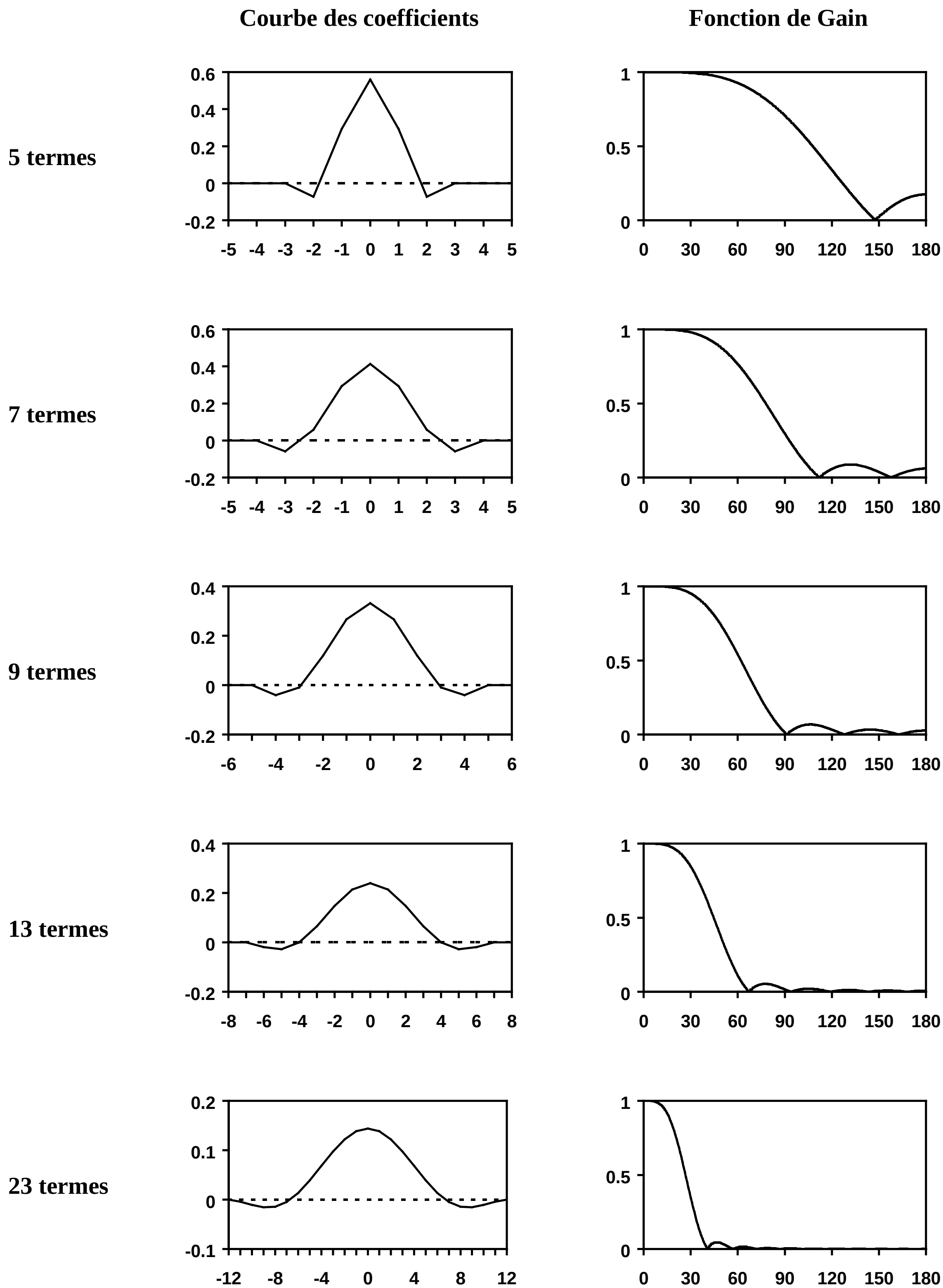
3.3 Les moyennes mobiles asymétriques de Musgrave.

L’application d’une moyenne mobile centrée d’ordre 2p+1 ne permet pas, par construction, d’avoir des estimations de la série lissée pour les p premiers et les p derniers instants de la série, ce qui est pour le moins gênant. On est donc amené dans la pratique à utiliser des moyennes mobiles non-centrées pour effectuer ces estimations.

Musgrave (1964) a étudié ce problème dans le cadre de la méthode X11 et a proposé un ensemble de moyennes asymétriques qui complètent les moyennes mobiles de Henderson.

Supposons que la série à lisser se termine en juillet 1999. Si nous utilisons une moyenne mobile symétrique d’ordre 13, le dernier point lissé grâce à cette moyenne sera celui de janvier 1999, les six derniers mois devant être estimés par des moyennes mobiles asymétriques. Six mois plus tard, donc en janvier 1999, la valeur lissée pour juillet 1999 pourra être calculé avec la moyenne mobile symétrique d’ordre 13. Nous aurons donc deux estimations différentes de cette valeur et il serait souhaitable qu’elles ne soient pas trop différentes.

Figure 7 : Moyennes mobiles de Henderson centrées utilisées dans X11



L'idée de Musgrave est justement de construire des moyennes mobiles asymétriques qui minimisent les révisions des estimations. Pour cela, il pose les hypothèses suivantes :

- la série à lisser peut se modéliser linéairement sous la forme: $X_t = a + bt + \varepsilon_t$ où *a* et *b* sont des constantes, et les ε_t sont des variables aléatoires non corrélées, de moyenne nulle et de variance σ^2.
- on dispose d'une série de poids $\{w_1, w_2, \cdots, w_N\}$ de somme égale à 1 (c'est par exemple une moyenne mobile centrée de Henderson) et on cherche une série de poids $\{v_1, v_2, \cdots, v_M\}$, avec $M < N$, de somme aussi égale à 1.
- cette nouvelle moyenne mobile doit en outre minimiser les révisions des estimations, c'est-à-dire, par exemple, minimiser le critère: $E(\sum_{i=1}^{i=M} v_i X_i - \sum_{i=1}^{i=N} w_i X_i)^2$.

Sous ces hypothèses, on montre (Doherty 1992, Findley & *alii* 1998) que les poids peuvent être calculés explicitement en fonction du rapport $D = \frac{b^2}{\sigma^2}$:

$$v_j = w_j + \frac{1}{M}\sum_{i=M+1}^{N} w_i + \frac{\left[j - \frac{M+1}{2}\right]D}{1 + \frac{M(M-1)(M+1)}{12}D}\sum_{i=M+1}^{N}\left(i - \frac{M+1}{2}\right)w_i$$

3.3.1 Moyennes mobiles asymétriques de Musgrave associées aux moyennes symétriques de Henderson

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 $R = \frac{\bar{I}}{\bar{C}}$ où $\bar{I}$ désigne la moyenne des variations absolues au mois le mois dans la partie irrégulière de la série et $\bar{C}$ désigne la moyenne des variations absolues au mois le mois dans la tendance de la série[18]. En supposant la normalité des ε_t, on montre alors que: $D = \frac{4}{\pi}\left(\frac{1}{R}\right)^2$, ce qui permet de calculer numériquement les moyennes mobiles asymétriques.

[18] Se reporter à l'explication du tableau B7 pour plus de détail

Le tableau 3 présente les coefficients de ces moyennes mobiles calculés en fonction des valeurs suivantes du rapport $R = \frac{\bar{I}}{\bar{C}}$:

Henderson sur 5 termes	$R^{-1} = 0.001$
Henderson sur 7 termes	$R^{-1} = 4.5$
Henderson sur 9 termes	$R^{-1} = 1$
Henderson sur 13 termes	$R^{-1} = 3.5$
Henderson sur 23 termes	$R^{-1} = 4.5$

Ce sont ces moyennes mobiles asymétriques qui sont utilisées dans X12-ARIMA. X11-ARIMA88 utilise des moyennes mobiles asymétriques différentes pour la moyenne de Henderson sur 5 termes.

3.3.2 Remarque sur les moyennes mobiles de Musgrave

Bien que la série à lisser soit supposée suivre un modèle linéaire en fin de période, les moyennes mobiles de Musgrave ne conservent pas les droites mais seulement les constantes. En effet, pour que cela soit le cas, il aurait en outre fallu imposer sur les coefficients la contrainte supplémentaire $\sum_{i=-p}^{+f} i\theta_i = 0$.

Comme le montre la figure 8, les moyennes asymétriques de Musgrave fournissent des estimations « prudentes » de la série lissée en atténuant la croissance observée dans les derniers points de la série.

Figure 8 : Lissage d'une droite par une moyenne mobile de Henderson sur 13 termes complétée par des moyennes mobiles asymétriques de Musgrave

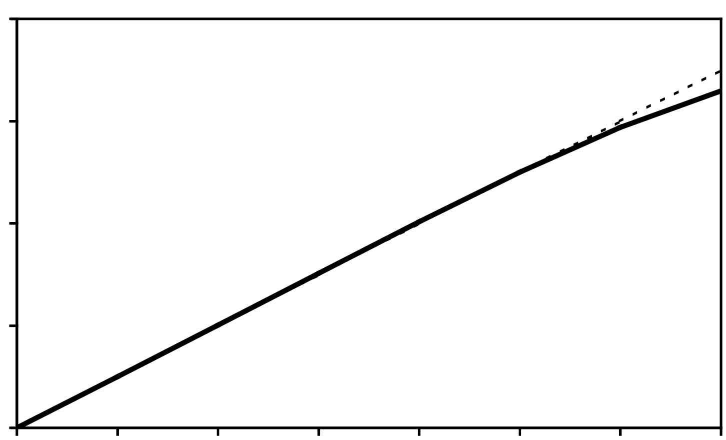


Tableau 3 : Coefficients des moyennes mobiles asymétriques de Musgrave associées aux moyennes de Henderson (X12-ARIMA)

La notation Hp_f signifie que la moyenne mobile est d'ordre *p+f+1* avec *p* points dans le passé et *f* points dans le futur.

Moyenne mobile de Henderson sur 23 termes :

i	H11_11	H11_10	H11_9	H11_8	H11_7	H11_6	H11_5	H11_4	H11_3	H11_2	H11_1	H11_0
-11	-0.00428	-0.00390	-0.00282	-0.00103	0.00108	0.00268	0.00258	-0.00065	-0.00861	-0.02293	-0.04520	-0.07689
-10	-0.01092	-0.01059	-0.00968	-0.00817	-0.00642	-0.00511	-0.00519	-0.00776	-0.01396	-0.02486	-0.04130	-0.06385
-9	-0.01569	-0.01542	-0.01467	-0.01344	-0.01205	-0.01103	-0.01109	-0.01300	-0.01744	-0.02491	-0.03554	-0.04893
-8	-0.01453	-0.01431	-0.01372	-0.01279	-0.01175	-0.01101	-0.01106	-0.01230	-0.01500	-0.01904	-0.02385	-0.02808
-7	-0.00495	-0.00479	-0.00436	-0.00372	-0.00303	-0.00258	-0.00261	-0.00319	-0.00413	-0.00475	-0.00373	0.00119
-6	0.01343	0.01354	0.01380	0.01416	0.01448	0.01465	0.01464	0.01472	0.01554	0.01834	0.02518	0.03925
-5	0.03893	0.03898	0.03908	0.03916	0.03913	0.03900	0.03902	0.03976	0.04233	0.04856	0.06121	0.08444
-4	0.06830	0.06830	0.06823	0.06802	0.06764	0.06723	0.06726	0.06866	0.07299	0.08264	0.10112	0.13350
-3	0.09740	0.09734	0.09711	0.09661	0.09587	0.09517	0.09522	0.09729	0.10337	0.11644	0.14074	0.18228
-2	0.12195	0.12184	0.12144	0.12066	0.11956	0.11858	0.11865	0.12137	0.12921	0.14571	0.17583	0.22652
-1	0.13832	0.13815	0.13759	0.13652	0.13507	0.13380	0.13389	0.13728	0.14687	0.16679	0.20273	0.26258
0	0.14406	0.14384	0.14312	0.14176	0.13995	0.13839	0.13850	0.14255	0.15390	0.17724	0.21901	0.28801
1	0.13832	0.13804	0.13716	0.13551	0.13334	0.13150	0.13163	0.13634	0.14945	0.17622	0.22380	0
2	0.12195	0.12162	0.12057	0.11864	0.11611	0.11399	0.11413	0.11951	0.13437	0.16456	0	0
3	0.09740	0.09701	0.09580	0.09358	0.09070	0.08829	0.08845	0.09449	0.11111	0	0	0
4	0.06830	0.06786	0.06649	0.06398	0.06075	0.05805	0.05823	0.06493	0	0	0	0
5	0.03893	0.03844	0.03690	0.03411	0.03052	0.02753	0.02773	0	0	0	0	0
6	0.01343	0.01288	0.01118	0.00810	0.00415	0.00088	0	0	0	0	0	0
7	-0.00495	-0.00555	-0.00742	-0.01078	-0.01509	0	0	0	0	0	0	0
8	-0.01453	-0.01519	-0.01721	-0.02087	0	0	0	0	0	0	0	0
9	-0.01569	-0.01640	-0.01859	0	0	0	0	0	0	0	0	0
10	-0.01092	-0.01169	0	0	0	0	0	0	0	0	0	0
11	-0.00428	0	0	0	0	0	0	0	0	0	0	0

Moyenne mobile de Henderson sur 13 termes :

i	H6_6	H6_5	H6_4	H6_3	H6_2	H6_1	H6_0
-6	-0.01935	-0.01643	-0.01099	-0.00813	-0.01603	-0.04271	-0.09186
-5	-0.02786	-0.02577	-0.02204	-0.02019	-0.02487	-0.03863	-0.05811
-4	0	0.00127	0.00330	0.00413	0.00267	0.00182	0.01202
-3	0.06549	0.06594	0.06626	0.06608	0.06784	0.07990	0.11977
-2	0.14736	0.14698	0.14559	0.14441	0.14939	0.17436	0.24390
-1	0.21434	0.21314	0.21004	0.20784	0.21605	0.25392	0.35315
0	0.24006	0.23803	0.23324	0.23002	0.24144	0.29223	0.42113
1	0.21434	0.21149	0.20498	0.20076	0.21540	0.27910	0
2	0.14736	0.14368	0.13547	0.13024	0.14810	0	0
3	0.06549	0.06099	0.05108	0.04483	0	0	0
4	0	-0.00532	-0.01694	0	0	0	0
5	-0.02786	-0.03401	0	0	0	0	0
6	-0.01935	0	0	0	0	0	0

Moyenne mobile de Henderson sur 9 termes :

H4_4	H4_3	H4_2	H4_1	H4_0
-0.04072	-0.03082	-0.02262	-0.04941	-0.15554
-0.00987	-0.00426	-0.00021	-0.01056	-0.03384
0.11847	0.11980	0.11969	0.12578	0.18536
0.26656	0.26361	0.25933	0.28187	0.42429
0.33114	0.32391	0.31547	0.35445	0.57972
0.26656	0.25504	0.24244	0.29786	0
0.11847	0.10267	0.08590	0	0
-0.00987	-0.02995	0	0	0
-0.04072	0	0	0	0

Moyenne mobile de Henderson sur 7 termes :

i	H3_3	H3_2	H3_1	H3_0
-3	-0.05874	-0.05314	-0.05421	-0.03379
-2	0.05874	0.05818	0.06101	0.11601
-1	0.29371	0.28699	0.29371	0.38329
0	0.41259	0.39972	0.41032	0.53449
1	0.29371	0.27468	0.28917	0
2	0.05874	0.03356	0	0
3	-0.05874	0	0	0

Moyenne mobile de Henderson sur 5 termes[19] :

X12-ARIMA			X11-ARIMA88	
H2_2	H2_1	H2_0	H2_1	H2_0
-0.07343	-0.03671	-0.18357	-0.073	-0.073
0.29371	0.29371	0.36713	0.294	0.403
0.55944	0.52273	0.81643	0.522	0.670
0.29371	0.22028	0	0.257	0
-0.07343	0	0	0	0

[19] Contrairement à X12-ARIMA, X11-ARIMA n'utilise pas des moyennes mobiles de Musgrave pour compléter la moyenne de Henderson sur 5 termes.

3.3.3 Moyennes mobiles asymétriques associées aux moyennes mobiles composées

Curieusement, bien que le premier travail de Musgrave ait porté sur la génération des filtres asymétriques associés aux moyennes mobiles composées servant à estimer les coefficients saisonniers, ses recommandations n'ont pas été appliquées dans la méthode X11.
Les filtres asymétriques associés aux moyennes 3x3, 3x5 et 3x9 figurent dans le tableau 4. Les filtres 2x4 et 2x12 ne sont pas complétés par des moyennes asymétriques.

Tableau 4 : Moyennes mobiles asymétriques associées aux moyennes mobiles composées [20].

La notation Sp_f signifie que la moyenne mobile est d'ordre *p+f+1* avec *p* points dans le passé et *f* points dans le futur.

Moyenne mobile 3x9 :

i	S5_5	S5_4	S5_3	S5_2	S5_1	S5_0
-5	1/27	35/1026	35/1026	33/1026	29/1026	52/1026
-4	2/27	75/1026	77/1026	81/1026	94/1026	115/1026
-3	3/27	114/1026	116/1026	136/1026	148/1026	177/1026
-2	3/27	116/1026	120/1026	136/1026	164/1026	202/1026
-1	3/27	117/1026	126/1026	147/1026	181/1026	227/1026
0	3/27	119/1026	131/1026	158/1026	197/1026	252/1026
1	3/27	120/1026	135/1026	167/1026	213/1026	0
2	3/27	121/1026	141/1026	177/1026	0	0
3	3/27	123/1026	145/1026	0	0	0
4	2/27	86/1026	0	0	0	0
5	1/27	0	0	0	0	0

Moyenne mobile 3x5 :

i	S3_3	S3_2	S3_1	S3_0
-3	1/15	4/60	4/60	9/60
-2	2/15	8/60	11/60	17/60
-1	3/15	13/60	15/60	17/60
0	3/15	13/60	15/60	17/60
1	3/15	13/60	15/60	0
2	2/15	9/60	0	0
3	1/15	0	0	0

Moyenne mobile 3x3 :

S2_2	S2_1	S2_0
1/9	3/27	5/27
2/9	7/27	11/27
3/9	10/27	11/27
2/9	7/27	0
1/9	0	0

[20] Dans X11-ARIMA et X12-ARIMA, les moyennes mobiles asymétriques liées à la 3x9 sont codées sous forme décimale avec 3 chiffres seulement après la virgule. La forme fractionnaire donnée ici est celle qui se rapproche le plus de cette expression décimale. Les autres moyennes mobiles sont codées sous forme fractionnaire dans les programmes.

3.4 Le filtre moyenne mobile X11

Si on ne tient pas compte de la procédure de détection et de correction des points atypiques et de la procédure d'estimation des effets de calendrier, la méthode X11 peut être vue comme l'application successive de plusieurs moyennes mobiles. L'opérateur qui permet de passer de la série brute à la série corrigée des variations saisonnières est donc lui même une moyenne mobile.

Ainsi, dans le cas d'une série mensuelle, l'algorithme de base décrit en 2.4 peut se résumer à l'application d'une seule moyenne mobile que l'on peut calculer matriciellement (Gouriéroux et Monfort, 1990).

Estimation de la Tendance-Cycle par moyenne mobile 2x12 :

$TC^{(1)} = M_{2x12}X$ avec $M_{2x12} : [13]\ ;\ \frac{1}{24}\{1,2,2,2,2,2,2\}$

Estimation de la composante Saisonnier-Irrégulier :

$$(S+I)^{(1)} = X - TC^{(1)} = [Id - M_{2x12}]X$$

Où Id représente l'opérateur identité qui transforme la série en elle-même. Ici, Id serait la moyenne mobile $[13]\ ;\ \{0,0,0,0,0,0,1\}$.

Estimation de la composante Saisonnière par moyenne mobile 3x3 sur chaque mois :

Appliquer la moyenne mobile 3x3 sur les valeurs d'un mois donné revient à appliquer sur la série des coefficients saisonniers la moyenne M_3, sur 49 mois, définie par :

$$M_3 : [49]\ ;\ \frac{1}{9}\{1,0,0,0,0,0,0,0,0,0,0,0,2,0,0,0,0,0,0,0,0,0,0,0,3\}$$

$$S^{(1)} = M_3(S+I)^{(1)} = M_3[Id - M_{2x12}]X$$

Les coefficients sont ensuite normalisés de telle sorte que leur somme sur toute période de 12 mois soit approximativement nulle :

$$Snorm^{(1)} = S^{(1)} - M_{2x12}S^{(1)} = [Id - M_{2x12}]M_3[Id - M_{2x12}]X = M_3[Id - M_{2x12}]^2 X$$

Estimation de la série corrigée des variations saisonnières :

$$Xsa^{(1)} = X - Snorm^{(1)} = X - M_3[Id - M_{2x12}]^2 X = \left(Id - M_3[Id - M_{2x12}]^2\right)X$$

Estimation de la Tendance-Cycle par moyenne mobile de Henderson sur 13 termes :

$$TC^{(2)} = H_{13}Xsa^{(1)} = H_{13}\left(Id - M_3[Id - M_{2x12}]^2\right)X$$

Estimation de la composante Saisonnier-Irrégulier :

$$(S+I)^{(2)} = X - TC^{(2)} = \left[Id - H_{13}\left(Id - M_3[Id - M_{2x12}]^2\right)\right]X$$

Estimation de la composante Saisonnière par moyenne mobile 3x5 sur chaque mois :

Appliquer la moyenne mobile 3x5 sur les valeurs d'un mois donné revient à appliquer sur la série des coefficients saisonniers la moyenne M_5, sur 73 mois, définie par :

$$M_5 : [73]\ ;\ \frac{1}{27}\{1,0,0,0,0,0,0,0,0,0,0,0,2,0,0,0,0,0,0,0,0,0,0,0,3,0,0,0,0,0,0,0,0,0,0,0,3\}$$

$$S^{(2)} = M_5(S+I)^{(2)} = M_5\left[Id - H_{13}\left(Id - M_3[Id - M_{2x12}]^2\right)\right]X$$

Les coefficients sont ensuite normalisés de telle sorte que leur somme sur toute période de 12 mois soit approximativement nulle.

$$S^{(2)} = M_5(S+I)^{(2)} = M_5\left[Id - H_{13}\left(Id - M_3[Id - M_{2x12}]^2\right)\right]X$$

$$Snorm^{(2)} = S^{(2)} - M_{2x12}S^{(2)} = (Id - M_{2x12})M_5\left[Id - H_{13}\left(Id - M_3[Id - M_{2x12}]^2\right)\right]X$$

Estimation de la série corrigée des variations saisonnières :

$$Xsa = X - Snorm^{(2)} = \left\{Id - (Id - M_{2x12})M_5\left[Id - H_{13}\left(Id - M_3[Id - M_{2x12}]^2\right)\right]\right\}X$$

L'ordre de cette moyenne mobile peut se calculer de proche en proche :

$$\text{Ordre}\,[Id - M_{2x12}]^2 = 2 * \text{Ordre}\,[Id - M_{2x12}] - 1 = 2*13 - 1 = 25$$

$$\text{Ordre}\left(Id - M_3\,[Id - M_{2x12}]^2\right) = \text{Ordre}\, M_3 + 25 - 1 = 49 + 25 - 1 = 73$$

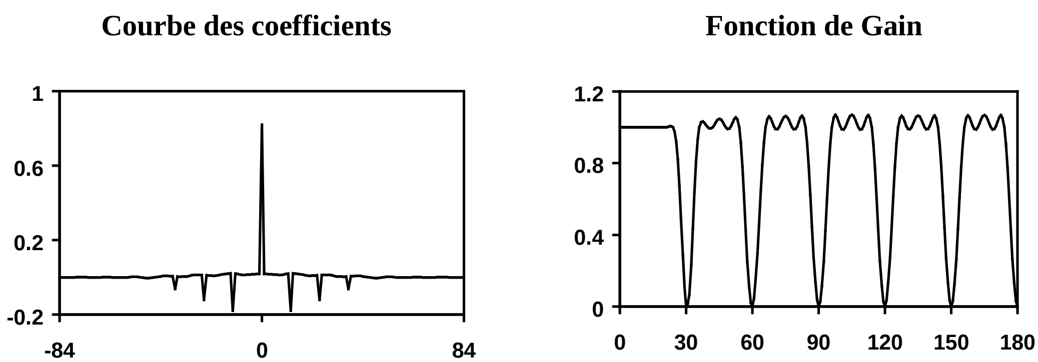
$$\text{Ordre}\left[H_{13}\left(Id - M_3\,[Id - M_{2x12}]^2\right)\right] = \text{Ordre}\ \mathrm{H}_{13} + 73 - 1 = 13 + 73 - 1 = 85$$

$$\text{Ordre}\left[M_5\left[Id - H_{13}\left(Id - M_3\,[Id - M_{2x12}]^2\right)\right]\right] = \text{Ordre}\ \mathrm{M}_5 + 85 - 1 = 73 + 85 - 1 = 157$$

$$\text{Ordre}\left[Id - (Id - M_{2x12})M_5\left[Id - H_{13}\left(Id - M_3\,[Id - M_{2x12}]^2\right)\right]\right] = \text{Ordre}\,(Id - M_{2x12}) + 157 - 1 = 169$$

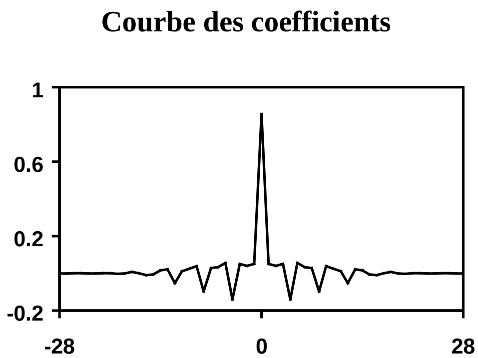
C'est donc une moyenne mobile d'ordre 169 dont les coefficients et la fonction de gain sont présentés à la figure 9. En toute rigueur, il faut disposer de 84 observations, soit 7 ans, de part et d'autre d'un point pour pouvoir utiliser ce filtre. 84 moyennes mobiles asymétriques doivent donc le compléter.

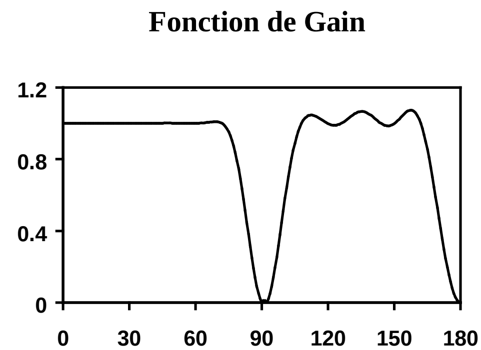
Figure 9 : Le filtre moyenne mobile mensuel symétrique de X11



La figure 10 présente les coefficients et la fonction de gain du filtre central X11 utilisé dans le cas trimestriel. Ce filtre est d'ordre 57 et nécessite aussi 7 ans de part et d'autre d'un point pour être utilisé.

Figure 10 : Le filtre moyenne mobile trimestriel symétrique de X11





Chapitre 4 LES DIFFÉRENTS TABLEAUX

Ce chapitre présente un exemple complet, et très détaillé, de désaisonnalisation par la méthode X11. La série qui sera décomposée est mensuelle : c'est dans ce cas que les options de traitement proposées par les logiciels sont les plus nombreuses et les plus complexes. Il s'agit de l'indice mensuel de la production industrielle en France dont l'évolution d'octobre 1985 à mars 1995 est représentée à la figure 11.

Avertissements

1. Dans la suite, nous nous référerons presque exclusivement au schéma de composition multiplicatif pour décrire le contenu des tableaux. Cependant, dans la traduction en formules mathématiques de ces explications, nous utiliserons un langage symbolique qui permettra de traiter les deux schémas. Ainsi, on aura les correspondances suivantes :

Symbole	Schéma Additif	Schéma Multiplicatif	Signification
_op	-	/	Ces deux premières lignes traduisent les opérations de base de chacun des deux schémas
_invop	+	*	
_xbar	0	1	Certains estimateurs, les coefficients saisonniers par exemple, sont supposés de moyenne nulle dans le cas additif et de moyenne égale à 1 dans le cas multiplicatif. Ces valeurs moyennes interviennent à de nombreux endroits dans l'algorithme de calcul de la méthode X11.
_mult	.	100 *	Dans certains tableaux, et pour le cas multiplicatif seulement, les estimations sont multipliées par 100 et s'interprètent en termes de pourcentages

2. Dans sa version actuelle, ce chapitre présente les parties B, C, D, E et F. La partie G des graphiques est avantageusement remplacée par les logiciels graphiques actuels. Les parties A (ajustements préalables de la série) et F (diagnostics de la désaisonnalisation) diffèrent sensiblement entre X11-ARIMA88 et X12-ARIMA.

3. Certains traitements relativement complexes se répètent dans la méthode X11. Ils sont détaillés au premier tableau où ils sont mis en œuvre. Il s'agit notamment :

Traitement	Tableaux
Repérage et correction des valeurs atypiques	B4, B9, B17, C17
Régression pour correction des effets de Jours Ouvrables	B14, B15, C14, C15
Extraction de la Tendance-Cycle	B7, C7, D7, D12

4. Les sorties commentées dans ce document peuvent s'obtenir en soumettant aux logiciels X11-ARIMA88 et X12-ARIMA les programmes suivants :

X12-ARIMA :

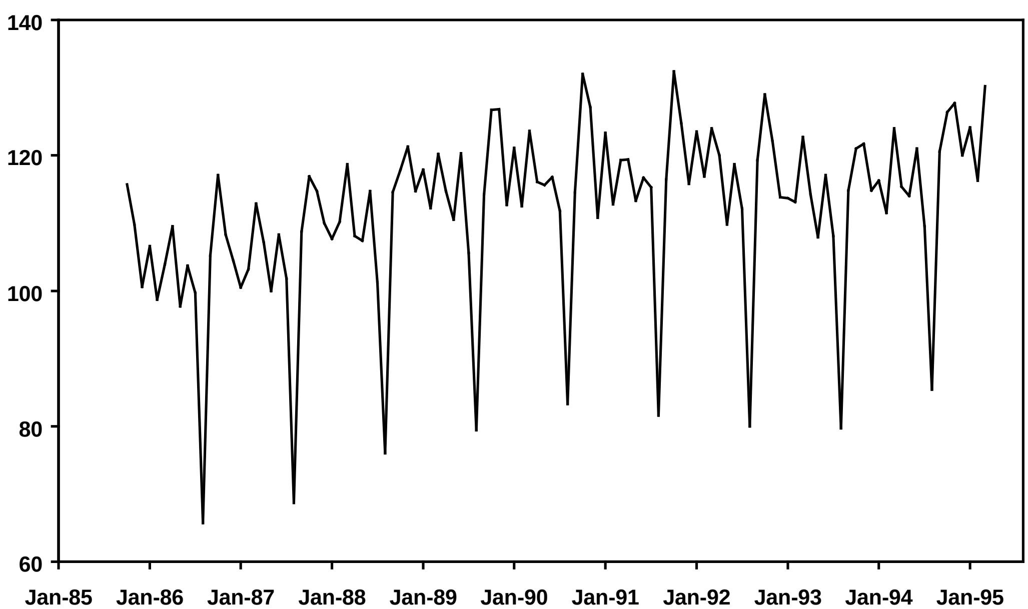
```
series{ data=(115.7 109.8 …..130.2) start= 1985.10 period= 12
        print=none decimals=3}
x11 {mode=mult print=(all) }
x11regression { variables=td print=(all)}
```

X11-ARIMA88 :

```
DATA ipi    12 85 10 ;
   (les données)
;
TITLE ipi;
RANGE 12 85  10  95  3  ;
SA (ipi, 0 ,1) TDR 2 00 00 CHART 1 PRTDEC 3 PRINT 5;
END;
```

Ces deux programmes ne donnent cependant pas tout à fait les mêmes diagnostics ; les différences sont expliquées à chaque étape concernée. Les données sont celles du tableau B1 (voir 4.1.1.3)

Figure 11 : Indice mensuel de la Production Industrielle en France, d'octobre 1985 à mars 1995



4.1 PARTIE B : Estimation préliminaire des points atypiques et des effets de calendrier

4.1.1 Tableau B1 : Série Brute ou Série Brute ajustée *a priori*

4.1.1.1 Description et mode de calcul

Ce tableau présente la série brute ou la série préalablement ajustée *a priori* par les éléments de la partie A. Ces facteurs correctifs sont : les coefficients d'ajustements permanents ou temporaires et les pondérations associées à chaque jour de la semaine.

4.1.1.2 Remarques

- Les facteurs d'ajustement *a priori* de la partie A figurent dans des tableaux différents selon le logiciel utilisé.

Facteurs d'ajustement	X11-ARIMA88	X12-ARIMA
Ajustements mensuels permanents	A2	A2p
Ajustements mensuels temporaires	A4	A2t
Effets mensuels dus aux poids journaliers fournis par l'utilisateur	A6	A4

- X11-ARIMA88 et X12-ARIMA proposent, à la suite de ce tableau B1, un test statistique sur l'existence d'une saisonnalité. Ce test ne peut être calculé qu'à partir de la première estimation de la composante Saisonnier-Irrégulier du tableau B3 (voir 4.1.3.1).

4.1.1.3 Exemple

B1 : SÉRIE BRUTE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOUT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	115.700	109.800	100.600
1986	106.600	98.700	103.900	109.500	97.700	103.700	99.700	65.700	105.200	117.100	108.300	104.400
1987	100.500	103.200	112.900	107.100	100.000	108.300	101.800	68.700	108.700	116.900	114.700	110.000
1988	107.700	110.200	118.700	108.100	107.400	114.700	101.200	76.000	114.600	117.900	121.300	114.700
1989	117.900	112.200	120.200	114.700	110.500	120.300	105.600	79.400	114.200	126.700	126.800	112.700
1990	121.100	112.500	123.600	116.100	115.600	116.800	111.800	83.300	114.600	132.000	127.100	110.800
1991	123.300	112.800	119.300	119.400	113.300	116.700	115.300	81.600	116.400	132.400	124.800	115.800
1992	123.500	116.900	124.000	120.000	109.800	118.700	112.100	80.000	119.300	129.000	122.100	113.800
1993	113.700	113.100	122.700	114.200	107.900	117.100	108.100	79.700	114.800	121.000	121.700	114.800
1994	116.300	111.500	124.000	115.400	114.000	121.000	109.500	85.400	120.600	126.400	127.700	120.000
1995	124.100	116.300	130.200	.	.	.	.	.	.	.	.	.

TEST DE LA PRÉSENCE D'UNE SAISONNALITÉ STABLE

	Somme des Carrés	D.D.L	Moyenne Quadratique	F-Value	PROB>F
Entre les Mois	10897.091	11	990.645	183.698	0.000
Résidu	485.351	90	5.393		
TOTAL	11382.442	101			

4.1.2 Tableau B2 : Estimation préliminaire de la Tendance-Cycle

4.1.2.1 Description et mode de calcul

La première estimation de la composante Tendance-Cycle est obtenue en appliquant aux données du tableau B1 une moyenne mobile centrée simple d'ordre 12. Cette moyenne mobile, qui par construction élimine les saisonnalités mensuelles constantes, est la composée d'une moyenne mobile simple sur 12 termes (et donc de coefficients 1/12) et d'une moyenne mobile simple sur 2 termes qui permet de "recentrer" le résultat. La moyenne mobile utilisée est donc une 2x12, sur 13 termes, et de coefficients {1/24, 1/12, 1/12, 1/12, 1/12, 1/12, 1/12, 1/12, 1/12, 1/12, 1/12, 1/12, 1/24}.

4.1.2.2 Remarques

- X11-ARIMA88 et X12-ARIMA proposent aussi une moyenne mobile centrée sur 24 termes (Cholette, 1981).
- Les 6 premiers et les 6 derniers points de la série, pour lesquels on ne peut obtenir d'estimation de la Tendance-Cycle à cause de la symétrie de la moyenne mobile, ne sont pas imputés à ce stade du calcul.

4.1.2.3 Exemple

B2 : TENDANCE-CYCLE - MOYENNE MOBILE SIMPLE SUR 12 TERMES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOUT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	101.458	101.454	101.550	101.454	101.388	101.950	102.225	102.221	102.508
1987	102.788	103.000	103.271	103.408	103.667	104.167	104.700	105.292	105.825	106.108	106.458	107.033
1988	107.275	107.554	108.104	108.392	108.708	109.179	109.800	110.308	110.454	110.792	111.196	111.558
1989	111.975	112.300	112.425	112.775	113.371	113.517	113.567	113.713	113.867	114.067	114.338	114.404
1990	114.517	114.938	115.117	115.354	115.588	115.521	115.533	115.638	115.471	115.429	115.471	115.371
1991	115.513	115.588	115.592	115.683	115.604	115.717	115.933	116.113	116.479	116.700	116.579	116.517
1992	116.467	116.267	116.321	116.300	116.046	115.850	115.358	114.792	114.579	114.283	113.963	113.817
1993	113.583	113.404	113.204	112.683	112.333	112.358	112.508	112.550	112.538	112.642	112.946	113.363
1994	113.583	113.879	114.358	114.825	115.300	115.767	116.308	116.833	117.292	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

Ainsi la valeur d'avril 1986, la première calculable, s'obtient à partir des valeurs du tableau B1 d'octobre 1985 à octobre 1986 (6 mois avant et 6 mois après) :

$$APR86 = \frac{115.7}{24} + \frac{109.8 + 100.6 + 106.6 + 98.7 + 103.9 + 109.5 + 97.7 + 103.7 + 99.7 + 65.7 + 105.2}{12} + \frac{117.1}{24} = 101.458$$

4.1.3 Tableau B3 : Estimation préliminaire de la composante Saisonnier-Irrégulier non modifiée

4.1.3.1 Description et mode de calcul

La composante Tendance-Cycle est enlevée de la série analysée, par soustraction ou division selon le schéma de composition adopté, pour obtenir une première estimation de la composante Saisonnier-Irrégulier.
On a donc : *B3 = B1 _op B2*

Ce tableau sert de base au *test de saisonnalité stable* édité à la suite du tableau B1. Ce test est un test d'analyse de la variance à un facteur : on dispose de k échantillons (ici les estimations Saisonnier-Irrégulier pour chacun de nos k=12 mois ou de nos k=4 trimestres) de tailles respectives $n_1, n_2, ..., n_k$. Ces échantillons correspondent chacun à un niveau différent d'un facteur A, ici la *saisonnalité*. On suppose que ce facteur influe uniquement sur les moyennes des distributions et non sur leur variance. Il s'agit donc d'un test de confusion des k moyennes $\bar{x}_1, \bar{x}_2, ..., \bar{x}_k$. Si on considère chaque échantillon comme issu d'une variable aléatoire X_i suivant une loi de moyenne m_i et d'écart-type σ, le problème est de tester :

$$\begin{cases} H_0 : m_1 = m_2 = ... = m_k \\ H_1 : m_i \neq m_j \ \text{ pour au moins un couple } (i, j) \end{cases}$$

L'équation dite *d'analyse de la variance* s'écrit :

$$\frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_j}\left(X_{ij} - \bar{X}\right)^2 = \frac{1}{n}\sum_{i=1}^{k} n_i\left(\bar{X}_i - \bar{X}\right)^2 + \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_j}\left(X_{ij} - \bar{X}_i\right)^2 \quad \text{soit} \quad \frac{1}{n}S^2 = \frac{S_A^2 + S_R^2}{n}$$

La variance totale se décompose donc en variance des moyennes, variance due au facteur saisonnalité, et en une variance résiduelle. Si l'hypothèse H_0 est vraie, on montre que la quantité $F_S = \frac{S_A^2 / (k-1)}{S_R^2 / (n-k)}$ suit une loi de Fisher-Snedecor $F(k-1; n-k)$ à *k-1* et *n-k* degrés de liberté.

D'où le test : si la quantité F_S calculée à partir des données du tableau B3 est supérieure à la valeur critique d'une loi de Fisher-Snedecor, on conclut à une influence significative du facteur saisonnalité (i.e. les moyennes mensuelles ou trimestrielles ne sont pas toutes égales).

4.1.3.2 Remarques

- Là encore, pas d'estimation pour les 6 valeurs de début et les 6 valeurs de fin de la série.
- Certaines hypothèses du modèle d'analyse de la variance utilisé pour tester l'existence d'une saisonnalité sont sans doute violées. Ainsi la composante irrégulière peut être autocorrélée à ce niveau de l'analyse. Il est donc usuel de choisir des valeurs critiques élevées et de ne rejeter l'hypothèse nulle que si la probabilité associée à la valeur de la statistique F_S est inférieure à un pour mille.
- Le rejet de l'hypothèse nulle tend à prouver l'existence d'une saisonnalité stable, c'est à dire que par exemple, le coefficient saisonnier du mois de janvier n'évolue pas. Accepter l'hypothèse nulle ne signifie pour autant pas qu'il n'y a pas de saisonnalité. Ainsi, une forte

évolution des coefficients saisonniers se traduirait par une variance telle qu’il serait impossible de distinguer statistiquement les moyennes.

4.1.3.3 Exemple

B3 : SÉRIE BRUTE DÉBARRASSÉE DE SA TENDANCE (estimation Saisonnier-Irrégulier)

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	107.926	96.300	102.117	98.271	64.801	103.188	114.551	105.947	101.845
1987	97.775	100.194	109.324	103.570	96.463	103.968	97.230	65.247	102.717	110.170	107.742	102.772
1988	100.396	102.460	109.802	99.731	98.796	105.057	92.168	68.898	103.753	106.416	109.087	102.816
1989	105.291	99.911	106.916	101.707	97.468	105.976	92.985	69.825	100.293	111.075	110.900	98.510
1990	105.749	97.879	107.369	100.647	100.011	101.107	96.769	72.035	99.246	114.356	110.071	96.038
1991	106.742	97.588	103.208	103.213	98.007	100.850	99.454	70.277	99.932	113.453	107.052	99.385
1992	106.039	100.545	106.602	103.181	94.618	102.460	97.175	69.691	104.120	112.877	107.141	99.985
1993	100.103	99.732	108.388	101.346	96.053	104.220	96.082	70.813	102.010	107.420	107.751	101.268
1994	102.392	97.911	108.431	100.501	98.873	104.521	94.146	73.096	102.821	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

La valeur de avril86 s’obtient donc simplement : $APR86 = 100*(109.500/101.458) = 107.926$

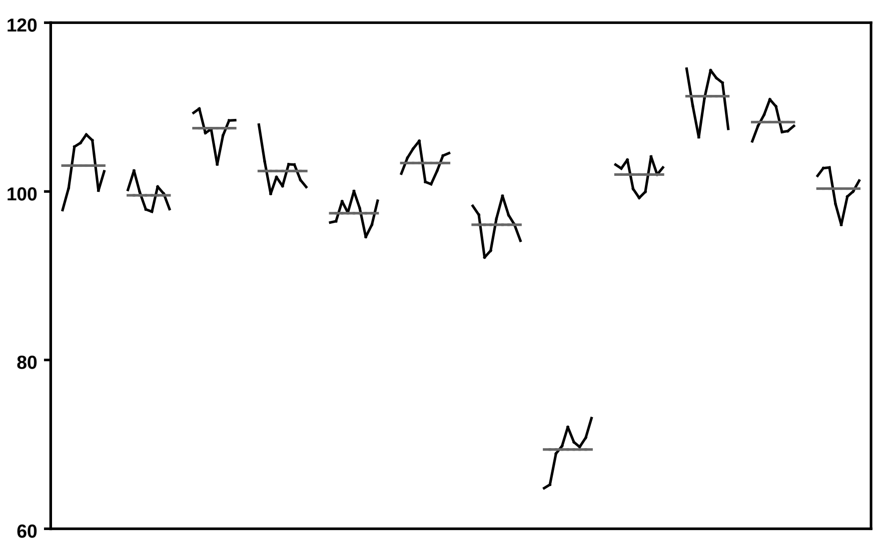
TEST DE LA PRÉSENCE D'UNE SAISONNALITÉ STABLE

	Somme des Carrés	D.D.L	Moyenne Quadratique	F-Value	PROB>F
Entre les Mois	10897.091	11	990.645	183.698	0.000
Résidu	485.351	90	5.393		
TOTAL	11382.442	101			

La valeur de la statistique F_S étant très élevée, l’hypothèse d’égalité de la moyenne des coefficients saisonniers est rejetée. Comme le graphique ci dessous le montre, cela est essentiellement dû au mois d’août qui présente un coefficient saisonnier très faible.

Figure 12 : Composantes Saisonnier-Irrégulier de chaque mois.

(Ce graphique représente les 12 colonnes du tableau B3 : la première courbe montre de la composante Saisonnier-Irrégulier du mois de janvier pour les années 1987 à 1994 ; la droite représente la valeur moyenne de ces 8 valeurs ; les autres courbes correspondent aux autres mois).



4.1.4 Tableau B4 : Valeurs de remplacement pour les points atypiques de la composante Saisonnier-Irrégulier

4.1.4.1 Description et mode de calcul

Ce tableau présente les résultats de la procédure automatique de détection et de correction des points "atypiques" de la composante Saisonnier-Irrégulier présente dans la méthode X11. Y figurent les valeurs de remplacement proposées pour les points aberrants détectés ainsi que, en marge du tableau, le écart-types mobiles ayant servi à les déterminer.
C'est sans aucun doute le tableau le plus difficile à reconstituer dans la mesure où il résulte d'un algorithme assez complexe comportant six étapes.

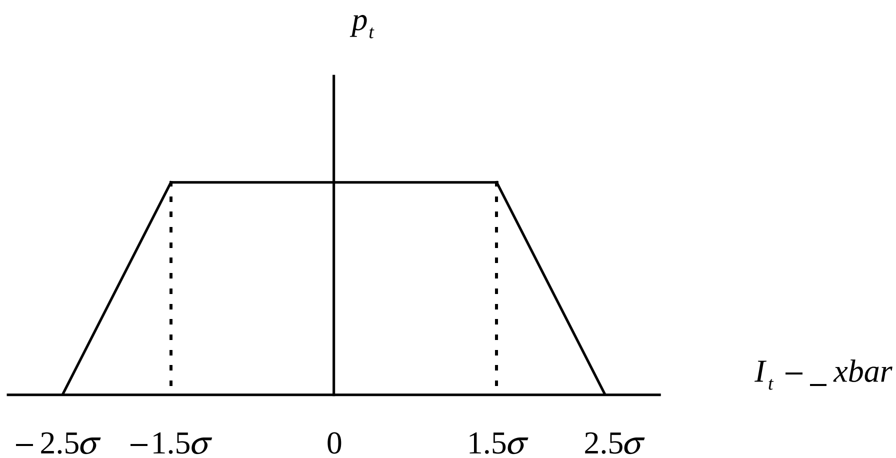
- ***Étape 1 : estimation de la composante saisonnière***.
  Cette composante saisonnière est estimée par lissage de la composante Saisonnier-Irrégulier mais en s'intéressant aux valeurs de chaque mois : on lisse les valeurs correspondant au mois de janvier, puis celles du mois de février et ainsi de suite, en employant une moyenne 3x3, de coefficients {1/9,2/9,3/9,2/9,1/9}. Cette moyenne mobile, symétrique, porte sur 5 termes et ne permet donc pas d'estimer les valeurs des coefficients saisonniers des 2 premières et dernières années. Celles-ci sont alors calculées grâce à des moyennes mobiles asymétriques *ad hoc*. On obtient alors une série de coefficients provisoires *fspro*.

- ***Étape 2 : normalisation des coefficients saisonniers***.
  Ces coefficients saisonniers provisoires sont alors normalisés de telle sorte que, sur une année d'observations, leur moyenne soit approximativement égale à 0 (pour un schéma additif) ou à 1 (pour un schéma multiplicatif). Pour cela, on calcule une moyenne mobile centrée sur 12 termes, donc une 2x12, de la série *fspro*, notée $M_{2x12}(fspro)$; les six valeurs qui manquent au début (à la fin) de cette série sont prises égales à la première (dernière) valeur calculée par cette moyenne mobile. La série *fsnorm* des coefficients normalisés est alors définie par :
  fsnorm = fspro _op $M_{2x12}(fspro)$

- ***Étape 3 : estimation de la composante irrégulière.***
  L'estimation initiale des facteurs saisonniers est enlevée de la composante Saisonnier-Irrégulier pour obtenir une estimation de la composante irrégulière.
  Irreg = SI _op *fsnorm*

- ***Étape 4 : Calcul d'un écart-type mobile.***
  Un écart-type mobile de la composante irrégulière est calculé sur des intervalles de cinq ans. Chaque écart-type est associé à l'année centrale qui a permis de le calculer. Les valeurs de cette année centrale qui s'éloignent, en valeur absolue de leur écart à la moyenne *_xbar*, de plus de 2.5 fois cet écart-type sont considérées comme aberrantes et affectées d'un poids nul. L'écart-type mobile est alors recalculé sans tenir compte de ces valeurs (cela a donc pour effet d'obtenir une estimation plus robuste de cet écart-type).
  Pour les deux premières années, c'est l'écart-type associé à la troisième année qui est utilisé dans les comparaisons. De même, pour les deux dernières années, c'est l'écart-type associé à l'antépénultième année qui est considéré.

- ***Étape 5 : Détection des valeurs atypiques et pondération de l'irrégulier.***
  On affecte à chaque valeur de la composante irrégulière un poids, fonction de l'écart-type σ qui lui est associé, calculé de la façon suivante (voir graphique) :
  - Les valeurs qui sont, en valeur absolue de leur écart à la moyenne *_xbar*, plus grandes que 2.5σ sont affectées d'un poids nul.
  - Les valeurs qui sont, en valeur absolue de leur écart à la moyenne *_xbar*, plus petites que 1.5σ sont affectées d'un poids égal à 1.
  - Celles qui sont, en valeur absolue de leur écart à la moyenne *_xbar*, comprises entre 1.5σ et 2.5σ reçoivent un poids variant linéairement entre 0 et 1, en fonction de leur position.



- ***Étape 6 : Correction des valeurs atypiques de la composante Saisonnier-Irrégulier.***

Une valeur de la composante Saisonnier-Irrégulier dont l'irrégulier ne reçoit pas une pondération intégrale, et qui a donc été considérée comme atypique, est corrigée et remplacée par une moyenne pondérée de cinq valeurs :

- la valeur elle-même affectée de son poids,
- les deux valeurs précédentes, pour le même mois, ayant une pondération intégrale,
- et les deux valeurs suivantes, pour le même mois, ayant une pondération intégrale.

Pour les deux premières et deux dernières années, les valeurs de remplacement sont calculées comme la moyenne pondérée de la valeur considérée et des quatre valeurs les plus proches, du même mois, ayant reçu une pondération 1.[21]

4.1.4.2 Remarques

Sur les moyennes mobiles asymétriques :

- Les moyennes mobiles asymétriques utilisées pour compléter la moyenne mobile 3x3 diffèrent très légèrement, pour des problèmes d'arrondis, selon le logiciel utilisé.

[21] En début et en fin de série, c'est-à-dire dans les deux premières et deux dernières années, X11 utilise les 3 plus proches valeurs de pondération intégrale pour corriger un point atypique (et non les 4 plus proches comme dans X11-ARIMA88 ou X12-ARIMA). De plus, il y a ici des erreurs de calcul dans le logiciel X11.

Sur le calcul des poids et la correction des valeurs atypiques :

- Le caractère atypique d'une valeur I_t de l'irrégulier est déterminé en comparant la valeur de $|I_t - \_xbar|$ aux bornes $\lambda_{Inf} * \sigma_t$ et $\lambda^{Sup} * \sigma_t$ où σ_t est l'écart-type de l'année de l'observation I_t et λ_{Inf} et λ^{Sup} sont des paramètres fixés par l'utilisateur (par défaut, 1.5 et 2.5)
- S’il n’existe pas quatre points du même mois ayant reçu une pondération intégrale, la valeur atypique est remplacée par la moyenne des valeurs du mois.

Sur le calcul des écart-types mobiles :

- L'écart-type est calculé en supposant connue la moyenne de l'irrégulier *_xbar*, c'est à dire 0 pour un schéma additif et 1 pour un schéma multiplicatif. Dans ce cas, un estimateur sans biais de la variance est donné par :
  $\sigma^2 = \frac{1}{n}\sum_i (x_i - \_xbar)^2$, où n est le nombre d’observations utilisées (voir ci-après).
  N'oublions pas que nous cherchons à repérer des valeurs atypiques. En utilisant une estimation de la moyenne, on risquerait de voir cette estimation fortement influencée par ces mêmes valeurs atypiques.
- Lors du calcul du second écart-type correspondant à une année, on exclut les valeurs x de l'irrégulier de l'année vérifiant la condition $|x - \_xbar| > \lambda^{Sup} * \sigma$.
- L'écart-type est en principe calculé sur 5 années complètes d'observations. En début et en fin de série, en particulier à cause des valeurs manquantes générées par l'utilisation de moyennes mobiles symétriques, on fait une petite exception.
  Ainsi, si votre série brute commence en janvier 1970, la première estimation de la composante Saisonnier-Irrégulier du tableau B3 commencera en juillet 1970. Pour X11-ARIMA88 et X12-ARIMA, l’écart-type de 1972 et sera calculé à partir des observations de 1970 et des cinq premières années complètes (de 1971 à 1975) donc sur 66 observations. C'est cet écart-type qui sera alors attribué aux années 1970 et 1971. L'écart-type de 1973 sera lui calculé sur moins d'observations, les 60 correspondants aux années 1971 à 1975.[22]

4.1.4.3 Exemple

Voici le tableau en sortie de X11. Nous allons, pour mieux le comprendre, détailler les étapes de son calcul à partir de tableaux malheureusement non récupérables dans les versions usuelles des logiciels de la famille X11 (tableaux numérotés ici de B4a à B4f).

[22] Notons que pour le logiciel X11, le premier écart-type sur cinq ans calculable sera celui de 1973 et sa valeur sera aussi affectée aux années 1970, 1971 et 1972. Les six premières observations ne seront donc pas intégrées dans ce calcul

B4 : VALEURS DE REMPLACEMENT POUR LES POINTS EXTRÊMES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC	ÉCART -TYPE
1985	.	.	.	.	.	.	.	.	.	.	.	.	1.427
1986	.	.	.	102.584	.	.	.	.	.	112.451	.	.	1.427
1987	103.375	.	.	.	.	.	.	.	.	.	.	.	1.427
1988	.	.	.	101.798	.	.	95.684	.	.	112.038	.	.	1.427
1989	.	.	.	.	.	103.387	.	.	.	.	.	.	1.371
1990	.	.	.	.	.	.	.	70.119	.	.	.	99.580	1.396
1991	.	.	106.783	.	.	.	96.339	.	.	.	.	.	1.294
1992	.	.	.	.	97.354	.	.	.	101.594	.	.	.	1.285
1993	104.841	.	.	.	.	.	.	.	.	112.788	.	.	1.285
1994	.	.	.	.	98.075	.	.	70.649	.	.	.	.	1.285
1995	.	.	.	.	.	.	.	.	.	.	.	.	1.285

- ***Étape 1 : estimation de la composante saisonnière***.

B4a : FACTEURS SAISONNIERS PROVISOIRES (MM 3x3)

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	104.634	96.829	103.416	96.717	65.741	103.101	111.260	107.260	102.403
1987	100.235	101.065	109.073	103.497	97.137	103.993	95.716	66.587	102.838	110.433	107.976	102.070
1988	101.580	100.703	108.659	102.035	97.840	104.254	94.664	68.186	102.190	110.072	108.951	100.988
1989	103.631	99.799	107.513	101.407	98.276	103.897	94.835	69.652	101.058	110.933	109.534	99.488
1990	105.305	99.071	106.529	101.632	98.266	102.721	96.058	70.544	100.673	112.380	109.149	98.523
1991	105.466	98.806	105.874	102.261	97.422	102.209	97.257	70.547	100.981	112.591	108.248	98.887
1992	104.438	99.118	106.533	102.201	96.762	102.572	97.057	70.709	102.035	111.776	107.601	99.724
1993	102.973	99.232	107.361	101.810	96.629	103.467	96.238	71.054	102.537	110.761	107.373	100.397
1994	102.135	99.140	108.075	101.342	96.936	104.017	95.496	71.535	102.731	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

Les données du tableau B3 sont lissées colonne par colonne (mois par mois), avec une moyenne mobile 3x3 de coefficients {1/9,2/9,3/9,2/9,1/9}. Ainsi, les valeurs de la composante Saisonnier-Irrégulier du mois d'avril sont :

	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
Avril	.	107.926	103.570	99.731	101.707	100.647	103.213	103.181	101.346	100.501	

Le facteur saisonnier du mois d'avril 1988 sera donc estimé par :

$$APR88 = \frac{107.926 + 2*103.570 + 3*99.731 + 2*101.707 + 100.647}{9} = 102.035$$

Cette moyenne mobile symétrique peut s'appliquer pour estimer les valeurs des coefficients saisonniers des années 1988 à 1992. Pour le début de la série (années 1986 et 1987) et la fin de la série (années 1993 et 1994), on utilise des moyennes asymétriques prédéfinies (voir tableau 4, 3.3.3) :

- $APR86 = \frac{107.926*11 + 103.570*11 + 99.731*5}{27} = 104.634$

  (le point courant et deux points dans le futur)
- $APR87 = \frac{107.926*7 + 103.570*10 + 99.731*7 + 101.707*3}{27} = 103.497$

  (un point dans le passé, le point courant et deux points dans le futur)
- $APR93 = \frac{100.501*7 + 101.346*10 + 103.181*7 + 103.213*3}{27} = 101.810$

  (un point dans le futur, le point courant et deux points dans le passé)
- $APR94 = \frac{100.501*11 + 101.346*11 + 103.181*5}{27} = 101.342$

  (le point courant et deux points dans le passé)

- ***Étape 2 : normalisation des coefficients saisonniers***.

B4b: MOYENNE MOBILE CENTRÉE SUR 12 MOIS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	100.097	100.097	100.097	100.097	100.097	100.097	100.097	100.062	100.099
1987	100.082	100.075	100.099	100.054	100.050	100.066	100.108	100.149	100.116	100.038	100.007	100.047
1988	100.014	100.037	100.076	100.034	100.060	100.055	100.096	100.144	100.058	99.984	99.976	99.979
1989	99.972	100.040	100.054	100.043	100.103	100.064	100.072	100.111	100.040	100.008	100.017	99.968
1990	99.970	100.058	100.079	100.123	100.168	100.111	100.078	100.073	100.035	100.034	100.025	99.969
1991	99.997	100.047	100.060	100.082	100.053	100.031	100.003	99.973	100.014	100.039	100.009	99.996
1992	100.003	100.001	100.052	100.062	100.001	100.009	99.983	99.927	99.966	99.984	99.962	99.994
1993	99.997	99.977	100.013	99.991	99.939	99.958	99.951	99.912	99.938	99.948	99.942	99.977
1994	99.969	99.958	99.987	99.987	99.987	99.987	99.987	99.987	99.987	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

Au tableau B4a, on applique une moyenne mobile centrée sur 12 mois. Le premier terme calculable est donc celui d'octobre 1986 et le dernier, celui de mars 1994. Ainsi :

$$OCT86 = \frac{104.634}{24} + \frac{96.829+103.416+96.717+65.741+103.101+111.260+107.260+102.403+100.235+101.065+109.073}{12} + \frac{103.497}{24} = 100.097$$

Les six premières valeurs, d'avril à septembre 1986, non calculables avec cette moyenne mobile symétrique, seront prises égales à la première valeur calculable, celle d'octobre 1986. On procède de même pour la fin de série : la valeur calculée pour mars 1994 (99.987) est répétée pour les six mois suivants. Les coefficients saisonniers normalisés sont alors obtenus en corrigeant le tableau B4a par le tableau B4b, pour obtenir :

B4c : FACTEURS SAISONNIERS NORMALISÉS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	104.532	96.735	103.315	96.623	65.678	103.001	111.152	107.193	102.301
1987	100.153	100.989	108.964	103.441	97.089	103.925	95.613	66.488	102.719	110.391	107.969	102.022
1988	101.566	100.666	108.577	102.001	97.782	104.196	94.573	68.088	102.131	110.089	108.977	101.009
1989	103.660	99.759	107.456	101.363	98.176	103.830	94.767	69.574	101.018	110.924	109.515	99.521
1990	105.337	99.014	106.445	101.507	98.102	102.607	95.984	70.493	100.638	112.342	109.122	98.554
1991	105.469	98.760	105.811	102.177	97.370	102.178	97.254	70.566	100.967	112.548	108.238	98.890
1992	104.435	99.117	106.478	102.138	96.761	102.563	97.074	70.761	102.070	111.794	107.642	99.730
1993	102.976	99.255	107.347	101.819	96.688	103.511	96.285	71.117	102.600	110.818	107.435	100.419
1994	102.166	99.182	108.089	101.355	96.949	104.031	95.509	71.545	102.745	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

Ainsi, $APR86 = 100*(104.634 / 100.097) = 104.532$

- ***Étape 3 : estimation de la composante irrégulière.***

Il suffit de diviser la composante Saisonnier-Irrégulier du tableau B3 par les facteurs saisonniers normalisés du tableau B4c ci-dessus.

B4d : IRRÉGULIER PROVISOIRE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	103.246	99.550	98.840	101.705	98.665	100.182	103.058	98.838	99.555
1987	97.625	99.213	100.330	100.125	99.355	100.041	101.692	98.134	99.998	99.800	99.790	100.735
1988	98.848	101.782	101.128	97.775	101.038	100.826	97.456	101.189	101.589	96.663	100.101	101.789
1989	101.573	100.152	99.498	100.339	99.279	102.066	98.120	100.360	99.282	100.137	101.264	98.985
1990	100.391	98.854	100.869	99.152	101.946	98.539	100.818	102.189	98.617	101.793	100.870	97.447
1991	101.207	98.814	97.540	101.014	100.654	98.701	102.261	99.590	98.974	100.804	98.904	100.500
1992	101.535	101.441	100.116	101.022	97.785	99.900	100.105	98.488	102.009	100.969	99.534	100.256
1993	97.210	100.481	100.970	99.535	99.344	100.685	99.789	99.573	99.425	96.934	100.294	100.845
1994	100.221	98.719	100.316	99.157	101.984	100.471	98.574	102.168	100.074	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

Ainsi, $APR86 = 100*(107.926/104.532) = 103.246$

- ***Étape 4 : Calcul d'un écart-type mobile.***

L'écart-type correspondant à l'année 1989 sera calculé à partir des données des années 1987 à 1991 (deux années avant, deux années après) selon la formule [23]:

$$\sigma_{89} = \sqrt{\frac{1}{60}\sum_{i=1}^{60}(x_i - 100)^2} = 1.3705$$

Ceux des années 1990 et 1991 sont calculés selon le même principe.
Pour X11-ARIMA88 et X12-ARIMA, l'écart-type de 1988 est calculé à partir de l'ensemble des observations disponibles de 1986 à 1991, soit 69 observations. C'est cet écart-type qui sera associé aux années 1986 et 87. De même, l'écart-type de 1992 utilise toutes les données de 1989 à 1994 ; il sera aussi associé aux années 1993 et 1994.
Les résultats du calcul de X11-ARIMA88 et X12-ARIMA figurent dans le tableau B4e, colonne Écart-type 1.

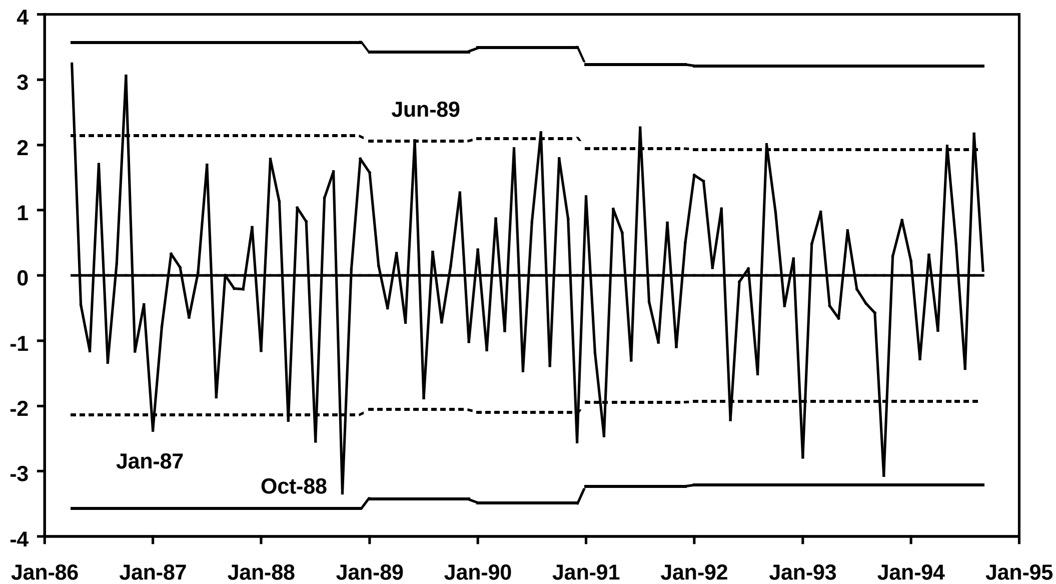
B4e: VALEURS DES ÉCART-TYPES MOBILES SUR CINQ ANS

ANNÉE	ÉCART-TYPE 1	ÉCART-TYPE 2
1985		
1986	1.4265	1.4265
1987	1.4265	1.4265
1988	1.4265	1.4265
1989	1.3705	1.3705
1990	1.3958	1.3958
1991	1.2941	1.2941
1992	1.2847	1.2847
1993	1.2847	1.2847
1994	1.2847	1.2847
1995		

Ce premier calcul sert à repérer d'éventuels points atypiques. Pour une année donnée, une valeur sera considérée comme aberrante si elle s'éloigne, en valeur absolue de son écart à la moyenne théorique (ici 1), de plus de 2.5 fois l'écart-type correspondant à cette année. Le graphique suivant représente l'écart de l'irrégulier à sa moyenne théorique et les deux "limites de confiance". Comme on peut le constater, aucune valeur n'est considérée comme très atypique. Dans le cas contraire, on aurait éliminé ces valeurs et refait le calcul de l'écart-type (voir par exemple le cas du tableau B17, 4.1.18.3). Ce nouveau calcul conduit donc à la colonne Écart-type 2 du tableau B4e identique à la colonne 1.

[23] où la moyenne théorique est considérée égale à 100 pour tenir compte du fait que les valeurs de l'irrégulier ont été elles-mêmes multipliées par 100

Figure 13 : Écart à la moyenne de l'irrégulier et "limites de confiance" (tableau B4)



Étape 5 : Détection des valeurs atypiques et pondération de l'irrégulier.

Le graphique ci-dessus permet de situer les valeurs de l'irrégulier par rapport aux limites de confiance supérieures et inférieures calculées à partir des écart-types estimés auparavant. Toutes les valeurs situées entre ces limites sont considérées comme atypiques et vont donc être corrigées ; les poids associés à chacune de ces valeurs figurent dans le tableau B4f.

B4f : POIDS ASSOCIÉS AUX VALEURS DE L'IRRÉGULIER

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	22.419	100.000	100.000	100.000	100.000	100.000	35.633	100.000	100.000
1987	83.535	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	100.000	94.011	100.000	100.000	71.692	100.000	100.000	16.096	100.000	100.000
1989	100.000	100.000	100.000	100.000	100.000	99.217	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	100.000	100.000	100.000	100.000	100.000	93.192	100.000	100.000	100.000	67.113
1991	100.000	100.000	59.928	100.000	100.000	100.000	75.258	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	100.000	100.000	77.601	100.000	100.000	100.000	93.624	100.000	100.000	100.000
1993	32.820	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	11.350	100.000	100.000
1994	100.000	100.000	100.000	100.000	95.570	100.000	100.000	81.282	100.000	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

Aucune valeur n'est jugée très atypique. Par contre, on a par exemple pour janvier 1987 :

$$\left|JAN87 - 100\right| = \left|97.625 - 100\right| \;= 2.375$$

$$1.5 * \sigma_{87} = 1.5 * 1.4265 = 2.13975 < \;\; 2.375 < \;\; 2.5 * \sigma_{87} = 2.5 * 1.4265 = 3.5663$$

On va attribuer à cette valeur jugée moyennement atypique, un poids proportionnel à l'écart à la moyenne constaté de :

$$poids(JAN87) = \frac{3.5663 - 2.375}{3.5663 - 2.1398} = 0.835$$

De même, pour la valeur de juin 1989, légèrement au-delà de la limite de confiance inférieure, on a :

$$\left|JUN89-100\right| = \left|102.066-100\right| = 2.066$$

$$1.5*\sigma_{89} = 1.5*1.3705 = 2.056 < \ 2.066 \ < \ 2.5*\sigma_{89} = 2.5*1.3705 = 3.426$$

$$poids(JUN89) = \frac{3.426-2.066}{3.426-2.056} = 0.992$$

- ***Étape 6 : Correction des valeurs atypiques de la composante Saisonnier-Irrégulier.***

La correction de la composante Saisonnier-Irrégulier (tableau B3) se fait enfin à partir de ces poids. Ainsi, la valeur de juin 1989 sera remplacée par la moyenne de cette valeur affectée de son poids et des deux valeurs antérieures et postérieures du même mois ayant reçu une pondération intégrale, donc non jugées atypiques. Il s'agit donc, comme le montre le tableau B4f, des valeurs des mois de juin 1987, 1988 et 1990, 1991. Ce qui conduit à :

$$SI(JUN89) = \frac{103.968+105.057+105.976*0.992+101.107+100.850}{4+0.992} = 103.387$$

La valeur de janvier 1987 a été jugée elle aussi atypique. Mais, comme elle se trouve en début de série, elle est corrigée de façon légèrement différente et est remplacée par la moyenne de cette valeur affectée de son poids et des quatre valeurs les plus proches du même mois ayant reçu une pondération intégrale. Il s'agit donc dans ce cas, d'après le tableau B4f, des valeurs de janvier 1988, 1989, 1990 et 1991. On a donc, aux erreurs d'arrondis près :

$$SI(JAN87) = \frac{97.775*0.835+100.396+105.291+105.749+106.742}{4+0.835} = 103.375$$

4.1.5 Tableau B5 : Estimation de la composante saisonnière

4.1.5.1 Description et mode de calcul

Cette estimation est obtenue à partir des valeurs de la composante Saisonnier-Irrégulier du tableau B3 corrigées par les valeurs du tableau B4. On procède en deux étapes identiques aux 2 premières étapes du tableau B4 :

- ***Étape 1 : estimation de la composante saisonnière par moyenne mobile 3x3***.
- ***Étape 2 : normalisation des coefficients saisonniers***.

4.1.5.2 Exemple

L'estimation est faite à partir de la composante Saisonnier-Irrégulier corrigée :

B4g : COMPOSANTE SAISONNIER-IRRÉGULIER CORRIGÉE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	102.584	96.300	102.117	98.271	64.801	103.188	112.451	105.947	101.845
1987	103.375	100.194	109.324	103.570	96.463	103.968	97.230	65.247	102.717	110.170	107.742	102.772
1988	100.396	102.460	109.802	101.798	98.796	105.057	95.684	68.898	103.753	112.038	109.087	102.816
1989	105.291	99.911	106.916	101.707	97.468	103.387	92.985	69.825	100.293	111.075	110.900	98.510
1990	105.749	97.879	107.369	100.647	100.011	101.107	96.769	70.119	99.246	114.356	110.071	99.580
1991	106.742	97.588	106.783	103.213	98.007	100.850	96.339	70.277	99.932	113.453	107.052	99.385
1992	106.039	100.545	106.602	103.181	97.354	102.460	97.175	69.691	101.594	112.877	107.141	99.985
1993	104.841	99.732	108.388	101.346	96.053	104.220	96.082	70.813	102.010	112.788	107.751	101.268
1994	102.392	97.911	108.431	100.501	98.075	104.521	94.146	70.649	102.821	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

Étape 1 : estimation de la composante saisonnière.

Les données du tableau précédent sont lissées colonne par colonne (mois par mois), avec une moyenne mobile 3x3 coefficients {1/9,2/9,3/9,2/9,1/9}, pour aboutir au tableau B5a.

Le facteur saisonnier du mois d'avril 1988 sera donc estimé par :

$$APR88 = \frac{102.584 + 2*103.570 + 3*101.798 + 2*101.707 + 100.647}{9} = 102.131$$

Cette moyenne mobile symétrique peut s'appliquer pour estimer les valeurs des coefficients saisonniers des années 1988 à 1992. Pour le début de la série (années 1986 et 1987) et la fin de la série (années 1993 et 1994), on utilise des moyennes asymétriques prédéfinies (tableau 4, 3.3.3) :

- $APR86 = \frac{102.584*11 + 103.570*11 + 101.798*5}{27} = 102.840$

  (le point courant et deux points dans le futur)
- $APR87 = \frac{102.584*7 + 103.570*10 + 101.798*7 + 101.707*3}{27} = 102.648$

  (un point dans le passé, le point courant et deux points dans le futur)
- $APR93 = \frac{100.501*7 + 101.346*10 + 103.181*7 + 103.213*3}{27} = 101.810$

  (un point dans le futur, le point courant et deux points dans le passé)

- $APR94 = \frac{100.501*11+101.346*11+103.181*5}{27} = 101.342$

  (le point courant et deux points dans le passé)

B5a : FACTEURS SAISONNIERS PROVISOIRES (MM 3x3)

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	102.840	96.829	103.416	97.368	65.741	103.101	111.445	107.260	102.403
1987	102.516	101.065	109.073	102.648	97.137	103.706	96.627	66.587	102.838	111.346	107.976	102.070
1988	103.032	100.703	108.659	102.131	97.840	103.678	95.836	67.973	102.190	111.712	108.951	101.382
1989	104.253	99.799	107.911	101.866	98.276	103.034	95.270	69.226	101.058	112.182	109.534	100.276
1990	105.305	99.071	107.323	101.862	98.570	102.146	95.757	69.906	100.393	113.004	109.149	99.704
1991	105.993	98.806	107.066	102.261	98.030	101.921	96.219	70.121	100.420	113.188	108.248	99.674
1992	105.492	99.118	107.328	102.201	97.585	102.572	96.365	70.225	101.192	113.168	107.601	100.117
1993	104.728	99.232	107.758	101.810	97.132	103.467	95.892	70.420	101.881	112.948	107.373	100.397
1994	104.065	99.140	108.075	101.342	97.118	104.017	95.496	70.538	102.263	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

- ***Étape 2 : normalisation des coefficients saisonniers***.

B5b : MOYENNE MOBILE CENTRÉE SUR 12 MOIS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	100.247	100.247	100.247	100.247	100.247	100.247	100.247	100.251	100.276
1987	100.258	100.262	100.286	100.271	100.297	100.313	100.321	100.327	100.295	100.256	100.264	100.292
1988	100.258	100.282	100.313	100.301	100.357	100.369	100.392	100.405	100.336	100.294	100.301	100.292
1989	100.242	100.270	100.275	100.248	100.292	100.270	100.268	100.281	100.226	100.202	100.214	100.189
1990	100.172	100.221	100.221	100.228	100.246	100.206	100.211	100.229	100.207	100.213	100.207	100.175
1991	100.185	100.213	100.223	100.232	100.202	100.163	100.141	100.133	100.157	100.166	100.145	100.153
1992	100.186	100.197	100.233	100.265	100.237	100.229	100.215	100.188	100.211	100.212	100.177	100.196
1993	100.213	100.202	100.239	100.258	100.239	100.242	100.226	100.194	100.203	100.197	100.177	100.199
1994	100.206	100.194	100.215	100.215	100.215	100.215	100.215	100.215	100.215	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

Au tableau B5a, on applique une moyenne mobile centrée sur 12 mois. Le premier terme calculable est donc celui d'octobre 1986 et le dernier, celui de mars 1994. Ainsi :

$$\begin{aligned} OCT86 = {} & \frac{102.840}{24} \\ & + \frac{96.829+103.416+97.368+65.741+103.101+111.445+107.260+102.403+102.516+101.065+109.073}{12} \\ & + \frac{102.648}{24} \\ = {} & 100.247 \end{aligned}$$

Les six premières valeurs, d'avril à septembre 1986, non calculables avec cette moyenne mobile symétrique, seront prises égales à la première valeur calculable, celle d'octobre 1986. On procède de même pour la fin de série : la valeur calculée pour mars 1994 (100.215) est répétée pour les six mois suivants.

Les coefficients saisonniers normalisés sont alors obtenus en divisant le tableau B5a par le tableau B5b :

B5 : COEFFICIENTS SAISONNIERS NORMALISÉS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	111.171	106.991	102.120
1986	102.253	100.801	108.761	102.587	96.590	103.161	97.128	65.580	102.847	111.171	106.991	102.120
1987	102.253	100.801	108.761	102.370	96.850	103.382	96.319	66.370	102.536	111.062	107.692	101.773
1988	102.767	100.419	108.320	101.824	97.492	103.297	95.462	67.699	101.848	111.385	108.624	101.087
1989	104.002	99.530	107.614	101.614	97.991	102.757	95.016	69.032	100.830	111.956	109.301	100.086
1990	105.124	98.853	107.086	101.630	98.328	101.936	95.555	69.746	100.185	112.764	108.924	99.529
1991	105.797	98.596	106.828	102.024	97.832	101.755	96.083	70.028	100.262	113.000	108.091	99.521
1992	105.295	98.924	107.078	101.931	97.355	102.338	96.158	70.093	100.980	112.928	107.411	99.922
1993	104.505	99.033	107.501	101.548	96.900	103.218	95.676	70.284	101.675	112.725	107.183	100.197
1994	103.852	98.948	107.843	101.124	96.910	103.793	95.291	70.387	102.044	112.725	107.183	100.197
1995	103.852	98.948	107.843	.	.	.	.	.	.	.	.	.

$APR86 = 100 * (102.840 / 100.247) = 102.587$

Les valeurs manquantes d'octobre 1985 à mars 1986 sont obtenues en dupliquant la première valeur calculée pour le mois considéré. De même, pour les valeurs d'octobre 1994 à mars 1995, on duplique la dernière valeur calculée pour le mois considéré.

4.1.6 Tableau B6 : Estimation de la série corrigée des variations saisonnières

4.1.6.1 Description et mode de calcul

Cette estimation s'obtient simplement en retirant à la série de départ, du tableau B1, l'estimation de la composante saisonnière du tableau B5 : *B6 = B1 _op B5*

4.1.6.2 Exemple

B6: SÉRIE DÉSAISONNALISÉE PROVISOIRE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	104.074	102.626	98.511
1986	104.251	97.916	95.530	106.739	101.149	100.522	102.648	100.183	102.288	105.333	101.224	102.232
1987	98.286	102.380	103.805	104.620	103.253	104.757	105.691	103.511	106.011	105.257	106.507	108.084
1988	104.800	109.740	109.583	106.164	110.163	111.039	106.011	112.261	112.520	105.849	111.669	113.467
1989	113.363	112.730	111.695	112.878	112.766	117.072	111.140	115.019	113.260	113.169	116.010	112.603
1990	115.197	113.805	115.421	114.238	117.566	114.582	117.000	119.433	114.388	117.058	116.687	111.324
1991	116.544	114.406	111.675	117.031	115.810	114.687	120.000	116.525	116.096	117.168	115.458	116.357
1992	117.289	118.172	115.804	117.727	112.784	115.988	116.579	114.134	118.143	114.232	113.676	113.889
1993	108.799	114.205	114.138	112.459	111.352	113.449	112.985	113.398	112.909	107.340	113.544	114.574
1994	111.987	112.685	114.982	114.117	117.635	116.578	114.911	121.329	118.184	112.131	119.142	119.764
1995	119.498	117.536	120.731	.	.	.	.	.	.	.	.	.

Avec par exemple : $APR86 = 100*(109.500/102.587) = 106.739$

4.1.7 Tableau B7 : Estimation de la composante Tendance-Cycle

4.1.7.1 Description et mode de calcul

Ce tableau présente une estimation de la composante Tendance-Cycle réalisée à partir de la série désaisonnalisée du tableau précédent. Il s'agit donc ici d'un problème de lissage et pour le résoudre, le programme va utiliser une moyenne mobile de Henderson.

- ***Étape 1 : choix de la moyenne mobile, calcul du ratio*** $\bar{I}/\bar{C}$

Pour une série mensuelle, X11 utilise, à cette étape et selon les cas, une moyenne mobile de Henderson sur 9 ou 13 termes. Le choix de l'ordre de la moyenne mobile est, sauf intervention de l'utilisateur, automatique et fondé sur la valeur d'un indicateur dit "ratio $\bar{I}/\bar{C}$" mesurant en quelque sorte l'importance de la composante irrégulière dans la série ; plus elle est importante, plus l'ordre de la moyenne mobile choisie sera grand.
Pour calculer ce ratio, on calcule une première décomposition de la série CVS (Corrigée des Variations Saisonnières) à l'aide d'une moyenne mobile de Henderson sur 13 termes. A cette étape du calcul, on ne se préoccupe pas des 6 points "perdus" en début et fin de série.
On a donc : *Tendance-Cycle = C =* H_{13}*(Série CVS)* et *Irrégulier = I = Série CVS _op Tendance-Cycle*
On calcule ensuite, pour chaque série *C* et *I*, la moyenne de la valeur absolue des taux de croissance mensuels (schéma multiplicatif) ou des croissances mensuelles (schéma additif), notés $\bar{C}$ et $\bar{I}$. On a donc :

$$\bar{C} = \frac{1}{n-1}\sum_{t=2}^{n}\left|C_t \ \_op \ C_{t-1} - \ \_xbar\right| \quad \text{et} \quad \bar{I} = \frac{1}{n-1}\sum_{t=2}^{n}\left|I_t \ \_op \ I_{t-1} - \ \_xbar\right|$$

On calcule ensuite le ratio $\bar{I}/\bar{C}$ et :

- Si le ratio est inférieur à 1, on choisira une moyenne mobile de Henderson à 9 termes.
- Sinon on choisira une moyenne mobile de Henderson à 13 termes.

- ***Étape 2 : lissage de la série CVS par une moyenne mobile de Henderson***

La série CVS du tableau B6 est alors lissée par la moyenne mobile de Henderson choisie. A ce stade, on estimera les points non calculables par la moyenne symétrique en début et fin de série (4 ou 6 selon le cas) grâce à des moyennes mobiles asymétriques *ad hoc*.

4.1.7.2 Remarques

- Notez que le calcul du ratio se fait sans se préoccuper des 6 premiers et 6 derniers mois pour lesquels la moyenne mobile de Henderson sur 13 termes ne permet pas d'obtenir d'estimation de la Tendance-Cycle.
- A cette étape, le programme ne choisit qu'entre une moyenne à 9 termes et une moyenne à 13 termes.
- Il est possible de spécifier soit même la longueur de la moyenne mobile de Henderson à utiliser. Dans ce cas, X11 ARIMA donne le choix entre une moyenne mobile sur 9, 13 ou 23

termes. X12-ARIMA permet de choisir toute moyenne de Henderson d'ordre impair inferieur à 101.

- Les coefficients des moyennes mobiles utilisées (symétriques ou non) sont, aux arrondis près, les mêmes dans X11-ARIMA88 et X12-ARIMA. Les coefficients des moyennes mobiles symétriques sont calculés à partir de la formule exacte de Henderson. Pour les filtres asymétriques, X11-ARIMA88 utilise des valeurs à 7 décimales dérivées d'une formule établie par Laniel (1985) et X12-ARIMA utilise directement une formule due à Doherty (1991) et basée sur les travaux de Musgrave (1964).[24] Ces différents filtres sont présentés en 3.2.2 et 3.3.1.
- Notons enfin que le calcul de cette Tendance-Cycle se fait sans exclure de la série CVS les points jugés atypiques. Une option est cependant proposée à cette fin dans X11-ARIMA88 ; dans ce cas, le programme recherchera et corrigera ces valeurs atypiques de la même façon qu'au tableau B4 : estimation de la composante irrégulière en enlevant la Tendance-Cycle de la série CVS, calcul d'un écart-type mobile sur 5 ans …..

4.1.7.3 Exemple

- ***Étape 1 : choix de la moyenne mobile, calcul du ratio*** $\bar{I}/\bar{C}$

On lisse tout d'abord le tableau B6 par une moyenne mobile de Henderson sur 13 termes dont les coefficients figurent dans le tableau 3 (paragraphe 3.2.2 ,colonne H6_6).

Le premier terme calculable avec le filtre symétrique de Henderson est donc celui d'avril 1986, et on a, aux arrondis près :

$$
\begin{aligned}
APR86 = \; & 104.074*(-0.0193)+102.626*(-0.0279)+98.511*(0.0000)+\\
& 104.251*(0.0655)+97.916*(0.1474)+95.530*(0.2143)+\\
& 106.739*(0.2401)+101.149*(0.2143)+100.522*(0.1474)+\\
& 102.648*(0.0655)+100.183*(0.0000)+102.288*(-0.0279)+\\
& 105.333*(-0.0193) \qquad = 100.809
\end{aligned}
$$

A cette étape du calcul, on ne se préoccupe pas d'estimer les 6 points non calculables avec le filtre symétrique en début et en fin de série. On en déduit une estimation de la tendance (tableau B7a) et, par division avec le tableau B6, de la composante irrégulière (tableau B7b) :

B7a : TENDANCE-CYCLE (Moyenne Mobile de Henderson sur 13 termes)

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	100.809	101.258	101.649	102.031	102.287	102.241	102.092	101.939	101.700
1987	101.671	102.029	102.691	103.528	104.218	104.567	104.799	104.992	105.302	105.774	106.319	106.848
1988	107.460	107.972	108.320	108.737	109.126	109.403	109.568	109.760	110.159	110.671	111.282	111.855
1989	112.343	112.722	113.066	113.268	113.389	113.645	113.835	113.913	113.901	113.920	113.989	114.155
1990	114.366	114.521	114.877	115.348	115.889	116.472	116.816	116.862	116.614	116.065	115.337	114.704
1991	114.232	114.160	114.487	115.084	115.871	116.538	116.921	117.008	116.831	116.619	116.632	116.735
1992	116.829	116.824	116.503	116.091	115.767	115.602	115.688	115.709	115.381	114.694	113.877	113.108
1993	112.640	112.448	112.498	112.798	112.965	112.853	112.539	112.212	112.024	111.941	111.996	112.314
1994	112.954	113.648	114.346	115.193	116.069	116.819	117.188	117.307	117.362	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

[24] Dans X11, les coefficients des moyennes mobiles symétriques et asymétriques sont arrondis à la troisième décimale. On ne sait pas comment les filtres asymétriques ont été calculés.

B7b : IRRÉGULIER

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	105.882	99.893	98.891	100.605	97.944	100.046	103.174	99.298	100.524
1987	96.670	100.344	101.085	101.055	99.074	100.182	100.851	98.590	100.674	99.510	100.178	101.157
1988	97.525	101.638	101.166	97.634	100.951	101.495	96.753	102.279	102.143	95.643	100.348	101.441
1989	100.908	100.007	98.787	99.656	99.451	103.016	97.632	100.971	99.437	99.341	101.773	98.640
1990	100.727	99.375	100.473	99.037	101.447	98.378	100.158	102.200	98.091	100.856	101.171	97.053
1991	102.024	100.216	97.544	101.692	99.947	98.412	102.634	99.587	99.371	100.471	98.994	99.676
1992	100.394	101.154	99.400	101.409	97.423	100.334	100.770	98.639	102.393	99.598	99.823	100.691
1993	96.590	101.563	101.458	99.700	98.573	100.528	100.396	101.057	100.790	95.890	101.382	102.012
1994	99.144	99.152	100.556	99.066	101.350	99.793	98.057	103.428	100.701	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

Ainsi, $APR86 = 100 * (106.739/100.809) = 105.882$

Le schéma étant multiplicatif, on calcule :

$$\overline{C} = \frac{1}{n-1}\sum_{t=2}^{n}\left|C_t/C_{t-1} - 1\right| = \frac{1}{n-1}\sum_{t=2}^{n}\left|\frac{C_t - C_{t-1}}{C_{t-1}}\right| \quad \text{et} \quad \bar{I} = \frac{1}{n-1}\sum_{t=2}^{n}\left|I_t/I_{t-1} - 1\right| = \frac{1}{n-1}\sum_{t=2}^{n}\left|\frac{I_t - I_{t-1}}{I_{t-1}}\right|$$

B7c : CROISSANCE, EN VALEUR ABSOLUE ET EN %, DE LA TENDANCE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC	TOTAL
1985	.	.	.	.	.	.	.	.	.	.	.	.	0.000
1986	.	.	.	.	0.445	0.387	0.375	0.251	0.044	0.146	0.150	0.235	2.033
1987	0.028	0.352	0.649	0.814	0.667	0.335	0.222	0.184	0.295	0.449	0.514	0.498	5.008
1988	0.573	0.477	0.322	0.385	0.358	0.255	0.151	0.175	0.364	0.465	0.552	0.514	4.589
1989	0.437	0.337	0.305	0.178	0.106	0.226	0.168	0.069	0.011	0.017	0.061	0.146	2.060
1990	0.184	0.136	0.311	0.410	0.469	0.503	0.295	0.040	0.212	0.471	0.627	0.549	4.207
1991	0.412	0.063	0.286	0.522	0.684	0.575	0.329	0.075	0.151	0.182	0.011	0.089	3.379
1992	0.081	0.005	0.274	0.354	0.279	0.143	0.075	0.018	0.284	0.596	0.712	0.675	3.496
1993	0.414	0.170	0.045	0.266	0.148	0.099	0.278	0.291	0.167	0.074	0.049	0.284	2.285
1994	0.570	0.615	0.614	0.741	0.760	0.647	0.316	0.102	0.047	.	.	.	4.410
1995	.	.	.	.	.	.	.	.	.	.	.	.	0.000

Soit, en utilisant les totaux par ligne :

$$\overline{C} = \frac{2.033 + 5.008 + 4.589 + 2.060 + 4.207 + 3.379 + 3.496 + 2.285 + 4.410}{101} = 0.312$$

B7d : CROISSANCE, EN VALEUR ABSOLUE, DE L'IRRÉGULIER

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC	TOTAL
1985	.	.	.	.	.	.	.	.	.	.	.	.	0.000
1986	.	.	.	.	5.657	1.002	1.733	2.645	2.146	3.128	3.757	1.234	21.302
1987	3.833	3.801	0.738	0.029	1.961	1.118	0.668	2.242	2.114	1.156	0.671	0.978	19.307
1988	3.591	4.217	0.464	3.492	3.397	0.539	4.672	5.711	0.133	6.364	4.920	1.090	38.591
1989	0.526	0.893	1.219	0.879	0.205	3.585	5.226	3.420	1.519	0.097	2.448	3.079	23.096
1990	2.116	1.342	1.105	1.429	2.433	3.025	1.810	2.039	4.020	2.818	0.312	4.070	26.519
1991	5.122	1.772	2.666	4.252	1.715	1.536	4.290	2.968	0.218	1.107	1.470	0.689	27.805
1992	0.720	0.758	1.735	2.022	3.931	2.988	0.434	2.115	3.806	2.730	0.227	0.869	22.334
1993	4.073	5.148	0.103	1.733	1.130	1.984	0.131	0.658	0.264	4.862	5.728	0.622	26.435
1994	2.812	0.009	1.416	1.482	2.305	1.536	1.740	5.478	2.637	.	.	.	19.415
1995	.	.	.	.	.	.	.	.	.	.	.	.	0.000

$$\bar{I} = \frac{21.302 + 19.307 + 38.591 + 23.096 + 26.519 + 27.805 + 22.334 + 26.435 + 19.415}{101} = 2.226$$

donc $\dfrac{\bar{I}}{\overline{C}} = \dfrac{2.226}{0.312} = 7.14$

- ***Étape 2 : lissage de la série CVS par une moyenne mobile de Henderson***

Le ratio étant supérieur à 1, on choisit une moyenne mobile de Henderson sur 13 termes dont les coefficients et ceux des moyennes mobiles asymétriques associées figurent dans le tableau 3 (3.3.1).
L'estimation de la tendance pour octobre 1985 se fait, à partir de la série CVS du tableau B6, en utilisant le point courant et six points dans le futur auxquels on applique les coefficients de la moyenne mobile H13_0 du tableau H13.

$$OCT85 = 104.074*(0.42113)+102.626*(0.35315)+98.511*(0.24390)+104.251*(0.11977) +97.916*(0.01202)+95.530*(-0.05811)+ 106.739*(-0.09186) =102.405$$

Ce qui conduit au tableau B7 :

B7: TENDANCE-CYCLE
valeur de I/C : 7.1441
Une moyenne mobile de HENDERSON sur 13 termes a été choisie

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	102.405	101.784	101.095
1986	100.543	100.309	100.463	100.809	101.258	101.649	102.031	102.287	102.241	102.092	101.939	101.700
1987	101.671	102.029	102.691	103.528	104.218	104.567	104.799	104.992	105.302	105.774	106.319	106.848
1988	107.460	107.972	108.320	108.737	109.126	109.403	109.568	109.760	110.159	110.671	111.282	111.855
1989	112.343	112.722	113.066	113.268	113.389	113.645	113.835	113.913	113.901	113.920	113.989	114.155
1990	114.366	114.521	114.877	115.348	115.889	116.472	116.816	116.862	116.614	116.065	115.337	114.704
1991	114.232	114.160	114.487	115.084	115.871	116.538	116.921	117.008	116.831	116.619	116.632	116.735
1992	116.829	116.824	116.503	116.091	115.767	115.602	115.688	115.709	115.381	114.694	113.877	113.108
1993	112.640	112.448	112.498	112.798	112.965	112.853	112.539	112.212	112.024	111.941	111.996	112.314
1994	112.954	113.648	114.346	115.193	116.069	116.819	117.188	117.307	117.362	117.495	117.801	118.258
1995	118.787	119.246	119.901	.	.	.	.	.	.	.	.	.

4.1.8 Tableau B8 : Estimation de la composante Saisonnier-Irrégulier

4.1.8.1 Description et mode de calcul

Tableau similaire au tableau B3 : la composante Tendance-Cycle est enlevée de la série analysée, par soustraction ou division selon le schéma de composition adopté, pour obtenir une estimation de la composante Saisonnier-Irrégulier.
On a donc : *B8 = B1 _op B7*

4.1.8.2 Remarque

Contrairement au tableau B3, les points de début et de fin de série ayant été estimés par des moyennes mobiles de Henderson asymétriques pour la Tendance-Cycle, on dispose d'une estimation complète de la composante Saisonnier-Irrégulier.

4.1.8.3 Exemple

B8 : SÉRIE BRUTE DÉBARRASSÉE DE SA TENDANCE (estimation Saisonnier-Irrégulier)

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	112.983	107.876	99.511
1986	106.024	98.396	103.421	108.622	96.487	102.017	97.716	64.231	102.894	114.700	106.240	102.655
1987	98.848	101.148	109.941	103.451	95.953	103.570	97.138	65.434	103.227	110.518	107.883	102.950
1988	100.224	102.064	109.583	99.414	98.419	104.841	92.362	69.242	104.031	106.532	109.002	102.544
1989	104.946	99.537	106.309	101.264	97.453	105.856	92.766	69.702	100.263	111.218	111.239	98.725
1990	105.888	98.235	107.593	100.652	99.751	100.282	95.706	71.281	98.273	113.729	110.199	96.596
1991	107.938	98.809	104.204	103.750	97.781	100.139	98.614	69.739	99.631	113.533	107.003	99.199
1992	105.710	100.065	106.435	103.367	94.846	102.680	96.898	69.139	103.396	112.474	107.221	100.612
1993	100.941	100.580	109.069	101.243	95.517	103.763	96.055	71.026	102.478	108.092	108.664	102.213
1994	102.962	98.110	108.443	100.180	98.218	103.579	93.439	72.800	102.759	107.579	108.403	101.473
1995	104.473	97.530	108.590	.	.	.	.	.	.	.	.	.

La valeur d’avril 1986 s'obtient donc simplement : $APR86 = 100 * (109.500/100.809) = 108.622$

4.1.9 Tableau B9 : Valeurs de remplacement pour les valeurs atypiques de la composante Saisonnier-Irrégulier

4.1.9.1 Description et mode de calcul

Pour la seconde fois dans l'étape B, le programme va détecter et corriger automatiquement les valeurs atypiques de la composante Saisonnier-Irrégulier. La stratégie suivie est similaire à celle utilisée pour le tableau B4 (voir 4.1.4). La détection se fait à partir des données du tableau B8 et, contrairement à ce qui se passe pour le tableau B4, c'est une moyenne mobile 3x5, de coefficients {1/15,2/15,3/15,3/15,3/15,2/15,1/15}, qui est utilisée pour la première estimation de la composante saisonnière.

4.1.9.2 Remarques

- La moyenne mobile symétrique 3x5, sur 7 termes, ne permet pas d'estimer les coefficients saisonniers des 3 premières et 3 dernières années. Ce sont des moyennes mobiles asymétriques *ad hoc* qui sont utilisées (voir tableau 4, 3.3.3).
- Un problème pourrait intervenir dans l'application de ces moyennes mobiles asymétriques si on ne dispose pas d'assez d'années d'observations. Supposons que nous n'ayons, pour un mois donné, que 5 années d'observations. Le point central ne peut être estimé puisqu'on ne dispose ni de 3 points dans le futur, ni de 3 points dans le passé ce qui serait nécessaire pour pouvoir utiliser une moyenne mobile asymétrique. Dans ce cas, il sera estimé par la moyenne simple des 5 observations disponibles.
- Les autres remarques que l'on peut faire à ce stade sont les mêmes que celles concernant le tableau B4 (voir 4.1.4.2)

4.1.9.3 Exemple

Voici le tableau en sortie de X11. Nous allons, pour mieux le comprendre, détailler les étapes de son calcul à partir de tableaux malheureusement non récupérables dans les versions usuelles des logiciels de la famille X11 (tableaux numérotés ici de B9a à B9f).

B9: VALEURS DE REMPLACEMENT POUR LES POINTS EXTRÊMES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC	ECT
1985	.	.	.	.	.	.	.	.	.	.	.	.	2.077
1986	104.457	.	107.611	101.329	.	.	.	68.245	.	.	.	.	2.077
1987	103.337	.	.	.	.	.	.	.	.	.	.	.	2.077
1988	.	.	.	.	.	.	.	.	.	111.877	.	.	2.104
1989	.	.	.	.	.	.	.	.	.	.	.	.	1.885
1990	.	.	.	.	.	.	.	.	101.123	.	.	99.679	1.808
1991	105.353	.	106.753	.	.	.	95.836	.	.	.	.	.	1.609
1992	.	.	.	.	.	.	.	.	.	.	.	.	1.625
1993	104.314	.	.	.	.	.	.	.	.	.	.	.	1.603
1994	.	.	.	.	.	.	95.015	70.697	.	.	.	.	1.603
1995	.	.	.	.	.	.	.	.	.	.	.	.	1.603

- ***Étape 1 : estimation de la composante saisonnière***.

B9a: FACTEURS SAISONNIERS PROVISOIRES (mm 3x5)

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	111.803	107.583	101.831
1986	102.186	100.386	107.447	103.444	97.028	103.833	95.293	66.812	102.916	111.496	107.899	101.660
1987	102.573	100.199	107.399	103.147	97.231	103.700	95.192	67.258	102.471	111.238	108.370	101.295
1988	103.089	100.040	107.403	102.525	97.520	103.441	95.150	68.070	101.809	111.212	108.809	100.658
1989	103.899	99.859	107.240	102.025	97.711	103.004	95.241	68.959	101.314	111.314	108.970	100.077
1990	104.532	99.715	107.024	101.817	97.530	102.747	95.532	69.692	101.004	111.471	108.955	99.669
1991	104.905	99.449	106.899	101.861	97.314	102.464	95.807	70.265	101.078	111.463	108.699	99.674
1992	104.726	99.208	107.073	102.010	96.961	102.391	96.134	70.550	101.394	111.103	108.329	100.121
1993	104.205	99.083	107.544	101.929	96.722	102.550	96.058	70.779	101.975	110.432	108.036	100.701
1994	103.670	98.987	107.985	101.920	96.432	102.861	95.937	70.801	102.391	110.004	107.932	101.098
1995	103.230	98.939	108.361	.	.	.	.	.	.	.	.	.

Les données du tableau B8 sont lissées colonne par colonne (mois par mois), avec une moyenne mobile 3x5 dont les coefficients et ceux des moyennes asymétriques associées figurent dans le tableau 4 (3.3.3).

Ainsi, les valeurs de la composante Saisonnier-Irrégulier du mois d'avril sont :

	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
Avril	.	108.622	103.451	99.414	101.264	100.652	103.750	103.367	101.243	100.180	

Le facteur saisonnier du mois d'avril 1989 sera donc estimé par :

$$APR89 = \frac{108.622 + 103.451 * 2 + 99.414 * 3 + 101.264 * 3 + 100.652 * 3 + 103.750 * 2 + 103.367}{15} = 102.025$$

Cette moyenne mobile symétrique peut s'appliquer pour estimer les valeurs des coefficients saisonniers des années 1989 à 1991. Pour le début de la série (années 1986 à 1988) et la fin de la série (années 1992 à 1994), on utilise des moyennes asymétriques prédéfinies, par exemple :

$$APR87 = \frac{108.622 * 15 + 103.451 * 15 + 99.414 * 15 + 101.264 * 11 + 100.652 * 4}{60} = 103.147$$

(un point dans le passé, le point courant et trois points dans le futur)

- ***Étape 2 : normalisation des coefficients saisonniers***.

B9b: MOYENNE MOBILE CENTRÉE SUR 12 MOIS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.034	100.034	100.034
1986	100.034	100.034	100.034	100.034	100.034	100.041	100.050	100.058	100.048	100.034	100.030	100.033
1987	100.023	100.037	100.037	100.008	100.017	100.021	100.028	100.042	100.036	100.010	99.996	99.998
1988	99.985	100.017	100.023	99.995	100.012	100.004	100.011	100.037	100.023	99.995	99.982	99.972
1989	99.958	99.998	100.015	99.999	100.010	99.992	99.994	100.015	100.000	99.982	99.966	99.948
1990	99.949	99.992	100.009	100.003	100.009	99.991	99.990	99.994	99.978	99.974	99.967	99.946
1991	99.946	99.981	100.008	100.011	100.000	99.990	99.982	99.965	99.962	99.975	99.967	99.949
1992	99.960	99.985	100.010	100.009	99.978	99.981	99.978	99.951	99.966	99.982	99.969	99.965
1993	99.969	99.975	100.009	100.005	99.965	99.977	99.979	99.953	99.967	99.985	99.973	99.973
1994	99.981	99.977	99.995	99.995	99.973	99.985	99.983	99.963	99.976	99.976	99.976	99.976
1995	99.976	99.976	99.976	.	.	.	.	.	.	.	.	.

Au tableau B9a, on applique une moyenne mobile centrée sur 12 mois. Le premier terme calculable est donc celui de avril 1986 et le dernier, celui de septembre 1994. Ainsi :

$$
\begin{aligned}
APR86 = {} & \frac{111.803}{24} \\
& + \frac{107.583+101.831+102.186+100.386+107.447+103.444+97.028+103.833+95.293+66.812+102.916}{12} \\
& + \frac{111.496}{24} \\
= {} & 100.034
\end{aligned}
$$

Les six premières valeurs, de octobre 1985 à mars 1986, non calculables avec cette moyenne mobile symétrique, seront prises égales à la première valeur calculable, celle de avril 1986. On procède de même pour la fin de série : la valeur calculée pour septembre 1994 (99.976) est répétée pour les six mois suivants.
Les coefficients saisonniers normalisés sont alors obtenus en divisant le tableau B9a par le tableau B9b pour obtenir :

B9c: FACTEURS SAISONNIERS NORMALISÉS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	111.765	107.547	101.796
1986	102.151	100.352	107.411	103.409	96.994	103.791	95.246	66.774	102.866	111.458	107.867	101.627
1987	102.550	100.162	107.359	103.139	97.215	103.678	95.165	67.229	102.434	111.226	108.374	101.298
1988	103.104	100.023	107.378	102.530	97.509	103.437	95.140	68.045	101.786	111.217	108.828	100.686
1989	103.943	99.860	107.224	102.027	97.702	103.012	95.247	68.949	101.314	111.334	109.008	100.130
1990	104.585	99.724	107.014	101.814	97.522	102.756	95.542	69.697	101.027	111.499	108.990	99.723
1991	104.961	99.468	106.890	101.850	97.314	102.475	95.824	70.290	101.117	111.490	108.735	99.725
1992	104.768	99.223	107.062	102.001	96.982	102.410	96.155	70.584	101.429	111.123	108.363	100.155
1993	104.237	99.108	107.534	101.923	96.756	102.574	96.078	70.812	102.009	110.449	108.066	100.728
1994	103.689	99.010	107.990	101.925	96.458	102.876	95.953	70.827	102.415	110.030	107.958	101.122
1995	103.254	98.962	108.386	.	.	.	.	.	.	.	.	.

Ainsi, $APR86 = 100*(103.444/100.034) = 103.409$

- ***Étape 3 : estimation de la composante irrégulière.***

Il suffit de corriger la composante Saisonnier-Irrégulier du tableau B8 par les facteurs saisonniers normalisés du tableau B9c ci-dessus.

B9d: IRRÉGULIER PROVISOIRE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	101.090	100.306	97.755
1986	103.792	98.051	96.285	105.041	99.477	98.291	102.593	96.192	100.027	102.909	98.491	101.011
1987	96.390	100.984	102.405	100.303	98.702	99.896	102.073	97.330	100.774	99.363	99.547	101.631
1988	97.206	102.040	102.054	96.961	100.933	101.358	97.081	101.760	102.206	95.787	100.160	101.845
1989	100.965	99.676	99.147	99.252	99.745	102.761	97.395	101.092	98.962	99.896	102.047	98.597
1990	101.246	98.507	100.541	98.859	102.286	97.592	100.173	102.273	97.275	102.000	101.109	96.865
1991	102.836	99.338	97.488	101.865	100.480	97.721	102.912	99.216	98.531	101.832	98.408	99.472
1992	100.899	100.849	99.414	101.340	97.798	100.264	100.773	97.952	101.940	101.215	98.946	100.456
1993	96.838	101.485	101.427	99.333	98.720	101.160	99.976	100.302	100.459	97.866	100.554	101.475
1994	99.299	99.091	100.419	98.288	101.824	100.683	97.380	102.785	100.336	97.772	100.413	100.347
1995	101.180	98.553	100.188	.	.	.	.	.	.	.	.	.

Ainsi, $APR86 = 100*(108.622/103.409) = 105.041$

- ***Étape 4 :.Calcul d'un écart-type mobile.***

L'écart-type correspondant à l'année 1989 sera calculé à partir des données des années 1987 à 1991 (deux années avant, deux années après) selon la formule suivante :[25]

$$\sigma_{89} = \sqrt{\frac{1}{60}\sum_{i=1}^{60}(x_i - 100)^2} = 1.8846$$

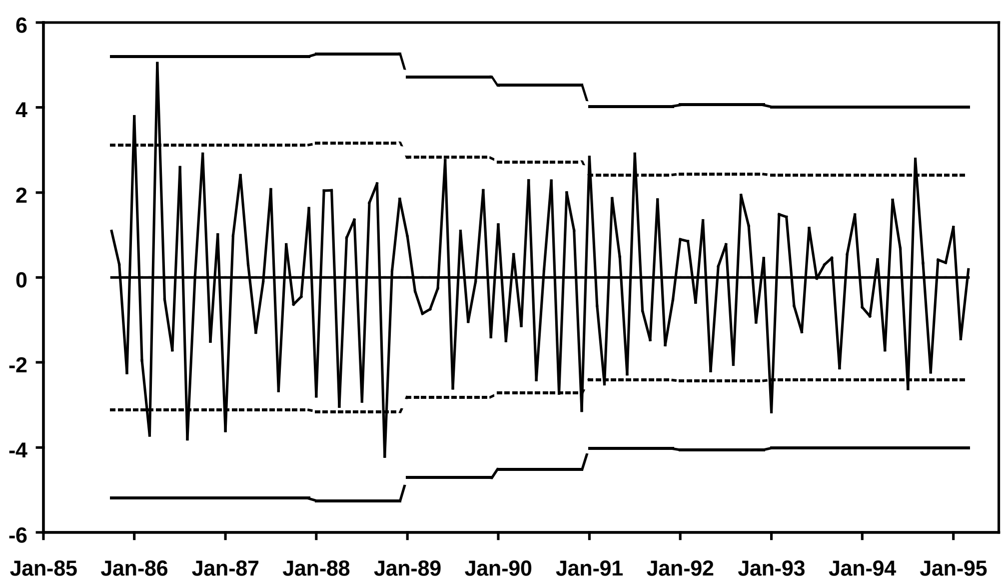
Ceux des années 1988, 1990, 1991 et 1992 sont calculés selon le même principe.
Pour X11-ARIMA88 et X12-ARIMA, l'écart-type de 1987 est calculé à partir de l'ensemble des observations disponibles de 1985 à 1990 soit 63 observations. Les résultats du calcul de X11-ARIMA88 et X12-ARIMA figurent dans le tableau B9e, colonne Écart-type 1.

B9e: VALEURS DES ÉCART-TYPES MOBILES SUR CINQ ANS

ANNÉE	ÉCART-TYPE 1	ÉCART-TYPE 2
1985	2.0774	2.0774
1986	2.0774	2.0774
1987	2.0774	2.0774
1988	2.1038	2.1038
1989	1.8846	1.8846
1990	1.8082	1.8082
1991	1.6093	1.6093
1992	1.6246	1.6246
1993	1.6030	1.6030
1994	1.6030	1.6030
1995	1.6030	1.6030

Ce premier calcul sert à repérer d'éventuels points atypiques. Pour une année donnée, une valeur sera considérée comme aberrante si elle s'éloigne, en valeur absolue de son écart à la moyenne théorique, de plus de 2.5 fois l'écart-type correspondant à cette année. Ici, d'après le graphique suivant, aucune valeur n'est détectée comme très atypique. Le nouveau calcul de l'écart-type conduit donc aux mêmes résultats (colonne Écart-type 2 du tableau B9e).

Figure 14 : Écart à la moyenne de l'irrégulier et "limites de confiance" (tableau B9)



[25] où la moyenne théorique est considérée égale à 100 pour tenir compte du fait que les valeurs de l'irrégulier ont été elles-mêmes multipliées par 100

- ***Étape 5 : Détection des valeurs atypiques et pondération de l'irrégulier.***

B9f: POIDS ASSOCIÉS AUX VALEURS DE L'IRRÉGULIER

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.000	100.000	100.000
1986	67.475	100.000	71.178	7.340	100.000	100.000	100.000	66.711	100.000	100.000	100.000	100.000
1987	76.235	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	49.731	100.000	100.000
1989	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	99.273	100.000	100.000	76.630
1991	73.758	100.000	93.885	100.000	100.000	100.000	69.057	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1993	52.737	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1994	100.000	100.000	100.000	100.000	100.000	100.000	86.573	76.236	100.000	100.000	100.000	100.000
1995	100.000	100.000	100.000	.	.	.	.	.	.	.	.	.

Les valeurs de l'irrégulier sont alors situées par rapport aux limites de confiance supérieures et inférieures calculées à partir des écart-types estimés auparavant. Toutes les valeurs situées au delà des limites de confiance inférieurs sont considérées comme atypiques et vont donc être corrigées, à des degrés divers, et les poids (multipliés par 100) associés à chacune de ces valeurs figurent dans le tableau B9f.

Ainsi, pour octobre 1988, on a :

$$|OCT88 - 100| = |95.787 - 100| = 4.213$$

$$1.5 * \sigma_{88} = 1.5 * 2.1038 = 3.1557 < 4.213 < 2.5 * \sigma_{88} = 2.5 * 2.1038 = 5.2595$$

On va attribuer à cette valeur jugée moyennement atypique, un poids proportionnel à l'écart à la moyenne constaté de :

$$poids(OCT88) = \frac{5.2595 - 4.213}{5.2595 - 3.1557} = 0.497$$

- ***Étape 6 : Correction des valeurs atypiques de la composante Saisonnier-Irrégulier.***

La correction de la composante Saisonnier-Irrégulier (tableau B8) se fait enfin à partir de ces poids. Ainsi, la valeur d'octobre 1988 sera remplacée par la moyenne de cette valeur affectée de son poids et des deux valeurs antérieures et postérieures du même mois ayant reçu une pondération intégrale, donc non jugées atypiques. Il s'agit donc, comme le montre le tableau B9f, des valeurs des mois d'octobre 1986, 1987, 1990 et 1991. Ce qui conduit à :

$$SI(OCT88) = \frac{114.700 + 110.518 + 106.532 * 0.497 + 111.218 + 113.729}{4 + 0.497} = 111.877$$

La valeur de janvier 1986 est jugée elle aussi atypique. Mais, comme elle se trouve en début de série, elle est corrigée de façon différente : elle est remplacée par la moyenne de cette valeur affectée de son poids et des quatre valeurs les plus proches du même mois ayant reçu une pondération intégrale. Il s'agit donc dans ce cas, d'après le tableau B9f, des valeurs de janvier 1988, 1989, 1990 et 1992 et on a :

$$SI(JAN86) = \frac{106.024 * 0.675 + 100.224 + 104.946 + 105.888 + 105.710}{4 + 0.675} = 104.457$$

4.1.10 Tableau B10 : Estimation de la composante saisonnière

4.1.10.1 Description et mode de calcul

Cette estimation est obtenue d'une façon similaire à celle conduisant au tableau B5 (voir 4.1.5) à partir des valeurs de la composante Saisonnier-Irrégulier du tableau B8 corrigées par les valeurs du tableau B9. On procède là encore en deux étapes : estimation de la composante saisonnière puis normalisation des coefficients saisonniers.
Comme au tableau B9, et à la différence du tableau B5, c'est une moyenne mobile 3x5 qui est utilisée pour l'estimation de la composante saisonnière.

4.1.10.2 Exemple

L'estimation est faite à partir de la composante Saisonnier-Irrégulier corrigée :

B9g: COMPOSANTE SAISONNIER-IRRÉGULIER CORRIGÉE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	112.983	107.876	99.511
1986	104.457	98.396	107.611	101.329	96.487	102.017	97.716	68.245	102.894	114.700	106.240	102.655
1987	103.337	101.148	109.941	103.451	95.953	103.570	97.138	65.434	103.227	110.518	107.883	102.950
1988	100.224	102.064	109.583	99.414	98.419	104.841	92.362	69.242	104.031	111.877	109.002	102.544
1989	104.946	99.537	106.309	101.264	97.453	105.856	92.766	69.702	100.263	111.218	111.239	98.725
1990	105.888	98.235	107.593	100.652	99.751	100.282	95.706	71.281	101.123	113.729	110.199	99.679
1991	105.353	98.809	106.753	103.750	97.781	100.139	95.836	69.739	99.631	113.533	107.003	99.199
1992	105.710	100.065	106.435	103.367	94.846	102.680	96.898	69.139	103.396	112.474	107.221	100.612
1993	104.314	100.580	109.069	101.243	95.517	103.763	96.055	71.026	102.478	108.092	108.664	102.213
1994	102.962	98.110	108.443	100.180	98.218	103.579	95.015	70.697	102.759	107.579	108.403	101.473
1995	104.473	97.530	108.590	.	.	.	.	.	.	.	.	.

- ***Étape 1 : estimation de la composante saisonnière***.

Les données du tableau précédent sont lissées colonne par colonne (mois par mois), avec une moyenne mobile 3x5 (voir tableau 4, 3.3.3), pour aboutir au tableau B10a.

Le facteur saisonnier du mois d'avril 1989 sera donc estimé par :

$$APR89 = \frac{101.329 + 2*103.451 + 3*99.414 + 3*101.264 + 3*100.652 + 2*103.750 + 103.367}{15} = 101.539$$

Pour le début de la série (années 1986 à 1988) et la fin de la série (années 1992 à 1994), on utilise des moyennes asymétriques prédéfinies, par exemple :

$$APR87 = \frac{101.329*15 + 103.451*15 + 99.414*15 + 101.264*11 + 100.652*4}{60} = 101.324$$

(un point dans le passé, le point courant et trois points dans le futur)

B10a: FACTEURS SAISONNIERS PROVISOIRES (mm 3x5)

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	112.605	107.583	101.831
1986	103.013	100.386	108.635	101.378	97.028	103.833	95.293	67.950	102.916	112.476	107.899	101.660
1987	103.304	100.199	108.447	101.324	97.231	103.700	95.192	68.261	102.661	112.396	108.370	101.501
1988	103.654	100.040	108.201	101.431	97.520	103.441	94.965	68.672	102.189	112.281	108.809	101.069
1989	104.048	99.859	107.859	101.539	97.711	103.004	94.871	69.227	101.884	112.383	108.970	100.694
1990	104.539	99.715	107.534	101.817	97.530	102.747	94.976	69.692	101.574	112.184	108.955	100.286
1991	104.837	99.449	107.408	101.861	97.314	102.464	95.356	70.125	101.648	111.819	108.699	100.291
1992	104.883	99.208	107.583	102.010	96.961	102.391	95.769	70.235	101.774	111.103	108.329	100.532
1993	104.591	99.083	107.883	101.929	96.722	102.550	95.943	70.253	102.165	110.432	108.036	100.906
1994	104.341	98.987	108.155	101.920	96.432	102.861	95.967	70.205	102.391	110.004	107.932	101.098
1995	104.185	98.939	108.361	.	.	.	.	.	.	.	.	.

- ***Étape 2 : normalisation des coefficients saisonniers***.

B10b: MOYENNE MOBILE CENTRÉE SUR 12 MOIS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.199	100.199	100.199
1986	100.199	100.199	100.199	100.199	100.207	100.213	100.218	100.222	100.206	100.196	100.202	100.205
1987	100.196	100.204	100.207	100.193	100.209	100.222	100.230	100.238	100.221	100.215	100.232	100.233
1988	100.213	100.221	100.218	100.194	100.207	100.207	100.206	100.215	100.193	100.183	100.196	100.185
1989	100.163	100.182	100.193	100.184	100.195	100.186	100.191	100.206	100.186	100.184	100.188	100.170
1990	100.164	100.187	100.194	100.173	100.164	100.146	100.141	100.143	100.126	100.123	100.116	100.095
1991	100.099	100.133	100.154	100.142	100.116	100.106	100.108	100.100	100.097	100.110	100.102	100.084
1992	100.098	100.120	100.130	100.105	100.060	100.055	100.053	100.035	100.042	100.052	100.038	100.035
1993	100.049	100.057	100.074	100.062	100.022	100.026	100.031	100.016	100.024	100.035	100.022	100.023
1994	100.037	100.036	100.043	100.035	100.013	100.016	100.018	100.009	100.016	100.016	100.016	100.016
1995	100.016	100.016	100.016	.	.	.	.	.	.	.	.	.

Au tableau B10a, on applique une moyenne mobile centrée sur 12 mois. Le premier terme calculable est donc celui de avril 1986 et le dernier, celui de septembre 1994. Ainsi :

$$
\begin{aligned}
APR86 = & \frac{112.605}{24} \\
& + \frac{107.583+101.831+103.013+100.386+108.635+101.378+97.028+103.833+95.293+67.950+102.916}{12} \\
& + \frac{112.746}{24} \\
& = 100.199
\end{aligned}
$$

Les six premières valeurs, de octobre 1985 à mars 1986, non calculables avec cette moyenne mobile symétrique, seront prises égales à la première valeur calculable, celle de avril 1986. On procède de même pour la fin de série : la valeur calculée pour septembre 1994 (100.016) est répétée pour les six mois suivants. Les coefficients saisonniers normalisés sont alors obtenus en divisant le tableau B10a par le tableau B10b, pour obtenir :

B10 : FACTEURS SAISONNIERS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	112.382	107.370	101.629
1986	102.809	100.187	108.419	101.177	96.828	103.613	95.086	67.799	102.704	112.255	107.681	101.452
1987	103.102	99.995	108.223	101.129	97.028	103.470	94.973	68.099	102.435	112.154	108.120	101.265
1988	103.433	99.820	107.966	101.235	97.319	103.227	94.770	68.525	101.992	112.076	108.596	100.882
1989	103.879	99.677	107.652	101.352	97.521	102.812	94.690	69.085	101.694	112.177	108.766	100.523
1990	104.368	99.529	107.326	101.641	97.371	102.597	94.842	69.593	101.446	112.046	108.828	100.190
1991	104.734	99.317	107.243	101.717	97.201	102.356	95.253	70.055	101.550	111.696	108.588	100.207
1992	104.780	99.089	107.443	101.902	96.902	102.335	95.718	70.210	101.731	111.046	108.287	100.497
1993	104.540	99.027	107.804	101.865	96.700	102.524	95.913	70.242	102.141	110.394	108.012	100.883
1994	104.302	98.952	108.108	101.884	96.419	102.844	95.950	70.199	102.374	109.987	107.915	101.082
1995	104.169	98.923	108.343	.	.	.	.	.	.	.	.	.

Avec par exemple : $APR86 = 100*(101.378/100.199) = 101.177$

4.1.11 Tableau B11 : Estimation de la série corrigée des variations saisonnières

4.1.11.1 Description et mode de calcul

Cette estimation s'obtient simplement en retirant à la série de départ, du tableau B1, l'estimation de la composante saisonnière du tableau B10 : *B11 = B1 _op B10*

4.1.11.2 Exemple

B11: SÉRIE DÉSAISONNALISÉE PROVISOIRE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	102.953	102.263	98.988
1986	103.687	98.516	95.832	108.226	100.901	100.084	104.853	96.904	102.430	104.316	100.575	102.906
1987	97.476	103.205	104.322	105.905	103.063	104.668	107.188	100.882	106.117	104.231	106.086	108.626
1988	104.125	110.399	109.942	106.781	110.359	111.115	106.785	110.908	112.362	105.197	111.698	113.697
1989	113.498	112.563	111.656	113.169	113.309	117.010	111.522	114.931	112.297	112.947	116.581	112.114
1990	116.032	113.033	115.163	114.225	118.722	113.843	117.881	119.696	112.967	117.809	116.789	110.589
1991	117.727	113.576	111.243	117.385	116.563	114.014	121.046	116.480	114.624	118.536	114.930	115.561
1992	117.866	117.975	115.410	117.760	113.310	115.991	117.114	113.944	117.270	116.168	112.755	113.238
1993	108.762	114.211	113.818	112.109	111.582	114.217	112.706	113.466	112.393	109.607	112.672	113.795
1994	111.503	112.681	114.700	113.266	118.234	117.654	114.123	121.655	117.803	114.923	118.334	118.716
1995	119.134	117.566	120.173	.	.	.	.	.	.	.	.	.

Avec par exemple : $APR86 = 100*(109.500/101.177) = 108.226$

4.1.12 Tableau B13 : Estimation de la composante irrégulière

4.1.12.1 Description et mode de calcul

Cette estimation s'obtient simplement en retirant à la série corrigée des variations saisonnière, du tableau B11, l'estimation de la composante Tendance-Cycle du tableau B7 : *B13 = B11 _op B7*

4.1.12.2 Exemple

B13: SÉRIE DES FACTEURS IRRÉGULIERS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.535	100.471	97.916
1986	103.127	98.213	95.390	107.358	99.648	98.460	102.766	94.737	100.185	102.178	98.661	101.186
1987	95.874	101.153	101.588	102.296	98.892	100.097	102.280	96.086	100.774	98.541	99.781	101.664
1988	96.897	102.248	101.498	98.202	101.130	101.564	97.460	101.047	101.999	95.053	100.374	101.647
1989	101.028	99.859	98.753	99.913	99.930	102.961	97.968	100.893	98.592	99.146	102.274	98.212
1990	101.457	98.700	100.249	99.026	102.444	97.743	100.912	102.425	96.873	101.503	101.259	96.413
1991	103.060	99.489	97.167	101.999	100.597	97.834	103.528	99.549	98.111	101.645	98.541	98.994
1992	100.887	100.985	99.061	101.438	97.878	100.337	101.233	98.474	101.637	101.286	99.015	100.115
1993	96.558	101.568	101.173	99.389	98.776	101.209	100.148	101.117	100.329	97.915	100.604	101.319
1994	98.715	99.149	100.309	98.327	101.865	100.715	97.384	103.706	100.376	97.811	100.452	100.387
1995	100.292	98.592	100.227	.	.	.	.	.	.	.	.	.

Avec par exemple : $APR86 = 100*(108.226/100.809) = 107.358$

4.1.13 Estimation de l'effet dû à la composition journalière du mois (Trading-Days)

Certaines séries économiques, comme le chiffre d'affaire mensuel d'un commerce de détail, sont fortement influencées par la composition journalière du mois : un samedi de plus ou de moins dans un mois peut faire varier le chiffre d'affaire mensuel de façon non négligeable. Ces "effets de jours ouvrables", comme la saisonnalité, peuvent rendre délicates les comparaisons des valeurs de la série d'un mois à l'autre, ou celles d'un même mois d'une année à l'autre, et c'est pourquoi, si ces effets sont jugés statistiquement significatifs, ils sont généralement ôtés de la série lors du processus de désaisonnalisation.
La composante irrégulière, celle du tableau B13 par exemple, ne contient par construction ni tendance, ni saisonnalité. Il est donc assez naturel d'y rechercher les effets de jours ouvrables que l'on explique par un modèle de régression linéaire.
Il est possible, à la demande de l'utilisateur, de faire une estimation et une correction automatique de ces effets. C'est l'objet des tableaux B14, B15, B16, B18 et B19 de la partie B et de leurs équivalents dans la partie C.

4.1.13.1 Quelques particularités de notre calendrier

Notre calendrier est basé sur le rythme solaire. La Terre réalise une révolution complète autour du soleil en 365 jours et 6 heures environ. Pour tenir compte de ces 6 heures supplémentaires, le calendrier admet 365 jours 3 années sur 4 et la quatrième année (dite *bissextile*) compte un jour supplémentaire, le 29 février. La règle retenue étant que l'année bissextile correspond à toute année dont l'expression numérique est parfaitement divisible par 4. Cette correction est cependant trop forte. C'est pourquoi, les années séculaires ne sont pas bissextiles, sauf si elles sont divisibles par 400 (1600, 2000, 2400, ...). Il subsiste néanmoins de la sorte une petite erreur estimée à un jour pour 4000 ans ... En termes de mois, notre calendrier est donc périodique de période 4, au moins entre 1901 et 2099.
Pour une date donnée, le jour correspondant subit un décalage temporel : si le premier janvier d'une année non bissextile est un samedi, il sera un dimanche l'année suivante et un lundi, si l'année de référence est bissextile. Il faut donc attendre 28 ans (4 fois 7) pour retrouver la même structure annuelle en termes de dates et de jours.

4.1.13.2 Les effets de jours ouvrables

Nous supposerons par la suite, en reprenant les notations de Findley et *alii* (1998), que le *j*-ième jour de la semaine a un effet α_j où, par exemple, *j*=1 désigne le lundi, *j*=2 le mardi ...., et *j*=7 le dimanche. Chaque α_j représente par exemple les ventes moyennes pour un jour *j*. Si D_{jt} représente le nombre de jours *j* dans le mois *t*, la longueur du mois sera $N_t = \sum_{j=1}^{7} D_{jt}$ et l'effet cumulatif pour ce mois, les ventes totales du mois, sera : $\sum_{j=1}^{7} \alpha_j D_{jt}$. Notons par ailleurs $\overline{\alpha} = \frac{1}{7}\sum_{j=1}^{7} \alpha_j$ l'effet journalier moyen, les ventes d'un jour en moyenne.

Comme par construction $\sum_{j=1}^{7} \left(\alpha_j - \overline{\alpha}\right) = 0$, on peut écrire :

$$\sum_{j=1}^{7}\alpha_j D_{jt} = \overline{\alpha}N_t + \sum_{j=1}^{7}\left(\alpha_j - \overline{\alpha}\right)D_{jt} = \overline{\alpha}N_t + \sum_{j=1}^{6}\left(\alpha_j - \overline{\alpha}\right)\left(D_{jt} - D_{7t}\right) \qquad (1)$$

Ainsi, l'effet cumulatif du mois se décompose en un effet directement lié à la longueur du mois et un effet net de chaque jour de la semaine.

On peut remarquer que la somme $\sum_{j=1}^{7}\left(\alpha_j - \overline{\alpha}\right)D_{jt}$ ne fait en réalité intervenir que les jours de la semaine apparaissant 5 fois dans le mois : tout mois contient 4 semaines complètes, pour lesquelles par définition l'effet lié aux jours s'annule, plus 0, 1, 2 ou 3 jours qui participent à l'effet jours ouvrables du mois.

4.1.13.3 Le modèle de régression

Pour être homogène avec notre variable à expliquer B13 qui ne contient ni saisonnalité ni tendance, l'équation (1) doit être corrigée de ces effets.

- La partie $\overline{\alpha}N_t$ de l'équation contient potentiellement de telles composantes parce que les mois sont de longueur différente et que, comme on l'a vu ci-dessus, la variable N_t est périodique de période 48 mois (4 ans). On peut résumer ces effets par la quantité $\overline{\alpha}N_t^*$ où N_t^* représente la moyenne, sur 4 ans, de la longueur du mois t. En d'autres termes, N_t^* est égal à 30 ou 31 si le mois considéré n'est pas un mois de Février et à 28.25 dans le cas contraire. Ainsi, on a :
  $\overline{\alpha}N_t = \overline{\alpha}N^*{}_t + \overline{\alpha}(N_t - N_t^*)$ équation dont la seconde partie est nulle sauf pour le mois de février.
- La seconde partie de l'équation fait intervenir les D_{jt}, nombre de fois où le jour *j* est présent dans le mois *t*. Ces variables sont périodiques de période 336 mois (28 ans) et de moyennes égales : il apparaît en moyenne sur un cycle, dans un mois de mars, autant de lundi, mardi, mercredi .....[26]. Dans la seconde partie de l'équation c'est la différence $D_{jt} - D_{7t}$ qui intervient et, comme ces variables ont même comportement, il n'y a dans cette différence ni saisonnalité, ni tendance.

La façon dont on corrige l'équation (1) de ces effets dépend du schéma de composition adopté.

- Pour un **schéma multiplicatif**, on élimine les effets saisonnier et tendanciel en divisant l'équation (1) par $\overline{\alpha}N_t^*$. En posant $\beta_j = \alpha_j / \overline{\alpha} - 1$, cela conduit à :

$$\frac{N_t}{N^*{}_t} + \sum_{j=1}^{6}\beta_j\left(\frac{D_{jt} - D_{7t}}{N^*{}_t}\right) = \frac{\sum_{j=1}^{7}(\beta_j + 1)D_{jt}}{N^*{}_t} \qquad (2)$$

  Si $\hat{I}_t$ est une estimation de la composante irrégulière, X11 et ses descendants directs estiment les coefficients $\beta_1, \beta_2, \cdots, \beta_6$ (et $\beta_7 = -\sum_{j=1}^{6}\beta_j$) en ajustant par les moindres carrés ordinaires le modèle :

[26] Pour un mois de 31 jours, cette moyenne commune est de 4.428574, pour un mois de 30 jours de 4.285714 et pour un mois de février, de 4.035714.

$$N_t^{\ *}\hat{I}_t - N_t = \sum_{j=1}^{6} \beta_j (D_{jt} - D_{7t}) + e_t \qquad (3)$$

Ce qui correspond au modèle proposé par Young (1965).

- Pour un schéma additif, on doit soustraire logiquement $\overline{\alpha} N_t^{\ *}$ à l'équation (1). Ce qui conduit à :

$$\hat{I}_t = \beta_0 (N_t - N_t^{\ *}) + \sum_{j=1}^{6} \beta_j (D_{jt} - D_{7t}) + e_t \qquad (4)$$

où $\beta_0 = \overline{\alpha}$ et $\beta_j = \alpha_j - \overline{\alpha}$ pour $1 \le j \le 6$

Dans X11 et X11-ARIMA88, le premier régresseur $N_t - N^*{}_t$ est omis et on n'estime alors que 6 paramètres.

- ***Estimation des paramètres***

Dans le cas de X11 et X11-ARIMA88 le modèle peut donc s'écrire, que le schéma soit additif ou multiplicatif :

$$Y_t = \sum_{j=1}^{6} \beta_j (D_{jt} - D_{7t}) + e_t = \sum_{j=1}^{6} \beta_j Z_{jt} + e_t$$

La résolution par moindres carrés ordinaires[27] conduit aux résultats suivants, en notant $\hat{\sigma}^2 = \dfrac{\hat{e}'*\hat{e}}{n-6}$ où les $\hat{e}$ sont les résidus de la régression :

Paramètre	Variance
$\hat{\beta} = [Z'Z]^{-1} Z'Y$	$\hat{s}^2(\hat{\beta}_j) = \hat{\sigma}^2 [Z'Z]^{-1}_{jj}$
$\beta_7 = -\sum_{j=1}^{6} \beta_j$	$\hat{s}^2(\hat{\beta}_j) = \hat{\sigma}^2 \sum_{i=1}^{6} \sum_{j=1}^{6} [Z'Z]^{-1}_{ij}$

Les tests de Student, pour tester la nullité d'un coefficient, et de Fisher, pour tester l'existence d'un effet global dû aux jours ouvrables sont aussi disponibles :

Hypothèse	Statistique	Loi suivie
$\beta_j = k$	$t_j = \dfrac{\hat{\beta}_j - k}{\hat{s}(\hat{\beta}_j)}$	Loi de Student à n-6 degrés de liberté
$\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$	$F = \dfrac{\hat{\beta}' \hat{Z}' \hat{Z} \hat{\beta}}{6 \hat{\sigma}^2}$	Loi de Fisher à 6 et n-6 degrés de liberté

[27] L'utilisation des MCO et des tests associés repose sur l'hypothèse supplémentaire que la composante irrégulière que l'on cherche à expliquer, ne présente que peu ou pas d'autocorrélation.

4.1.14 Tableau B14 : Valeurs de la composante irrégulière exclues de la régression pour Jours Ouvrables

4.1.14.1 Description et mode de calcul

Si l'utilisateur le demande, le programme peut estimer dans la série un effet dû à la composition journalière du mois. Cet effet de calendrier va être recherché dans la composante irrégulière et évalué par régression linéaire. Auparavant, X11 va repérer les valeurs extrêmes de la composante irrégulière et les exclure des calculs, rendant ainsi plus robustes les résultats de la régression. La recherche se fait en trois phases :

- ***Étape 1 : Calcul de la moyenne de l'irrégulier par type de mois***.

On distingue ici 15 types différents de mois :

- les mois de 31 jours commençant un lundi, un mardi, un mercredi, ..., soit donc 7 catégories ;
- les mois de 30 jours commençant un lundi, un mardi, un mercredi, ..., soit 7 autres catégories ;
- les mois de février de 28 jours.

Les valeurs de la composante irrégulière sont ainsi regroupées en 15 classes pour lesquelles on calcule la moyenne m_i. On a donc N données qui se répartissent en 15 groupes d'effectif n_i, $(i = 1 \ldots 15)$ avec :

$$\sum_{i=1}^{15} n_i = N \text{ et } m_i = \frac{1}{n_i} \sum_{j=1}^{n_i} I_{ij}$$

- ***Étape 2 : Premier calcul d'un écart-type global et repérage des valeurs atypiques***

Dans chaque classe, on calcule les carrés des écarts à la moyenne de la classe, dont la moyenne fournit une estimation de l'écart-type global : $\sigma^2 = \frac{1}{N} \sum_{i=1}^{15} \sum_{j=1}^{n_i} \left(I_{ij} - m_i\right)^2$.

Une valeur I_{ij} de l'irrégulier est jugée atypique si elle s'éloigne trop de la moyenne m_i de la classe à laquelle elle appartient, en d'autres termes si elle vérifie : $\left|I_{ij} - m_i\right| \geq \lambda * \sigma$ où σ est l'écart-type global calculé ci-dessus et λ un paramètre modifiable par l'utilisateur et fixé par défaut à 2.5.

- ***Étape 3 : Calcul final de l'écart-type global et repérage des valeurs atypiques***

Les deux étapes sont recommencées en excluant les valeurs atypiques repérées dans la première itération : calcul des moyennes par classes et estimation de l'écart-type. Les valeurs atypiques ainsi repérées et qui seront exclues de la régression pour jours ouvrables sont indiquées dans le tableau B14.

4.1.14.2 Remarques

- Du point de vue théorique, on se place ici dans le cadre d'un modèle d'analyse de la variance à un facteur et on cherche à évaluer "l'effet jours ouvrables" de chacun des 15 types de mois distingués. On suppose donc que, dans chaque classe *i* l'irrégulier du tableau B13 suit une loi normale de moyenne variable m_i et d'écart-type constant σ. Dans ce cas, l'estimation faite à l'étape 2 est une estimation biaisée de l'écart-type inconnu mais constant σ. Il faudrait en fait diviser par N-15 pour obtenir une estimation sans biais.

- Pour les mois de février d'années bissextiles, on ne calcule pas de moyenne de classe et celle-ci est prise égale à _xbar (0 pour un schéma additif, 1 pour un schéma multiplicatif).
- Dans l'étape 3, les valeurs détectées comme atypiques lors de l'étape précédente sont comparées à la moyenne théorique _xbar.
- Toutes les observations, hormis les valeurs détectées comme atypiques, sont utilisées pour la régression.

4.1.15 Exemple

- ***Étape 1 : Calcul de la moyenne de l'irrégulier par type de mois***.

On classe tout d'abord les mois selon leur nombre de jour et le jour correspondant au 1er du mois, en distinguant 15 classes :

B14a : Répartition des mois par type

Longueur	1er du mois	Mois concernés	Moyenne
28		FEB86 FEB87 FEB89 FEB90 FEB91 FEB93 FEB94 FEB95	99.590
	Dimanche	JUN86 NOV87 APR90 SEP91 NOV92	98.879
	Lundi	SEP86 JUN87 APR91 JUN92 NOV93	100.644
	Mardi	APR86 SEP87 NOV88 SEP92 JUN93 NOV94	101.967
30	Mercredi	APR87 JUN88 NOV89 APR92 SEP93 JUN94	101.436
	Jeudi	SEP88 JUN89 NOV90 APR93 SEP94	101.197
	Vendredi	NOV85 APR88 SEP89 JUN90 NOV91 APR94	98.646
	Samedi	NOV86 APR89 SEP90 JUN91	98.320
	Dimanche	DEC85 MAR87 MAY88 JAN89 OCT89 JUL90 DEC91 MAR92 AUG93 MAY94 JAN95	100.277
	Lundi	DEC86 AUG88 MAY89 JAN90 OCT90 JUL91 MAR93 AUG94	101.691
	Mardi	OCT85 JUL86 DEC87 MAR88 AUG89 MAY90 JAN91 OCT91 DEC92 MAR94	101.493
31	Mercredi	JAN86 OCT86 JUL87 MAR89 AUG90 MAY91 JAN92 JUL92 DEC93 MAR95	101.303
	Jeudi	MAY86 JAN87 OCT87 DEC88 MAR90 AUG91 OCT92 JUL93 DEC94	99.703
	Vendredi	AUG86 MAY87 JAN88 JUL88 DEC89 MAR91 MAY92 JAN93 OCT93 JUL94	97.310
	Samedi	MAR86 AUG87 OCT88 JUL89 DEC90 AUG92 MAY93 JAN94 OCT94	97.188

Pour chaque groupe, on calcule la moyenne des valeurs de l'irrégulier B13, moyennes figurant dans la dernière colonne du tableau ci-dessus. Ainsi, pour les mois de février à 28 jours, on a :

$$m_{Fév} = \frac{98.213 + 101.153 + 99.859 + 98.700 + 99.489 + 101.568 + 99.149 + 98.592}{8} = 99.590$$

- ***Étape 2 : Premier calcul d'un écart-type global et repérage des valeurs atypiques***

On calcule ensuite le carré de l'écart de chaque valeur de l'irrégulier à la moyenne de la classe à laquelle il appartient. Ainsi, aux erreurs d'arrondis près, on a :

$JUN86 = (98.460 - 98.879)^2 \approx 0.175$ et $FEB88 = (102.248 - 100)^2 \approx 5.052$.

La moyenne de ces valeurs (voir tableau B14c) fourni la première estimation de l'écart-type.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{15}\sum_{j=1}^{n_i}\left(I_{ij} - m_i\right)^2} = 1.2389$$

Cet écart-type va servir à déterminer la borne à partir de laquelle un point de l'irrégulier sera considéré comme atypique. Cette borne est, par défaut égale à 2.5 fois la valeur de l'écart-type, soit 3.097.

B14b : Écarts à la moyenne en valeur absolue

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	0.958	1.825	2.361
1986	1.825	1.378	1.798	5.391	0.055	0.419	1.273	2.572	0.459	0.875	0.341	0.505
1987	3.829	1.563	1.310	0.860	1.582	0.547	0.977	1.101	1.194	1.162	0.903	0.171
1988	0.413	2.248	0.005	0.444	0.853	0.128	0.150	0.645	0.802	2.134	1.594	1.944
1989	0.751	0.269	2.550	1.593	1.761	1.764	0.781	0.599	0.054	1.131	0.838	0.902
1990	0.234	0.890	0.546	0.148	0.951	0.903	0.634	1.123	1.448	0.188	0.063	0.774
1991	1.567	0.102	0.143	1.355	0.706	0.486	1.837	0.154	0.768	0.152	0.105	1.283
1992	0.416	0.985	1.216	0.002	0.568	0.307	0.070	1.287	0.330	1.582	0.137	1.378
1993	0.752	1.978	0.518	1.808	1.589	0.759	0.445	0.840	1.107	0.605	0.041	0.016
1994	1.528	0.441	1.184	0.319	1.588	0.721	0.074	2.015	0.821	0.623	1.515	0.684
1995	0.015	0.999	1.075	.	.	.	.	.	.	.	.	.

B14c : Carrés des écarts à la moyenne

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	0.917	3.332	5.576
1986	3.330	1.898	3.231	29.061	0.003	0.175	1.620	6.617	0.211	0.766	0.116	0.255
1987	14.661	2.442	1.717	0.740	2.504	0.300	0.955	1.212	1.425	1.350	0.815	0.029
1988	0.171	5.052	0.000	0.197	0.727	0.016	0.022	0.415	0.644	4.555	2.540	3.779
1989	0.564	0.072	6.501	2.536	3.101	3.112	0.609	0.359	0.003	1.280	0.702	0.813
1990	0.055	0.792	0.298	0.022	0.905	0.815	0.403	1.260	2.095	0.035	0.004	0.600
1991	2.456	0.010	0.021	1.835	0.498	0.236	3.374	0.024	0.590	0.023	0.011	1.646
1992	0.173	0.971	1.478	0.000	0.323	0.094	0.005	1.656	0.109	2.504	0.019	1.899
1993	0.565	3.912	0.268	3.267	2.524	0.575	0.198	0.706	1.225	0.366	0.002	0.000
1994	2.334	0.195	1.401	0.102	2.522	0.520	0.005	4.060	0.675	0.389	2.296	0.468
1995	0.000	0.998	1.156	.	.	.	.	.	.	.	.	.

Les deux seuls points qui s'éloignent, en valeur absolue, de plus de 3.097 (soit 2.5σ) sont les valeurs de avril 1986 (mois de 30 jours commençant un mardi) et janvier 1987 (mois de 31 jours commençant un jeudi). Ces deux points sont donc exclus et les calculs de ces deux premières étapes refaits.

- ***Étape 3 : Calcul final de l'écart-type global et repérage des valeurs atypiques***

B14d : Répartition des valeurs de l'irrégulier non exclues par type de mois

Longueur	1er du mois	Mois concernés	Moyenne
28		FEB86 FEB87 FEB89 FEB90 FEB91 FEB93 FEB94 FEB95	99.590
30	Dimanche	JUN86 NOV87 APR90 SEP91 NOV92	98.879
	Lundi	SEP86 JUN87 APR91 JUN92 NOV93	100.644
	Mardi	SEP87 NOV88 SEP92 JUN93 NOV94	100.889
	Mercredi	APR87 JUN88 NOV89 APR92 SEP93 JUN94	101.436
	Jeudi	SEP88 JUN89 NOV90 APR93 SEP94	101.197
	Vendredi	NOV85 APR88 SEP89 JUN90 NOV91 APR94	98.646
	Samedi	NOV86 APR89 SEP90 JUN91	98.320
31	Dimanche	DEC85 MAR87 MAY88 JAN89 OCT89 JUL90 DEC91 MAR92 AUG93 MAY94 JAN95	100.277
	Lundi	DEC86 AUG88 MAY89 JAN90 OCT90 JUL91 MAR93 AUG94	101.691
	Mardi	OCT85 JUL86 DEC87 MAR88 AUG89 MAY90 JAN91 OCT91 DEC92 MAR94	101.493
	Mercredi	JAN86 OCT86 JUL87 MAR89 AUG90 MAY91 JAN92 JUL92 DEC93 MAR95	101.303
	Jeudi	MAY86 OCT87 DEC88 MAR90 AUG91 OCT92 JUL93 DEC94	100.182
	Vendredi	AUG86 MAY87 JAN88 JUL88 DEC89 MAR91 MAY92 JAN93 OCT93 JUL94	97.310
	Samedi	MAR86 AUG87 OCT88 JUL89 DEC90 AUG92 MAY93 JAN94 OCT94	97.188

Seules les moyennes des types de mois concernés par les valeurs exclues sont affectées et le calcul des carrés des écarts à la moyenne conduit à la nouvelle estimation de l'écart-type:

$$\sigma = \sqrt{\frac{1}{N'} \sum_{i=1}^{15} \sum_{j=1}^{n_i} \left(I_{ij} - m_i\right)^2} = 1.0505$$

Ainsi, la nouvelle moyenne pour les mois de 30 jours commençant un mardi est 100.889 au lieu de 101.967 et le carré de l'écart à cette nouvelle moyenne pour septembre 1987 sera de $(100.774 - 100.889)^2 = (0.115)^2 = 0.01323$ au lieu de $(100.774 - 100.967)^2 = (0.193)^2 = 0.03725$.

B14e : Écarts à la moyenne en valeur absolue

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	0.958	1.825	2.361
1986	1.825	1.378	1.798	7.358	0.534	0.419	1.273	2.572	0.459	0.875	0.341	0.505
1987	4.126	1.563	1.310	0.860	1.582	0.547	0.977	1.101	0.115	1.641	0.903	0.171
1988	0.413	2.248	0.005	0.444	0.853	0.128	0.150	0.645	0.802	2.134	0.515	1.465
1989	0.751	0.269	2.550	1.593	1.761	1.764	0.781	0.599	0.054	1.131	0.838	0.902
1990	0.234	0.890	0.067	0.148	0.951	0.903	0.634	1.123	1.448	0.188	0.063	0.774
1991	1.567	0.102	0.143	1.355	0.706	0.486	1.837	0.633	0.768	0.152	0.105	1.283
1992	0.416	0.985	1.216	0.002	0.568	0.307	0.070	1.287	0.748	1.104	0.137	1.378
1993	0.752	1.978	0.518	1.808	1.589	0.320	0.033	0.840	1.107	0.605	0.041	0.016
1994	1.528	0.441	1.184	0.319	1.588	0.721	0.074	2.015	0.821	0.623	0.437	0.205
1995	0.015	0.999	1.075	.	.	.	.	.	.	.	.	.

Notez enfin que dans le tableau B14e, les points considérés comme extrêmes dans l’étape précédente sont comparés à la moyenne théorique (ici 100 pour tenir compte du fait que l’irrégulier est multiplié par 100). Ainsi, l’écart absolu pour janvier 1987 est : $|95.874 - 100| = 4.126$

B14f : Carrés des écarts à la moyenne

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	0.917	3.332	5.576
1986	3.330	1.898	3.231	54.144	0.285	0.175	1.620	6.617	0.211	0.766	0.116	0.255
1987	17.022	2.442	1.717	0.740	2.504	0.300	0.955	1.212	0.013	2.692	0.815	0.029
1988	0.171	5.052	0.000	0.197	0.727	0.016	0.022	0.415	0.644	4.555	0.266	2.147
1989	0.564	0.072	6.501	2.536	3.101	3.112	0.609	0.359	0.003	1.280	0.702	0.813
1990	0.055	0.792	0.004	0.022	0.905	0.815	0.403	1.260	2.095	0.035	0.004	0.600
1991	2.456	0.010	0.021	1.835	0.498	0.236	3.374	0.401	0.590	0.023	0.011	1.646
1992	0.173	0.971	1.478	0.000	0.323	0.094	0.005	1.656	0.560	1.218	0.019	1.899
1993	0.565	3.912	0.268	3.267	2.524	0.102	0.001	0.706	1.225	0.366	0.002	0.000
1994	2.334	0.195	1.401	0.102	2.522	0.520	0.005	4.060	0.675	0.389	0.191	0.042
1995	0.000	0.998	1.156	.	.	.	.	.	.	.	.	.

La nouvelle borne pour juger du caractère atypique d'un point est : 2.5*1.0505 = 2.626 et les seules valeurs de l'irrégulier s'éloignant suffisamment de leur moyenne sont celles de Avril 1986 et Janvier 1987. Ces valeurs sont consignées dans le tableau B14 :

B14 : VALEURS DE L'IRRÉGULIER EXCLUES DE LA RÉGRESSION POUR JOURS OUVRABLES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	107.358	.	.	.	.	.	.	.	.
1987	95.874	.	.	.	.	.	.	.	.	.	.	.
1988	.	.	.	.	.	.	.	.	.	.	.	.
1989	.	.	.	.	.	.	.	.	.	.	.	.
1990	.	.	.	.	.	.	.	.	.	.	.	.
1991	.	.	.	.	.	.	.	.	.	.	.	.
1992	.	.	.	.	.	.	.	.	.	.	.	.
1993	.	.	.	.	.	.	.	.	.	.	.	.
1994	.	.	.	.	.	.	.	.	.	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

4.1.16 Tableau B15 : Régression préliminaire pour Jours Ouvrables

4.1.16.1 Description et mode de calcul

Des poids journaliers vont être à présent estimés grâce à une régression par moindres carrés ordinaires sur les données non jugées atypiques du tableau B13 selon la méthodologie exposée au paragraphe 4.1.13.

- ***Écriture du modèle de régression***

On a, selon le schéma et avec les notations du paragraphe 4.1.13, le modèle suivant :

- Pour un schéma multiplicatif : $N_t^* \hat{I}_t - N_t = \sum_{j=1}^{6} \beta_j (D_{jt} - D_{7t}) + e_t$
- Pour un schéma additif : $\hat{I}_t = \sum_{j=1}^{6} \beta_j (D_{jt} - D_{7t}) + e_t$

- ***Dérivation des poids journaliers***

Dans le cas additif, les poids qui seront utilisés dans la suite de la désaisonnalisation sont les $\hat{\beta}_j$ calculés ci-dessus.
Dans le cas multiplicatif, on ajoute 1 aux estimations ci-dessus ou bien les poids journaliers *a priori* s'ils existent.

4.1.16.2 Remarques

- Dans le cas additif, le logiciel X12-ARIMA introduit une variable explicative supplémentaire et le modèle estimé est alors $\hat{I}_t = \beta_0 (N_t - N_t^*) + \sum_{j=1}^{6} \beta_j (D_{jt} - D_{7t}) + e_t$. Le principe de l'estimation reste cependant le même.
- Dans le cas multiplicatif, si un ajustement journalier *a priori* est fait, alors N_t et $N^*{}_t$ sont égaux à la somme mensuelle des poids journaliers *a priori*.
- X12-ARIMA permet d'introduire dans la régression d'autres variables prédéfinies ou définies par l'utilisateur. Parmi les variables prédéfinies, on peut citer l'effet de Pâques (voir chapitre 5), des variables indicatrices modélisant des points aberrants, des variables estimant les effets de fêtes américaines particulières (Thanksgiving, Labor Day).

4.1.16.3 Exemple

B15 : RÉGRESSION POUR JOURS OUVRABLES

	Poids Combinés	Poids a Priori	Coefficients de la régression	Std	Student	Prob >t
Lundi	1.081	1.000	0.081	0.093	0.872	0.192
Mardi	1.273	1.000	0.273	0.091	2.990	0.002
Mercredi	1.047	1.000	0.047	0.095	0.494	0.311
Jeudi	1.319	1.000	0.319	0.095	3.362	0.001
Vendredi	1.066	1.000	0.066	0.092	0.717	0.237
Samedi	0.565	1.000	-0.435	0.091	-4.772	0.000
Dimanche	0.649	1.000	-0.351	0.093	-3.760	0.000

	Somme des Carrés	D.D.L	Moyenne des Carrés	F-Value	Prob > F
RÉGRESSION	23.436	6	3.906	31.257	0.000
ERREUR	13.246	106	0.125		
TOTAL	36.682	112			

Il serait assez fastidieux de présenter ici l'ensemble des calculs intermédiaires de la régression aboutissant au tableau B15 et nous ne le ferons pas : les méthodes usuelles de régression par moindres carrés sont utilisées.

Si nous nous fixons un risque de première espèce de 1 % par exemple, les tests s'interprètent de la façon suivante :

- Le test de Fisher rejette l'hypothèse nulle d'égalité des coefficients journaliers ; on peut donc admettre l'existence d'un effet dû à la composition journalière du mois. En effet, la probabilité de trouver une valeur de la statistique de Fisher plus grande que celle trouvée (31.257) est quasiment nulle et donc inférieure à notre risque de première espèce. Nous sommes donc dans la région critique du test ce qui ne nous permet pas d'accepter l'hypothèse nulle d'égalité des coefficients journaliers.
- On interprète de la même manière les tests de Student mais en prenant garde à ce que la loi de Student est symétrique. Il faut donc comparer la valeur (Prob > |t|) à la moitié du risque de première espèce soit 0.005. Tous les tests conduisant à une valeur plus petite que 0.005 ne permettent pas d'accepter l'hypothèse de nullité du coefficient journalier. Dans notre cas, les coefficients de mardi, jeudi, samedi et dimanche sont jugés significativement différents de 0.

4.1.17 Tableau B16 : Coefficients d’ajustement pour jours ouvrables issus de la régression

4.1.17.1 Description et mode de calcul

Des estimations de la régression, on déduit des coefficients mensuels M_t d'ajustement pour jours ouvrables, directement en utilisant les équations (2) et (4) du paragraphe 4.1.13:

- Pour un schéma multiplicatif : $M_t = \frac{1}{N^*_t} \sum_{j=1}^{7} (\beta_j + 1) D_{jt}$
- Pour un schéma additif : $M_t = \sum_{j=1}^{7} \beta_j D_{jt}$

Où N^*_t est égal au nombre de jours du mois si des coefficients d'ajustement a priori ont été fournis et, dans le cas contraire, à 31, 30 ou 28.25 selon que le mois a 31 ou 30 jours ou est un mois de Février.

La composante irrégulière du tableau B13 est alors corrigée de ces effets de calendrier pour aboutir à un tableau "B16bis", malheureusement non imprimable dans les logiciels usuels :
B16bis = B13 _op B16

4.1.17.2 Exemple

B16 : COEFFICIENTS DE CORRECTION POUR JOURS OUVRABLES ISSUS DE LA RÉGRESSION

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	102.061	98.772	100.009
1986	101.393	99.115	97.726	101.067	99.840	99.099	102.061	97.678	101.180	101.393	97.380	101.294
1987	99.840	99.115	100.009	101.219	97.678	101.180	101.393	97.726	101.067	99.840	99.099	102.061
1988	97.678	102.941	102.061	98.772	100.009	101.219	97.678	101.294	101.283	97.726	101.067	99.840
1989	100.009	99.115	101.393	97.380	101.294	101.283	97.726	102.061	98.772	100.009	101.219	97.678
1990	101.294	99.115	99.840	99.099	102.061	98.772	100.009	101.393	97.380	101.294	101.283	97.726
1991	102.061	99.115	97.678	101.180	101.393	97.380	101.294	99.840	99.099	102.061	98.772	100.009
1992	101.393	101.116	100.009	101.219	97.678	101.180	101.393	97.726	101.067	99.840	99.099	102.061
1993	97.678	99.115	101.294	101.283	97.726	101.067	99.840	100.009	101.219	97.678	101.180	101.393
1994	97.726	99.115	102.061	98.772	100.009	101.219	97.678	101.294	101.283	97.726	101.067	99.840
1995	100.009	99.115	101.393	.	.	.	.	.	.	.	.	.

Prenons comme exemple le mois d’avril 1986.

	Lundi	Mardi	Mercredi	Jeudi	Vendredi	Samedi	Dimanche	Nb jours
poids	1.0809	1.2732	1.0469	1.3187	1.0663	0.5653	0.6487	
Nombre de APR86	4	5	5	4	4	4	4	30

On a donc :

$$APR86 = \frac{100*(4*1.0809+5*1.2732+5*1.0469+4*1.3187+4*1.0663+4*0.5653+4*0.6487)}{30} = 101.067$$

Ou encore, comme les jours apparaissant 5 fois sont le Mardi et le Mercredi :

$$APR86 = 100*\frac{28+1.2732+1.0469}{30} = 101.067$$

On obtient ensuite une valeur corrigée de la composante irrégulière (tableau B16bis ci-après) :

$$APR86 = 100 * (107.358 / 101.067) = 106.225$$

B16bis : IRRÉGULIER CORRIGÉ DES EFFETS JOURS OUVRABLES ISSUS DE LA RÉGRESSION

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	98.505	101.721	97.907
1986	101.710	99.090	97.610	106.225	99.808	99.356	100.691	96.989	99.016	100.774	101.316	99.894
1987	96.028	102.056	101.578	101.064	101.243	98.929	100.874	98.323	99.710	98.699	100.689	99.612
1988	99.200	99.326	99.448	99.423	101.121	100.341	99.776	99.756	100.707	97.265	99.314	101.810
1989	101.019	100.751	97.396	102.601	98.654	101.657	100.248	98.856	99.818	99.137	101.042	100.546
1990	100.161	99.582	100.410	99.927	100.376	98.958	100.902	101.018	99.479	100.207	99.977	98.657
1991	100.979	100.377	99.476	100.809	99.215	100.467	102.206	99.709	99.003	99.592	99.766	98.985
1992	99.501	99.870	99.053	100.216	100.204	99.167	99.842	100.766	100.564	101.448	99.916	98.093
1993	98.853	102.475	99.881	98.130	101.075	100.140	100.309	101.108	99.121	100.242	99.430	99.926
1994	101.013	100.034	98.284	99.550	101.856	99.502	99.699	102.382	99.104	100.087	99.392	100.548
1995	100.283	99.472	98.850	.	.	.	.	.	.	.	.	.

4.1.18 Tableau B17 : Poids préliminaires pour la correction de l'irrégulier

4.1.18.1 Description et mode de calcul

A partir de l'estimation de la composante irrégulière B16bis, ou du tableau B13 si aucune correction pour jours ouvrables n'est demandée, on cherche à identifier et à corriger les points atypiques. Pour cela, on utilise l'algorithme de détection des points atypiques et de calcul de poids correctifs détaillé aux tableaux B4 et B9 (voir respectivement 4.1.4 et 4.1.9). Comme on dispose déjà d'une estimation de l'irrégulier, seules les étapes 4 et 5 sont appliquées.

4.1.18.2 Remarques

Les remarques faites aux tableaux B4 et B9 concernant le calcul des écart-types mobiles et des poids correctifs sont aussi valables ici.

4.1.18.3 Exemple

- ***Calcul d'un écart-type mobile***

L'écart-type correspondant à l'année 1989 sera calculé à partir des données des années 1987 à 1991 (deux années avant, deux années après) selon la formule [28]:

$$\sigma_{89} = \sqrt{\frac{1}{60}\sum_{i=1}^{60}(x_i - 100)^2} = 1.198$$

Ceux des années 1988, 1990, 1991 et 1992 sont calculés selon le même principe.
Pour X11-ARIMA88 et X12-ARIMA, l'écart-type de 1987 est calculé à partir de l'ensemble des observations disponibles de 1985 à 1990, soit 63. Les résultats du calcul de X11-ARIMA88 et X12-ARIMA figurent dans le tableau B17a, colonne Écart-type 1.

B17a : VALEURS DES ÉCART-TYPES MOBILES SUR CINQ ANS

ANNÉE	ÉCART-TYPE 1	ÉCART-TYPE 2
1985	1.5282	1.2322
1986	1.5282	1.2322
1987	1.5282	1.2322
1988	1.5142	1.1965
1989	1.1979	1.0918
1990	1.0200	1.0200
1991	1.0173	0.9740
1992	0.9484	0.8527
1993	0.9399	0.8479
1994	0.9399	0.8479
1995	0.9399	0.8479

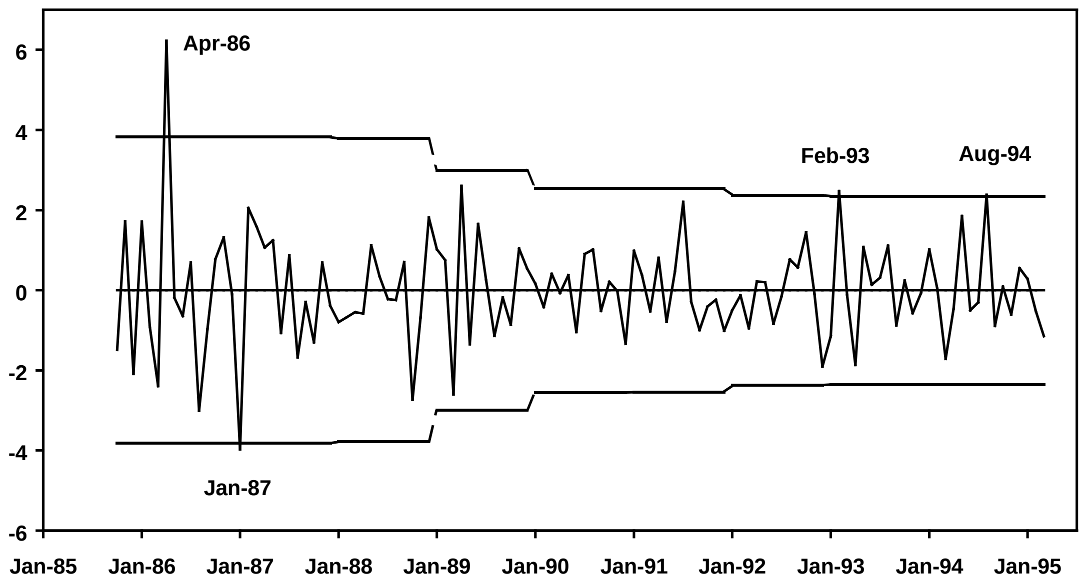
Ce premier calcul sert à repérer d'éventuels points atypiques. Comme le montre la figure 15, les valeurs de avril 1986, janvier 1987, février 1993 et août 1994 sont détectées comme très atypiques.
On a en effet par exemple :

$$|APR86 - 100| = |106.225 - 100| = 6.225 > 2.5 * \sigma_{86} = 2.5 * 1.5282 = 3.8205 \text{ et}$$

$$|AUG94 - 100| = |102.382 - 100| = 2.382 > 2.5 * \sigma_{94} = 2.5 * 0.94 = 2.35$$

[28] où la moyenne théorique est considérée égale à 100 pour tenir compte du fait que les valeurs de l'irrégulier ont été elles-mêmes multipliées par 100

Figure 15 : Écart à la moyenne de l'irrégulier et "limites supérieures de confiance" (tableau B17)



Ces points sont donc éliminés du second calcul de l'écart-type mobile qui conduit aux résultats de la colonne "Écart-type 2" du tableau B17a.

- ***Détection et correction des valeurs atypiques***

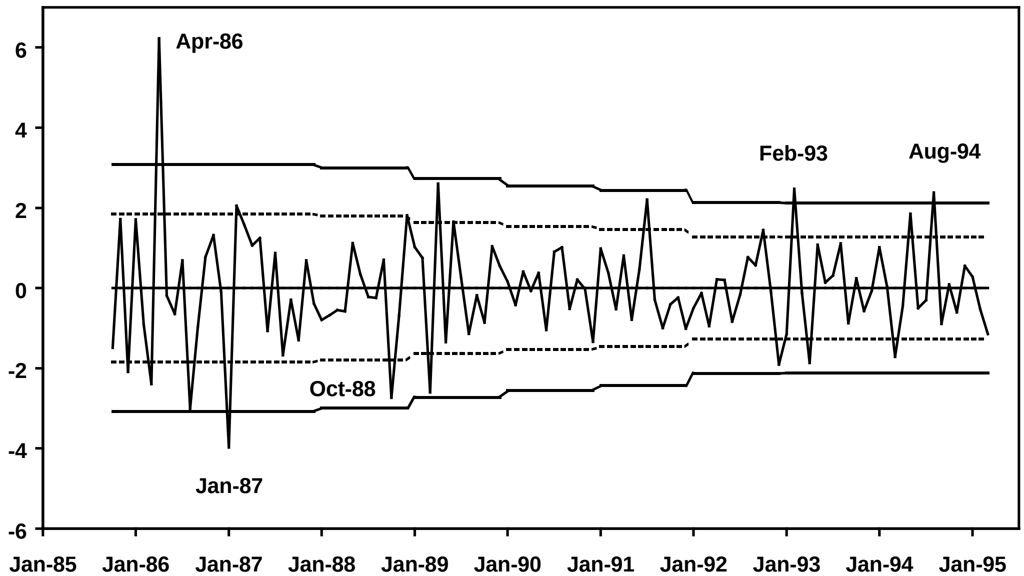
Les valeurs de l'irrégulier sont alors situées par rapport aux limites de confiance supérieures et inférieures calculées à partir des nouvelles estimations des écart-types. Toutes les valeurs situées au delà des limites de confiance inférieures (voir figure 16) sont considérées comme atypiques et vont donc être corrigées, à des degrés divers, et les poids associés à chacune de ces valeurs figurent dans le tableau B17.

B17 : POIDS PRÉLIMINAIRES POUR LA CORRECTION DE L'IRRÉGULIER

BORNES RETENUES : DE 1.5 A 2.5 SIGMA

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.000	100.000	80.138
1986	100.000	100.000	56.025	0.000	100.000	100.000	100.000	5.658	100.000	100.000	100.000	100.000
1987	0.000	83.133	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	21.455	100.000	98.701
1989	100.000	100.000	11.498	11.770	100.000	98.258	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1991	100.000	100.000	100.000	100.000	100.000	100.000	23.528	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	80.157	100.000	26.391
1993	100.000	0.000	100.000	29.466	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1994	100.000	100.000	47.606	100.000	31.094	100.000	100.000	0.000	100.000	100.000	100.000	100.000
1995	100.000	100.000	100.000	.	.	.	.	.	.	.	.	.

Figure 16 : Écart à la moyenne de l'irrégulier et "limites de confiance" (tableau B17)



Les valeurs jugées précédemment très atypiques le restent et sont affectées d'un poids nul. Ainsi :

$$\left|APR86-100\right| = \left|106.225-100\right| = 6.225 > 2.5*\sigma_{86} = 2.5*1.2322 = 3.0805$$

Pour octobre 1988, situé entre les deux limites de confiance et donc jugé moyennement atypique on a :

$$\left|OCT88-100\right| = \left|97.265-100\right| = 2.735$$

$$1.5*\sigma_{88} = 1.5*1.1965 = 1.79475 < 2.735 < 2.5*\sigma_{88} = 2.5*1.1965 = 2.99125$$

On va attribuer à cette valeur jugée moyennement atypique, un poids proportionnel à l'écart à la moyenne constaté de :

$$poids(OCT88) = \frac{2.99125-2.735}{2.99125-1.79475} = 0.214$$

4.1.19 Tableau B18 : Coefficients pour jours ouvrables combinés (issus de l'ajustement *a priori* et de la régression pour jours ouvrables)

4.1.19.1 Description et mode de calcul

Si on a fourni au préalable des coefficients journaliers de correction a priori des effets de jours ouvrables (schéma multiplicatif seulement) et si on a aussi demandé une régression pour jours ouvrables, le tableau B18 présente le résultat combiné de ces deux corrections, simple addition des deux effets. Ces poids journaliers combinés vont permettre d'estimer des coefficients de correction pour chaque mois, de la même façon qu'au tableau B16.

<u>Pour un schéma additif,</u> ce tableau n'est pas édité puisqu'on ne peut utiliser dans ce cas de correction *a priori*.

<u>Pour un schéma multiplicatif,</u> on calcule : $M_t = \frac{1}{N^*_t} \sum_{j=1}^{7} \alpha_j D_{jt}$

où D_{jt} est le nombre de jours j (lundi, mardi, mercredi ...., dimanche) contenu dans le mois i, $(\alpha_1, \alpha_2, \cdots, \alpha_7)$ les poids combinés de chaque jour (colonne "poids combinés" du tableau B15) et où N^*_t est égal au nombre de jours du mois si des coefficients d'ajustement a priori ont été fournis et, dans le cas contraire, à 31, 30 ou 28.25 selon que le mois a 31 ou 30 jours ou est un mois de Février.

4.1.19.2 Remarque

Dans X12-ARIMA, d'autres effets peuvent être estimés au moment de la régression sur la composante irrégulière (Pâques, Labor Day, Thanksgiving …..). Dans ce cas le tableau B18 prend en compte l'ensemble des effets.

4.1.19.3 Exemple

Dans notre cas, le tableau B18 est identique au tableau B16.

B18: COEFFICIENTS COMBINÉS DE CORRECTION POUR EFFET DE JOURS OUVRABLES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	102.061	98.772	100.009
1986	101.393	99.115	97.726	101.067	99.840	99.099	102.061	97.678	101.180	101.393	97.380	101.294
1987	99.840	99.115	100.009	101.219	97.678	101.180	101.393	97.726	101.067	99.840	99.099	102.061
1988	97.678	102.941	102.061	98.772	100.009	101.219	97.678	101.294	101.283	97.726	101.067	99.840
1989	100.009	99.115	101.393	97.380	101.294	101.283	97.726	102.061	98.772	100.009	101.219	97.678
1990	101.294	99.115	99.840	99.099	102.061	98.772	100.009	101.393	97.380	101.294	101.283	97.726
1991	102.061	99.115	97.678	101.180	101.393	97.380	101.294	99.840	99.099	102.061	98.772	100.009
1992	101.393	101.116	100.009	101.219	97.678	101.180	101.393	97.726	101.067	99.840	99.099	102.061
1993	97.678	99.115	101.294	101.283	97.726	101.067	99.840	100.009	101.219	97.678	101.180	101.393
1994	97.726	99.115	102.061	98.772	100.009	101.219	97.678	101.294	101.283	97.726	101.067	99.840
1995	100.009	99.115	101.393	.	.	.	.	.	.	.	.	.

4.1.20 Tableau B19 : Série brute corrigée des effets de jours ouvrables

4.1.20.1 Description et mode de calcul

La série du tableau B1, ou celle du tableau A1 si aucun ajustement préalable n'est demandé, est corrigée par les effets de jours ouvrables estimés précédemment (tableau B18). On a donc :

B19 = B1 _op B18

4.1.20.2 Exemple

B19: SÉRIE BRUTE CORRIGÉE DES EFFETS DE JOURS OUVRABLES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	113.364	111.165	100.591
1986	105.135	99.581	106.318	108.344	97.857	104.643	97.687	67.262	103.973	115.491	111.214	103.067
1987	100.661	104.121	112.890	105.810	102.377	107.037	100.401	70.299	107.552	117.088	115.743	107.779
1988	110.260	107.051	116.303	109.444	107.390	113.319	103.605	75.029	113.148	120.644	120.019	114.884
1989	117.889	113.202	118.548	117.786	109.089	118.776	108.058	77.797	115.620	126.689	125.273	115.379
1990	119.553	113.504	123.799	117.156	113.266	118.252	111.790	82.155	117.683	130.314	125.490	113.379
1991	120.810	113.807	122.136	118.007	111.743	119.840	113.827	81.731	117.459	129.727	126.352	115.790
1992	121.803	115.610	123.989	118.555	112.410	117.315	110.560	81.862	118.040	129.207	123.211	111.502
1993	116.403	114.110	121.133	112.753	110.411	115.864	108.274	79.693	113.418	123.876	120.280	113.223
1994	119.007	112.496	121.496	116.835	113.990	119.543	112.103	84.309	119.072	129.342	126.352	120.193
1995	124.089	117.338	128.411	.	.	.	.	.	.	.	.	.

On a donc par exemple : $APR86 \;= 100 * (109.500 \,/\, 101.067) \;=\; 108.344$

4.1.21 Tableau B20 : Valeurs de correction des points atypiques de l'irrégulier

4.1.21.1 Description et mode de calcul

Les valeurs de la composante irrégulière B16bis, ou B13 si on ne demande pas de régression pour jours ouvrables, détectées comme atypiques lors de la constitution du tableau B17, et pour lesquelles on a donc calculé un poids, sont corrigées de la façon suivante :

- Schéma additif : *B20 = B16bis * (1 - B17)*
- Schéma multiplicatif : *B20 = B16bis / [1 + B17 * (B16bis – 1)]*

Ou, en notation symbolique : *B20 = B16bis _op [_xbar + B17 * (B16bis – _xbar)]*

4.1.21.2 Remarques

Il s'agit bien des valeurs qui serviront à corriger la série initiale. Un point jugé très aberrant recevra donc un poids de 0 et une valeur de correction égale à la valeur de l'irrégulier ; en d'autres termes, dans ce cas, on enlèvera de la série initiale l'irrégulier correspondant à cette date (voir tableau C1).

4.1.21.3 Exemple

La valeur de octobre 1988, détectée comme atypique et affectée d'un poids égal à 0.21455 sera donc corrigée de la façon suivante, puisque nous avons ici un schéma multiplicatif :

$$OCT88 = 100 * \frac{0.97265}{1 + 0.21455 * (0.97265 - 1)} = 97.839$$

B20:VALEURS DE CORRECTION POUR LES POINTS JUGÉS ATYPIQUES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.000	100.000	99.577
1986	100.000	100.000	98.935	106.225	100.000	100.000	100.000	97.155	100.000	100.000	100.000	100.000
1987	96.028	100.341	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	97.839	100.000	100.023
1989	100.000	100.000	97.689	102.288	100.000	100.028	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1991	100.000	100.000	100.000	100.000	100.000	100.000	101.678	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.284	100.000	98.589
1993	100.000	102.475	100.000	98.674	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1994	100.000	100.000	99.093	100.000	101.272	100.000	100.000	102.382	100.000	100.000	100.000	100.000
1995	100.000	100.000	100.000	.	.	.	.	.	.	.	.	.

4.2 PARTIE C : Estimation finale des points atypiques et des effets de calendrier

4.2.1 Tableau C1 : Série brute ajustée pour tenir compte des ajustements *a priori*, de ceux liés à la correction pour jours ouvrables et des points atypiques détectés

4.2.1.1 Description et mode de calcul

Ce tableau présente la série brute corrigée des divers effets mis en évidence lors de la partie B, points jugés atypiques et effets liés aux jours ouvrables, et ajustée *a priori* par les éléments de la partie A.
Il est donc calculé à partir du tableau B19 qui prend en compte les effets dus aux jours ouvrables, ou du tableau B1 si la régression pour jours ouvrables n'est pas demandée, et du tableau B20 qui précise les corrections à apporter aux points jugés atypiques.
On a donc : *C1 = B19 _op B20*

4.2.1.2 Exemple

C1: SÉRIE ORIGINALE AJUSTÉE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	113.364	111.165	101.018
1986	105.135	99.581	107.463	101.995	97.857	104.643	97.687	69.231	103.973	115.491	111.214	103.067
1987	104.825	103.768	112.890	105.810	102.377	107.037	100.401	70.299	107.552	117.088	115.743	107.779
1988	110.260	107.051	116.303	109.444	107.390	113.319	103.605	75.029	113.148	123.308	120.019	114.858
1989	117.889	113.202	121.354	115.151	109.089	118.742	108.058	77.797	115.620	126.689	125.273	115.379
1990	119.553	113.504	123.799	117.156	113.266	118.252	111.790	82.155	117.683	130.314	125.490	113.379
1991	120.810	113.807	122.136	118.007	111.743	119.840	111.949	81.731	117.459	129.727	126.352	115.790
1992	121.803	115.610	123.989	118.555	112.410	117.315	110.560	81.862	118.040	128.841	123.211	113.098
1993	116.403	111.354	121.133	114.269	110.411	115.864	108.274	79.693	113.418	123.876	120.280	113.223
1994	119.007	112.496	122.608	116.835	112.558	119.543	112.103	82.348	119.072	129.342	126.352	120.193
1995	124.089	117.338	128.411	.	.	.	.	.	.	.	.	.

Ainsi, la valeur du mois d'avril 1986, jugée atypique, devient :
$APR86 = 100*(108.344/106.225) = 101.995$

4.2.2 Tableau C2 : Estimation préliminaire de la Tendance-Cycle

4.2.2.1 Description et mode de calcul

Une nouvelle estimation de la composante Tendance-Cycle est obtenue en appliquant aux données du tableau C1 une moyenne mobile centrée simple d'ordre 12, tout comme au tableau B2.

4.2.2.2 Remarques

- X11-ARIMA88 et X12-ARIMA proposent aussi une moyenne mobile centrée sur 24 termes (Cholette, 1981).
- Les 6 premiers et les 6 derniers points de la série ne sont pas imputés à ce stade du calcul.

4.2.2.3 Exemple

C2: TENDANCE-CYCLE - MOYENNE MOBILE SIMPLE SUR 12 TERMES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	101.181	101.272	101.359	101.432	101.593	101.994	102.379	102.726	103.014
1987	103.227	103.385	103.578	103.794	104.049	104.434	104.857	105.220	105.500	105.793	106.153	106.624
1988	107.019	107.350	107.780	108.273	108.710	109.183	109.796	110.370	110.837	111.285	111.594	111.890
1989	112.302	112.603	112.821	113.065	113.425	113.665	113.756	113.838	113.953	114.138	114.396	114.549
1990	114.684	115.021	115.289	115.526	115.686	115.612	115.581	115.646	115.589	115.555	115.527	115.530
1991	115.603	115.592	115.565	115.531	115.542	115.679	115.821	115.937	116.089	116.189	116.240	116.163
1992	115.999	115.947	115.977	115.964	115.796	115.553	115.216	114.814	114.517	114.220	113.958	113.814
1993	113.658	113.473	113.190	112.790	112.461	112.344	112.458	112.614	112.723	112.892	113.088	113.331
1994	113.644	113.914	114.260	114.723	115.204	115.748	116.250	116.663	117.107	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

Ainsi la valeur d'avril 1986 s'obtient à partir des valeurs du tableau C1 de octobre 1985 à octobre 1986 (6 mois avant et 6 mois après) :

$$
\begin{aligned}
APR86 &= \frac{113.364}{24} + \\
&\frac{111.165 + 101.018 + 105.135 + 99.581 + 107.463 + 101.995 + 97.857 + 104.643 + 97.687 + 69.231 + 103.973}{12} \\
&+ \frac{115.491}{24} \\
&= 101.181
\end{aligned}
$$

4.2.3 Tableau C4 : Estimation préliminaire de la composante Saisonnier-Irrégulier modifiée

4.2.3.1 Description et mode de calcul

La composante Tendance-Cycle est enlevée de la série analysée pour obtenir une estimation de la composante Saisonnier-Irrégulier.
On a donc : *C4 = C1 _op C2*

4.2.3.2 Remarque

- Là encore, pas d'estimation pour les 6 valeurs de début et les 6 valeurs de fin de la série.
- Ce tableau équivaut au tableau B4g : ici le calcul des valeurs de remplacement pour les points atypiques de la composante Saisonnier-Irrégulier n'a pas lieu car ces valeurs ont déjà été repondérées par les poids du tableau B20.

4.2.3.3 Exemple

C4: COMPOSANTE SAISONNIER-IRRÉGULIER MODIFIÉE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	100.804	96.628	103.240	96.308	68.146	101.940	112.807	108.262	100.051
1987	101.548	100.370	108.990	101.943	98.393	102.492	95.751	66.811	101.946	110.676	109.034	101.083
1988	103.028	99.722	107.908	101.082	98.786	103.788	94.362	67.980	102.085	110.804	107.550	102.652
1989	104.976	100.532	107.563	101.846	96.177	104.467	94.991	68.340	101.463	110.996	109.509	100.724
1990	104.246	98.681	107.381	101.411	97.908	102.284	96.720	71.041	101.812	112.772	108.623	98.138
1991	104.505	98.456	105.686	102.143	96.712	103.597	96.657	70.496	101.180	111.651	108.699	99.679
1992	105.003	99.709	106.908	102.234	97.075	101.525	95.959	71.300	103.076	112.801	108.119	99.370
1993	102.414	98.132	107.018	101.311	98.177	103.132	96.279	70.766	100.616	109.730	106.360	99.905
1994	104.719	98.755	107.306	101.840	97.703	103.279	96.433	70.586	101.678	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

La valeur de avril 1986 s'obtient donc simplement : $APR86 = 100*(101.995/101.181) = 100.804$

4.2.4 Tableau C5 : Estimation de la composante saisonnière

4.2.4.1 Description et mode de calcul

Cette estimation est obtenue à partir des valeurs de la composante Saisonnier-Irrégulier du tableau C4. On procède en deux étapes, similaires aux étapes 1 et 2 vues au tableau B4 :

- ***Étape 1 : estimation de la composante saisonnière*** par une moyenne mobile 3x3.

- ***Étape 2 : normalisation des coefficients saisonniers*** par une moyenne mobile centrée sur 12 termes.

4.2.4.2 Exemple

L'estimation est faite à partir de la composante Saisonnier-Irrégulier modifiée du tableau C4.

- ***Étape 1 : première estimation de la composante saisonnière***.

Les données du tableau précédent sont lissées colonne par colonne (mois par mois), avec une moyenne mobile 3x3 coefficients {1/9,2/9,3/9,2/9,1/9}, pour aboutir au tableau C5a.
Le facteur saisonnier du mois d'avril 1988 sera donc estimé par :

$$APR88 = \frac{100.804 + 101.943 * 2 + 101.082 * 3 + 101.846 * 2 + 101.411}{9} = 101.449$$

Cette moyenne mobile symétrique peut s'appliquer pour estimer les valeurs des coefficients saisonniers des années 1988 à 1992. Pour le début de la série (années 1986 et 1987) et la fin de la série (années 1993 et 1994), on utilise des moyennes asymétriques prédéfinies. Ainsi, par exemple, en avril 1987, on utilise un point dans le passé, le point courant et deux points dans le futur :

$$APR87 = \frac{100.804 * 7 + 101.943 * 10 + 101.082 * 7 + 101.846 * 3}{27} = 101.414$$

C5a: FACTEURS SAISONNIERS PROVISOIRES (mm 3x3)

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	101.319	97.746	103.037	95.720	67.571	101.969	111.568	108.445	100.953
1987	102.786	100.136	108.285	101.414	97.791	103.241	95.451	67.630	101.927	111.297	108.502	101.182
1988	103.284	99.984	108.040	101.449	97.782	103.423	95.288	68.159	101.870	111.260	108.514	101.084
1989	103.947	99.692	107.549	101.623	97.447	103.515	95.505	68.930	101.701	111.385	108.734	100.502
1990	104.414	99.272	107.051	101.725	97.263	103.144	95.975	70.008	101.765	111.913	108.662	99.694
1991	104.378	98.979	106.691	101.875	97.162	102.890	96.289	70.586	101.710	111.870	108.383	99.409
1992	104.201	98.860	106.758	101.874	97.402	102.622	96.322	70.895	101.812	111.704	107.869	99.452
1993	103.915	98.739	106.916	101.780	97.606	102.805	96.278	70.828	101.592	111.337	107.510	99.645
1994	103.833	98.678	107.115	101.698	97.780	102.895	96.282	70.792	101.504	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

- ***Étape 2 : normalisation des coefficients saisonniers***.

C5b: MOYENNE MOBILE CENTRÉE SUR 12 MOIS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.			
1986				99.965	99.965	99.965	99.965	99.965	99.965	99.965	99.971	99.981
1987	99.979	99.970	99.971	99.958	99.949	99.961	99.991	100.005	99.989	99.980	99.981	99.988
1988	99.989	100.004	100.024	100.020	100.019	100.015	100.039	100.054	100.022	100.009	100.002	99.992
1989	100.005	100.046	100.071	100.069	100.084	100.068	100.064	100.066	100.027	100.011	100.007	99.984
1990	99.988	100.053	100.100	100.125	100.144	100.107	100.072	100.059	100.031	100.023	100.025	100.010
1991	100.012	100.049	100.071	100.067	100.054	100.030	100.011	99.999	99.997	99.999	100.009	100.008
1992	99.998	100.013	100.030	100.027	99.999	99.979	99.969	99.952	99.953	99.956	99.961	99.977
1993	99.983	99.978	99.966	99.942	99.911	99.905	99.909	99.903	99.909	99.914	99.918	99.929
1994	99.932	99.931	99.926	99.926	99.926	99.926	99.926	99.926	99.926			
1995				.	.	.	.	.	.	.	.	.

Au tableau C5a, on applique une moyenne mobile centrée sur 12 mois. Le premier terme calculable est donc celui de octobre 1986 et le dernier, celui de mars 1994. Ainsi :

$$
\begin{aligned}
APR90 = {} & \frac{111.385}{24} \\
& + \frac{108.734 + 100.502 + 104.414 + 99.272 + 107.051 + 101.725 + 97.263 + 103.144 + 95.975 + 70.008 + 101.765}{12} \\
& + \frac{111.913}{24} \\
& = 100.125
\end{aligned}
$$

Les six premières valeurs, de avril à septembre 1986, non calculables avec cette moyenne mobile symétrique, seront prises égales à la première valeur calculable, celle de octobre 1986 (99.965). On procède de même pour la fin de série : la valeur calculée pour mars 1994 (99.926) est répétée pour les six mois suivants. Les coefficients saisonniers normalisés sont alors obtenus en divisant le tableau C5a par le tableau C5b, pour obtenir :

C5: FACTEURS SAISONNIERS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	111.607	108.476	100.972
1986	102.808	100.166	108.317	101.355	97.780	103.072	95.754	67.595	102.005	111.607	108.476	100.972
1987	102.808	100.166	108.317	101.457	97.841	103.282	95.459	67.626	101.938	111.319	108.522	101.194
1988	103.296	99.980	108.014	101.428	97.763	103.407	95.251	68.122	101.847	111.250	108.512	101.093
1989	103.943	99.646	107.473	101.553	97.366	103.444	95.444	68.885	101.673	111.373	108.726	100.517
1990	104.426	99.220	106.943	101.598	97.123	103.033	95.906	69.967	101.733	111.888	108.635	99.684
1991	104.365	98.930	106.615	101.807	97.110	102.858	96.278	70.587	101.714	111.870	108.373	99.401
1992	104.203	98.848	106.726	101.846	97.403	102.643	96.352	70.929	101.860	111.752	107.912	99.475
1993	103.933	98.760	106.952	101.839	97.692	102.904	96.365	70.896	101.685	111.433	107.599	99.716
1994	103.903	98.746	107.194	101.773	97.853	102.971	96.354	70.844	101.580	111.433	107.599	99.716
1995	103.903	98.746	107.194	.	.	.	.	.	.	.	.	.

Avec par exemple : $APR86 = 100*(101.319/99.965) = 101.355$

Les valeurs manquantes d'octobre 1985 à mars 1986 sont obtenues en dupliquant la première valeur calculée pour le mois considéré. De même, pour les valeurs d'octobre 1994 à mars 1995, on duplique la dernière valeur calculée pour le mois considéré.

4.2.5 Tableau C6 : Estimation de la série corrigée des variations saisonnières

4.2.5.1 Description et mode de calcul

Cette estimation s'obtient simplement en retirant à la série de départ, du tableau C1, l'estimation de la composante saisonnière du tableau C5 : *C6 = C1 _op C5*

4.2.5.2 Exemple

C6: SÉRIE DÉSAISONNALISÉE PROVISOIRE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	101.574	102.479	100.046
1986	102.264	99.416	99.212	100.632	100.078	101.524	102.019	102.421	101.929	103.480	102.524	102.075
1987	101.962	103.595	104.222	104.291	104.636	103.635	105.177	103.952	105.507	105.182	106.654	106.507
1988	106.742	107.073	107.674	107.903	109.848	109.586	108.771	110.140	111.096	110.839	110.605	113.616
1989	113.418	113.603	112.915	113.391	112.040	114.789	113.216	112.938	113.717	113.751	115.219	114.785
1990	114.486	114.397	115.761	115.313	116.621	114.771	116.562	117.420	115.679	116.469	115.515	113.738
1991	115.757	115.038	114.558	115.913	115.069	116.509	116.277	115.788	115.480	115.962	116.589	116.487
1992	116.890	116.957	116.175	116.406	115.407	114.294	114.745	115.415	115.885	115.292	114.177	113.695
1993	111.997	112.752	113.259	112.205	113.019	112.594	112.357	112.407	111.539	111.166	111.786	113.545
1994	114.536	113.924	114.379	114.800	115.029	116.094	116.345	116.238	117.220	116.071	117.429	120.535
1995	119.428	118.829	119.793	.	.	.	.	.	.	.	.	.

Avec par exemple : $APR86 = 100 * (101.995 / 101.355) = 100.632$

4.2.6 Tableau C7 : Estimation de la composante Tendance-Cycle

4.2.6.1 Description et mode de calcul

Ce tableau présente une estimation de la composante Tendance-Cycle réalisée à partir de la série désaisonnalisée du tableau précédent. Tout comme au tableau B7 le programme va utiliser une moyenne mobile de Henderson dont l'ordre dépendra de la valeur du ratio $\bar{I}/\bar{C}$.

- ***Étape 1 : choix de la moyenne mobile, calcul du ratio*** $\bar{I}/\bar{C}$
    - Si le ratio est inférieur à 1, on choisira une moyenne mobile de Henderson à 9 termes.
    - Si le ratio est supérieur à 3.5, on choisira une moyenne mobile de Henderson à 23 termes.
    - Sinon on choisira une moyenne mobile de Henderson à 13 termes.

- ***Étape 2 : lissage de la série CVS par une moyenne mobile de Henderson***

4.2.7 Remarques

- A cette étape, contrairement à la partie B, le programme choisit entre une moyenne à 9 termes, une moyenne à 13 termes ou une moyenne à 23 termes.
- La série du tableau C1 correspond à une série corrigée, au moins en partie, d'effets indésirables et en particulier de points extrêmes. Cela se fera évidemment sentir sur le numérateur du ratio $\bar{I}/\bar{C}$ qui, selon toute logique sera plus petit que celui calculé à la partie B.

4.2.7.1 Exemple

- ***Étape 1 : choix de la moyenne mobile, calcul du ratio*** $\bar{I}/\bar{C}$

On lisse tout d'abord le tableau C6 par une moyenne mobile de Henderson sur 13 termes dont les coefficients figurent dans le tableau 3 (3.3.1).

Le premier terme calculable est donc celui de avril 1986, et on a, aux arrondis près :

APR86= 101.574(-0.01935)+102.479*(0.02786)+100.046*(0.00000)*
+102.264(0.06549)+99.416*(0.14736)+99.212*(0.21434)*
+100.632(0.24006)+100.078*(0.21434)+101.524*(0.14736)*
+102.019(0.06549)+102.421*(0.00000)+101.929*(-0.02786)*
+103.480(-0.01935) =100.198*

A cette étape du calcul, on ne se préoccupe pas d'estimer les 6 points non calculables en début et en fin de série. On en déduit une estimation de la tendance (tableau C7a) et, par rapport avec le tableau C6, de la composante irrégulière (tableau C7b):

C7a : TENDANCE-CYCLE (Moyenne Mobile de Henderson sur 13 termes)

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	100.198	100.587	101.166	101.772	102.244	102.449	102.498	102.511	102.620
1987	102.881	103.274	103.699	104.070	104.311	104.422	104.534	104.765	105.141	105.593	106.028	106.428
1988	106.850	107.321	107.802	108.319	108.846	109.307	109.680	110.054	110.537	111.161	111.891	112.540
1989	112.988	113.249	113.355	113.338	113.257	113.240	113.358	113.564	113.811	114.073	114.323	114.591
1990	114.852	115.028	115.234	115.503	115.820	116.153	116.348	116.353	116.189	115.874	115.489	115.137
1991	114.922	114.947	115.150	115.426	115.680	115.834	115.909	115.959	116.015	116.137	116.345	116.579
1992	116.732	116.654	116.328	115.875	115.474	115.233	115.181	115.203	115.108	114.803	114.281	113.645
1993	113.087	112.747	112.638	112.664	112.668	112.524	112.229	111.935	111.829	111.980	112.387	112.970
1994	113.584	114.142	114.586	114.967	115.342	115.681	116.010	116.382	116.837	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

C7b : IRRÉGULIER

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	100.433	99.494	100.354	100.243	100.173	99.493	100.958	100.013	99.469
1987	99.107	100.311	100.504	100.212	100.311	99.247	100.615	99.224	100.349	99.610	100.590	100.074
1988	99.899	99.768	99.881	99.616	100.921	100.255	99.170	100.079	100.505	99.710	98.851	100.956
1989	100.381	100.313	99.612	100.046	98.926	101.368	99.875	99.449	99.917	99.718	100.784	100.169
1990	99.681	99.451	100.458	99.836	100.692	98.810	100.184	100.917	99.561	100.513	100.023	98.785
1991	100.727	100.079	99.485	100.422	99.472	100.583	100.317	99.852	99.539	99.849	100.210	99.921
1992	100.135	100.260	99.869	100.458	99.942	99.185	99.621	100.184	100.675	100.426	99.909	100.044
1993	99.036	100.004	100.552	99.593	100.312	100.063	100.115	100.422	99.740	99.274	99.466	100.509
1994	100.839	99.809	99.819	99.855	99.729	100.357	100.289	99.876	100.329	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

Le schéma étant multiplicatif, on calcule les taux de croissance (voir 4.1.7.3) :

C7c : CROISSANCE, EN VALEUR ABSOLUE, DE LA TENDANCE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC	TOTAL
1985	.	.	.	.	.	.	.	.	.	.	.	.	0.000
1986	.	.	.	.	0.389	0.575	0.600	0.464	0.200	0.048	0.012	0.106	2.394
1987	0.255	0.382	0.411	0.358	0.232	0.106	0.108	0.221	0.358	0.431	0.412	0.377	3.650
1988	0.397	0.441	0.448	0.479	0.486	0.424	0.342	0.340	0.439	0.564	0.657	0.580	5.598
1989	0.398	0.231	0.094	0.015	0.072	0.015	0.104	0.182	0.218	0.230	0.219	0.235	2.012
1990	0.228	0.153	0.178	0.233	0.274	0.288	0.167	0.004	0.141	0.271	0.333	0.305	2.576
1991	0.187	0.022	0.177	0.239	0.221	0.133	0.064	0.044	0.048	0.106	0.179	0.201	1.620
1992	0.132	0.067	0.280	0.389	0.346	0.209	0.045	0.019	0.083	0.265	0.454	0.557	2.845
1993	0.491	0.301	0.097	0.023	0.004	0.128	0.262	0.261	0.095	0.134	0.364	0.519	2.679
1994	0.543	0.491	0.389	0.332	0.326	0.294	0.285	0.321	0.390	.	.	.	3.372
1995	.	.	.	.	.	.	.	.	.	.	.	.	0.000

Soit, en utilisant les totaux par ligne :

$$\overline{C} = \frac{2.394 + 3.650 + 5.598 + 2.012 + 2.576 + 1.620 + 2.845 + 2.679 + 3.372}{101} = 0.2648$$

(en B7, cette quantité était égale à 0.312).

C7d : CROISSANCE, EN VALEUR ABSOLUE, DE L'IRRÉGULIER

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC	TOTAL
1985	.	.	.	.	.	.	.	.	.	.	.	.	0.000
1986	.	.	.	.	0.935	0.865	0.111	0.069	0.679	1.473	0.937	0.544	5.614
1987	0.364	1.215	0.193	0.290	0.098	1.061	1.379	1.383	1.134	0.736	0.984	0.513	9.350
1988	0.176	0.130	0.113	0.265	1.309	0.660	1.082	0.916	0.426	0.791	0.862	2.130	8.859
1989	0.570	0.067	0.699	0.436	1.120	2.469	1.473	0.427	0.471	0.199	1.069	0.610	9.611
1990	0.487	0.230	1.012	0.619	0.858	1.869	1.391	0.732	1.344	0.956	0.488	1.237	11.223
1991	1.966	0.643	0.593	0.941	0.946	1.117	0.264	0.464	0.314	0.311	0.362	0.288	8.210
1992	0.214	0.124	0.390	0.591	0.515	0.757	0.440	0.564	0.491	0.248	0.514	0.135	4.983
1993	1.007	0.978	0.547	0.954	0.722	0.248	0.052	0.306	0.678	0.468	0.193	1.049	7.202
1994	0.328	1.021	0.010	0.035	0.126	0.630	0.068	0.411	0.453	.	.	.	3.083
1995	.	.	.	.	.	.	.	.	.	.	.	.	0.000

Et $\bar{I} = \frac{5.614 + 9.350 + 8.859 + 9.611 + 11.223 + 8.210 + 4.983 + 7.202 + 3.083}{101} = 0.6746$

(en B7, cette quantité était égale à 2.226).

donc $\frac{\bar{I}}{\bar{C}} = \frac{0.6746}{0.2648} = 2.547$

(en B7, cette quantité était égale à 7.14).

- ***Étape 2 : lissage de la série CVS par une moyenne mobile de Henderson***

Le ratio étant supérieur à 1 et inférieur à 3.5, on choisit une moyenne mobile de Henderson sur 13 termes dont les coefficients et ceux des moyennes mobiles asymétriques associées figurent dans le tableau 3 (3.3.1). L'estimation de la tendance pour octobre 1985 se fait en utilisant le point courant et six points dans le futur auxquels on applique les coefficients de la moyenne mobile H6_0 du tableau 3.

Oct85 = 101.574(0.42113)+102.479*(0.35315)+100.046*(0.24390)+102.264*(0.11977)*
+99.416(0.01202)+99.212*(-0.05811)+ 100.632*(-0.09186) =101.801*

Ce qui conduit au tableau C7 :

C7: TENDANCE-CYCLE
valeur de I/C: 2.5476; Une moyenne mobile de HENDERSON sur 13 termes a été choisie

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	101.801	101.494	101.102
1986	100.683	100.300	100.105	100.198	100.587	101.166	101.772	102.244	102.449	102.498	102.511	102.620
1987	102.881	103.274	103.699	104.070	104.311	104.422	104.534	104.765	105.141	105.593	106.028	106.428
1988	106.850	107.321	107.802	108.319	108.846	109.307	109.680	110.054	110.537	111.161	111.891	112.540
1989	112.988	113.249	113.355	113.338	113.257	113.240	113.358	113.564	113.811	114.073	114.323	114.591
1990	114.852	115.028	115.234	115.503	115.820	116.153	116.348	116.353	116.189	115.874	115.489	115.137
1991	114.922	114.947	115.150	115.426	115.680	115.834	115.909	115.959	116.015	116.137	116.345	116.579
1992	116.732	116.654	116.328	115.875	115.474	115.233	115.181	115.203	115.108	114.803	114.281	113.645
1993	113.087	112.747	112.638	112.664	112.668	112.524	112.229	111.935	111.829	111.980	112.387	112.970
1994	113.584	114.142	114.586	114.967	115.342	115.681	116.010	116.382	116.837	117.399	118.026	118.651
1995	119.188	119.603	119.876	.	.	.	.	.	.	.	.	.

4.2.8 Tableau C9 : Estimation de la composante Saisonnier-Irrégulier

4.2.8.1 Description et mode de calcul

Tableau similaire au tableau C4 : la composante Tendance-Cycle est enlevée de la série analysée, par soustraction ou division selon le schéma de composition adopté, pour obtenir une nouvelle estimation de la composante Saisonnier-Irrégulier. On a donc : *C9 = C1 _op C7*

4.2.8.2 Remarque

Contrairement au tableau C4, les points de début et de fin de série ayant été estimés par des moyennes mobiles de Henderson asymétriques pour la Tendance-Cycle, on dispose d'une estimation complète de la composante Saisonnier-Irrégulier.

4.2.8.3 Exemple

C9 : SÉRIE BRUTE DÉBARRASSÉE DE SA TENDANCE (estimation Saisonnier-Irrégulier)

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	111.358	109.528	99.917
1986	104.422	99.284	107.350	101.794	97.285	103.438	95.986	67.712	101.487	112.676	108.490	100.436
1987	101.889	100.478	108.863	101.672	98.145	102.504	96.046	67.101	102.294	110.886	109.163	101.269
1988	103.191	99.748	107.886	101.039	98.663	103.670	94.461	68.175	102.362	110.928	107.264	102.059
1989	104.338	99.959	107.056	101.600	96.320	104.859	95.325	68.505	101.589	111.060	109.579	100.687
1990	104.093	98.675	107.433	101.432	97.795	101.807	96.083	70.609	101.286	112.462	108.660	98.473
1991	105.124	99.008	106.066	102.236	96.597	103.458	96.584	70.483	101.245	111.701	108.601	99.323
1992	104.344	99.104	106.586	102.313	97.346	101.807	95.988	71.059	102.548	112.228	107.814	99.518
1993	102.932	98.764	107.542	101.425	97.997	102.968	96.476	71.195	101.420	110.624	107.024	100.224
1994	104.774	98.558	107.001	101.625	97.587	103.338	96.632	70.757	101.913	110.173	107.055	101.300
1995	104.112	98.106	107.120	.	.	.	.	.	.	.	.	.

La valeur d'avril 1986 s'obtient donc simplement : $APR86 = 100*(101.995/100.198) = 101.794$

4.2.9 Tableau C10 : Estimation de la composante saisonnière

4.2.9.1 Description et mode de calcul

Cette estimation est obtenue d'une façon similaire à celle conduisant au tableau B10, à partir des valeurs de la composante Saisonnier-Irrégulier du tableau C9. On procède là encore en deux étapes : estimation de la composante saisonnière, à l'aide d'une moyenne mobile 3x5, puis normalisation des coefficients saisonniers.

4.2.9.2 Exemple

- ***Étape 1 : estimation de la composante saisonnière***.

Les données du tableau C9 sont lissées colonne par colonne (mois par mois), avec une moyenne mobile 3x5 (voir tableau 4, 3.3.3), pour aboutir au tableau C10a.

Le facteur saisonnier du mois d'avril 1989 sera donc estimé par :

$$APR89 = \frac{101.794 + 2*101.672 + 3*101.039 + 3*101.600 + 3*101.432 + 2*102.236 + 102.313}{15} = 101.609$$

Pour le début de la série (années 1986 à 1988) et la fin de la série (années 1992 à 1994), on utilise des moyennes asymétriques prédéfinies, par exemple :

$$APR87 = \frac{15*101.794 + 15*101.672 + 15*101.039 + 11*101.600 + 4*101.432}{60} = 101.515$$

(un point dans le passé, le point courant et trois points dans le futur)

C10a: FACTEURS SAISONNIERS PROVISOIRES (mm 3x5)

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	111.533	108.791	100.768
1986	103.343	99.855	107.886	101.516	97.775	103.452	95.472	67.789	101.979	111.471	108.766	100.829
1987	103.444	99.782	107.814	101.515	97.702	103.414	95.505	68.014	101.913	111.482	108.683	100.793
1988	103.591	99.690	107.656	101.543	97.583	103.378	95.578	68.423	101.831	111.464	108.696	100.607
1989	103.844	99.500	107.394	101.609	97.497	103.212	95.656	69.054	101.788	111.562	108.556	100.320
1990	104.037	99.325	107.134	101.707	97.353	103.120	95.826	69.704	101.726	111.566	108.457	100.007
1991	104.212	99.074	106.956	101.777	97.340	102.925	96.044	70.319	101.702	111.576	108.183	99.808
1992	104.226	98.877	106.908	101.836	97.352	102.850	96.288	70.688	101.694	111.411	107.930	99.835
1993	104.184	98.721	106.950	101.846	97.462	102.783	96.386	70.882	101.784	111.232	107.627	100.035
1994	104.092	98.627	107.027	101.855	97.486	102.818	96.398	70.925	101.853	111.112	107.493	100.194
1995	104.000	98.570	107.126	.	.	.	.	.	.	.	.	.

- ***Étape 2 : normalisation des coefficients saisonniers***.

C10b: MOYENNE MOBILE CENTRÉE SUR 12 MOIS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.011	100.011	100.011
1986	100.011	100.011	100.011	100.011	100.007	100.009	100.015	100.016	100.010	100.007	100.004	99.999
1987	99.999	100.010	100.017	100.014	100.011	100.006	100.011	100.013	100.003	99.998	99.994	99.987
1988	99.989	100.009	100.023	100.018	100.018	100.011	100.014	100.016	99.998	99.990	99.989	99.978
1989	99.974	100.004	100.029	100.031	100.029	100.011	100.007	100.008	99.990	99.983	99.981	99.971
1990	99.975	100.009	100.033	100.031	100.027	100.010	100.004	100.001	99.983	99.979	99.981	99.972
1991	99.973	100.008	100.033	100.032	100.021	100.001	99.994	99.986	99.976	99.976	99.979	99.977
1992	99.984	100.009	100.024	100.017	100.000	99.990	99.989	99.981	99.976	99.979	99.984	99.985
1993	99.987	99.999	100.011	100.007	99.987	99.983	99.987	99.979	99.979	99.982	99.984	99.986
1994	99.988	99.990	99.995	99.993	99.982	99.983	99.986	99.980	99.982	99.982	99.982	99.982
1995	99.982	99.982	99.982	.	.	.	.	.	.	.	.	.

Au tableau C10a, on applique une moyenne mobile centrée sur 12 mois. Le premier terme calculable est donc celui de avril 1986 et le dernier, celui de septembre 1994. Ainsi :

$$
\begin{aligned}
APR86 &= \frac{111.533}{24} \\
&+ \frac{108.791+100.768+103.343+99.855+107.886+101.516+97.775+103.452+95.472+67.789+101.979}{12} \\
&+ \frac{111.471}{24} \\
&= 100.011
\end{aligned}
$$

Les six premières valeurs, de octobre 1985 à mars 1986, non calculables avec cette moyenne mobile symétrique, seront prises égales à la première valeur calculable, celle de avril 1986. On procède de même pour la fin de série : la valeur calculée pour septembre 1994 (99.982) est répétée pour les six mois suivants. Les coefficients saisonniers normalisés sont alors obtenus en divisant le tableau C10a par le tableau C10b, pour obtenir :

C10: FACTEURS SAISONNIERS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	111.521	108.779	100.758
1986	103.332	99.844	107.875	101.505	97.768	103.444	95.457	67.778	101.968	111.463	108.761	100.829
1987	103.444	99.772	107.796	101.500	97.691	103.408	95.495	68.004	101.910	111.485	108.689	100.806
1988	103.603	99.681	107.631	101.524	97.565	103.366	95.565	68.412	101.833	111.475	108.709	100.629
1989	103.870	99.497	107.364	101.578	97.468	103.200	95.649	69.048	101.799	111.580	108.576	100.348
1990	104.063	99.316	107.098	101.675	97.327	103.110	95.822	69.703	101.743	111.590	108.478	100.034
1991	104.240	99.066	106.921	101.744	97.319	102.924	96.050	70.329	101.727	111.603	108.206	99.832
1992	104.243	98.868	106.882	101.818	97.353	102.860	96.298	70.701	101.718	111.435	107.947	99.849
1993	104.198	98.722	106.939	101.839	97.474	102.801	96.399	70.896	101.806	111.252	107.645	100.048
1994	104.105	98.636	107.033	101.862	97.504	102.835	96.411	70.940	101.872	111.133	107.512	100.212
1995	104.019	98.588	107.145	.	.	.	.	.	.	.	.	.

Avec, par exemple : APR86=100*(101.516/100.011)=101.505

4.2.10 Tableau C11 : Estimation de la série corrigée des variations saisonnières

4.2.10.1 Description et mode de calcul

Cette estimation s'obtient simplement en retirant à la série de départ, du tableau B1 (et non pas C1), l'estimation de la composante saisonnière du tableau C10 : *C11 = B1 _op C10*
Ainsi, puisque l'on part de la série B1, cette série corrigée des variations saisonnières comporte les valeurs extrêmes détectées précédemment.

4.2.10.2 Exemple

C11: SÉRIE DÉSAISONNALISÉE PROVISOIRE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	103.747	100.938	99.844
1986	103.163	98.854	96.315	107.876	99.931	100.248	104.445	96.934	103.169	105.058	99.576	103.541
1987	97.154	103.436	104.735	105.517	102.364	104.731	106.603	101.023	106.663	104.858	105.530	109.121
1988	103.955	110.553	110.284	106.477	110.080	110.965	105.897	111.092	112.537	105.763	111.582	113.983
1989	113.507	112.768	111.956	112.919	113.370	116.569	110.404	114.992	112.182	113.550	116.784	112.309
1990	116.371	113.275	115.408	114.187	118.775	113.277	116.674	119.507	112.636	118.290	117.167	110.762
1991	118.284	113.863	111.578	117.353	116.421	113.385	120.042	116.026	114.424	118.635	115.336	115.995
1992	118.474	118.238	116.015	117.857	112.786	115.400	116.410	113.152	117.285	115.762	113.111	113.972
1993	109.120	114.564	114.739	112.138	110.696	113.910	112.138	112.418	112.764	108.762	113.057	114.745
1994	111.714	113.042	115.852	113.290	116.919	117.665	113.576	120.384	118.384	113.738	118.777	119.746
1995	119.305	117.965	121.517	.	.	.	.	.	.	.	.	.

Avec par exemple : $APR86 = 100*(109.500/101.505) = 107.876$

4.2.11 Tableau C13 : Estimation de la composante irrégulière

4.2.11.1 Description et mode de calcul

Cette estimation s'obtient simplement en retirant à la série corrigée des variations saisonnière, du tableau C11, l'estimation de la composante Tendance-Cycle du tableau C7 : *C13 = C11 _op C7*

4.2.11.2 Exemple

C13: SÉRIE DES FACTEURS IRRÉGULIERS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	101.911	99.452	98.755
1986	102.463	98.559	96.214	107.663	99.347	99.093	102.626	94.806	100.703	102.497	97.137	100.898
1987	94.433	100.157	100.999	101.390	98.133	100.296	101.979	96.428	101.448	99.303	99.531	102.530
1988	97.290	103.011	102.302	98.300	101.134	101.516	96.550	100.944	101.809	95.144	99.724	101.282
1989	100.459	99.576	98.765	99.630	100.100	102.940	97.394	101.258	98.569	99.542	102.153	98.008
1990	101.323	98.475	100.151	98.861	102.552	97.524	100.281	102.711	96.942	102.085	101.453	96.200
1991	102.926	99.057	96.898	101.670	100.640	97.885	103.566	100.058	98.629	102.151	99.133	99.499
1992	101.492	101.357	99.731	101.710	97.672	100.145	101.066	98.220	101.892	100.836	98.976	100.288
1993	96.492	101.612	101.865	99.533	98.250	101.232	99.920	100.431	100.835	97.127	100.597	101.571
1994	98.354	99.036	101.105	98.542	101.367	101.715	97.902	103.439	101.325	96.882	100.637	100.923
1995	100.098	98.631	101.369	.	.	.	.	.	.	.	.	.

Avec par exemple : $APR86 = 100*(107.876/100.198) = 107.663$

4.2.12 Tableau C14 : Valeurs de la composante irrégulière exclues de la régression pour Jours Ouvrables

4.2.12.1 Description et mode de calcul

Les tableaux C14, C15 et C16 portent sur l'estimation finale de l'effet de la composition journalière du mois. Au tableau C14, X11 va repérer les valeurs extrêmes de la composante irrégulière et les exclure des calculs. La recherche se fait en deux phases :

- ***Étape 1 : Calcul d'un écart-type global et repérage des valeurs atypiques***

L'étape B nous fournit une première estimation de l'effet jours ouvrables, effet résumé dans le tableau B16 (ou B18 si une correction *a priori* a été demandée par l'utilisateur). La composante irrégulière du tableau C13 est alors corrigée de cet effet pour obtenir un résidu : *Res = C13 – B16*[29].
On calcule ensuite une estimation de la variance de ce résidu, par la moyenne des carrés de ce résidu, ce qui revient à supposer que les résidus sont de moyenne nulle.

On a donc : $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(C13_i - B16_i)^2 = \frac{1}{N}\sum_{i=1}^{N}(R_i)^2$

Une valeur I_i de l'irrégulier est jugée atypique si le résidu associé R_i est trop grand et, plus précisément, si : $|R_i| \geq \lambda * \sigma$ où σ est l'écart-type global calculé ci-dessus et λ un paramètre modifiable par l'utilisateur et fixé par défaut à 2.5.

- ***Étape 2 : Calcul final de l'écart-type global et repérage des valeurs atypiques***

On refait les calculs ci-dessus en excluant les valeurs atypiques repérées. On obtient alors un nouvel écart-type et de nouvelles valeurs atypiques qui seront exclues de la régression pour jours ouvrables ; ces valeurs figurent dans le tableau C14.

4.2.12.2 Remarques

- Le calcul fait ici est donc sensiblement différent de celui fait en B14. L'utilisation des coefficients correcteurs du tableau B16 permet, dans la mesure où ces coefficients ont été calculés à partir de la structure journalière de chaque mois, de distinguer tous les types de mois et non 15 types seulement comme à l'étape B.
- Toutes les observations, hormis les valeurs détectées comme atypiques, sont utilisées pour la régression. Dans X11-ARIMA88, lorsqu'une extrapolation ARIMA est demandée, les coefficients journaliers sont estimés sur l'ensemble des données disponibles jusqu'au dernier mois de décembre[30].

[29] Il s'agit bien ici d'une soustraction dans la mesure où le modèle posé pour l'effet jours ouvrables est un modèle de régression linéaire.
[30] Sauf si on demande aussi une correction de l'effet de Pâques ; dans ce cas, tous les points sont utilisés.

4.2.12.3 Exemple

- ***Étape 1 : Premier calcul d'un écart-type global et repérage des valeurs atypiques***

On corrige la composante irrégulière du tableau C13 par les coefficients pour jours ouvrables du tableau B16 pour obtenir les résidus en valeur absolue du tableau C14a.

C14a : COMPOSANTE IRRÉGULIÈRE CORRIGÉE DES EFFETS POUR JOURS OUVRABLES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	0.149	0.680	1.254
1986	1.070	0.556	1.511	6.596	0.493	0.006	0.565	2.872	0.477	1.104	0.243	0.396
1987	5.407	1.042	0.990	0.171	0.455	0.884	0.586	1.298	0.381	0.536	0.432	0.469
1988	0.388	0.070	0.241	0.472	1.125	0.298	1.128	0.350	0.526	2.581	1.343	1.442
1989	0.450	0.460	2.628	2.250	1.193	1.657	0.331	0.803	0.203	0.467	0.935	0.330
1990	0.029	0.640	0.312	0.238	0.491	1.248	0.272	1.318	0.438	0.791	0.170	1.525
1991	0.865	0.058	0.781	0.489	0.753	0.505	2.272	0.218	0.469	0.090	0.361	0.510
1992	0.098	0.241	0.278	0.492	0.007	1.036	0.327	0.494	0.824	0.996	0.123	1.773
1993	1.187	2.497	0.572	1.750	0.524	0.165	0.080	0.422	0.383	0.552	0.584	0.178
1994	0.629	0.079	0.955	0.230	1.358	0.496	0.223	2.145	0.041	0.844	0.430	1.084
1995	0.089	0.485	0.024	.	.	.	.	.	.	.	.	.

Ainsi on a, pour le mois d'avril 1986 : $APR86 = 107.663 - 101.067 = 6.596$
La moyenne des carrés des éléments du tableau C14a fournit une première estimation de l'écart-type :

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(R_i\right)^2} = 1.2302$$

Les deux seuls points qui s'éloignent, en valeur absolue, de plus de 3.076 (soit 2.5σ) sont les valeurs de avril 1986 et janvier 1987. Ces deux points sont donc exclus.

- ***Étape 2 : Calcul final de l'écart-type global et repérage des valeurs atypiques***

Le nouvel écart-type se calcule simplement en enlevant les résidus concernés et on obtient $\sigma = 0.9439$
Ce nouveau calcule conduit à une borne égale à 2.360 (soit 2.5σ) et donc à l'élimination de 6 points : APR86, AUG86, JAN87, OCT88, MAR89 et FEB93.

C14 : VALEURS DE L'IRRÉGULIER EXCLUES DE LA RÉGRESSION POUR JOURS OUVRABLES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	107.663	.	.	.	94.806	.	.	.	.
1987	94.433	.	.	.	.	.	.	.	.	.	.	.
1988	.	.	.	.	.	.	.	.	.	95.144	.	.
1989	.	.	98.765	.	.	.	.	.	.	.	.	.
1990	.	.	.	.	.	.	.	.	.	.	.	.
1991	.	.	.	.	.	.	.	.	.	.	.	.
1992	.	.	.	.	.	.	.	.	.	.	.	.
1993	.	101.612	.	.	.	.	.	.	.	.	.	.
1994	.	.	.	.	.	.	.	.	.	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

4.2.13 Tableau C15 : Régression finale pour Jours Ouvrables

4.2.13.1 Description et mode de calcul

Des poids journaliers vont être à présent estimés grâce à une régression par moindres carrés ordinaires sur les données non jugées atypiques du tableau C13, selon une méthodologie identique à celle présentée au tableau B15 (voir 4.1.16.1) et que nous ne reprendrons pas ici.

4.2.13.2 Exemple

C15 : RÉGRESSION POUR JOURS OUVRABLES

	Poids Combinés	Poids a Priori	Coefficients de la régression	Std	Student	Prob >t
Lundi	1.092	1.000	0.092	0.067	1.373	0.086
Mardi	1.242	1.000	0.242	0.066	3.649	0.000
Mercredi	1.083	1.000	0.083	0.068	1.210	0.114
Jeudi	1.356	1.000	0.356	0.068	5.215	0.000
Vendredi	1.076	1.000	0.076	0.068	1.126	0.131
Samedi	0.518	1.000	-0.482	0.066	-7.281	0.000
Dimanche	0.632	1.000	-0.368	0.067	-5.458	0.000

	Somme des Carrés	D.D.L	Moyenne des Carrés	F-Value	Prob > F
RÉGRESSION	26.115	6	4.352	68.245	0.000
ERREUR	6.505	106	0.064		
TOTAL	32.620	112			

Si nous nous fixons un risque de première espèce de 1 % par exemple, les tests s'interprètent de la façon suivante :

- Le test de Fisher rejette l'hypothèse nulle d'égalité des coefficients journaliers ; on peut donc admettre l'existence d'un effet dû à la composition journalière du mois. En effet, la probabilité de trouver une valeur de la statistique de Fisher plus grande que celle trouvée (68.245) est quasiment nulle et donc inférieure à notre risque de première espèce. Nous sommes donc dans la région critique du test ce qui ne nous permet pas d'accepter l'hypothèse nulle d'égalité des coefficients journaliers.
- On interprète de la même manière les tests de Student mais en prenant garde à ce que la loi de Student est symétrique. Il faut donc comparer la valeur (Prob > |t|) à la moitié du risque de première espèce soit 0.005. Tous les tests conduisant à une valeur plus petite que 0.005 ne permettent pas d'accepter l'hypothèse de nullité du coefficient journalier. Dans notre cas, les coefficients de mardi, jeudi, samedi et dimanche sont jugés significativement différents de 0.
- Enfin, on peut remarquer que par rapport aux résultats du tableau B15, la précision des estimation s'est accrue grâce à la correction des valeurs atypiques.

4.2.14 Tableau C16 : Coefficients d'ajustement pour jours ouvrables issus de la régression

4.2.14.1 Description et mode de calcul

Des estimations de la régression, on déduit le tableau C16, similaire au tableau B16, des coefficients mensuels M_i d'ajustement pour jours ouvrables.

La composante irrégulière du tableau C13 est alors corrigée de ces effets de calendrier pour aboutir à un tableau "C16bis", malheureusement non imprimable dans les logiciels usuels :

C16bis = C13 _op C16

4.2.14.2 Remarques

- X11-ARIMA88 présente en plus un tableau C16A répétant les coefficients de chaque jour obtenus par la régression
- X12-ARIMA et X11-ARIMA88 présentent aussi un tableau C16C estimant les coefficients de correction des douze prochains mois.

4.2.14.3 Exemple

C16 : COEFFICIENTS DE CORRECTION POUR JOURS OUVRABLES ISSUS DE LA RÉGRESSION

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	102.198	98.646	99.895
1986	101.662	99.115	97.557	101.084	99.839	99.083	102.198	97.504	101.116	101.662	97.167	101.347
1987	99.839	99.115	99.895	101.463	97.504	101.116	101.662	97.557	101.084	99.839	99.083	102.198
1988	97.504	102.982	102.198	98.646	99.895	101.463	97.504	101.347	101.441	97.557	101.084	99.839
1989	99.895	99.115	101.662	97.167	101.347	101.441	97.557	102.198	98.646	99.895	101.463	97.504
1990	101.347	99.115	99.839	99.083	102.198	98.646	99.895	101.662	97.167	101.347	101.441	97.557
1991	102.198	99.115	97.504	101.116	101.662	97.167	101.347	99.839	99.083	102.198	98.646	99.895
1992	101.662	100.947	99.895	101.463	97.504	101.116	101.662	97.557	101.084	99.839	99.083	102.198
1993	97.504	99.115	101.347	101.441	97.557	101.084	99.839	99.895	101.463	97.504	101.116	101.662
1994	97.557	99.115	102.198	98.646	99.895	101.463	97.504	101.347	101.441	97.557	101.084	99.839
1995	99.895	99.115	101.662	.	.	.	.	.	.	.	.	.

Prenons comme exemples les mois d'avril 1986 et février 1989.

	Lundi	Mardi	Mercredi	Jeudi	Vendredi	Samedi	Dimanche	Nb jours
poids	1.092	1.242	1.083	1.356	1.076	0.518	0.632	
Nombre de								
APR86	4	5	5	4	4	4	4	30
FEB89	4	4	4	4	4	4	4	28

Les coefficients de correction pour ces mois seront donc, aux arrondis près :

$$APR86 = \frac{100*(4*1.0925 + 5*1.2424 + 5*1.0828 + 4*1.3562 + 4*1.0761 + 4*0.5176 + 4*0.6325)}{30} \approx 101.084$$

ou plus simplement, puisque Mardi et Mercredi sont les seuls jours intervenant 5 fois dans le mois :

$$APR86 = \frac{100*(28 + 1.2424 + 1.0828)}{30} \approx 101.084$$

et, pour un mois de Février de 28 jours : $FEB89 = \frac{28}{28.25} = 99.115$

On obtient ensuite une valeur corrigée de la composante irrégulière (tableau C16bis ci-après) :

$$APR86 = 100 * \frac{107.663}{101.084} = 106.509 \text{ et } FEB89 = 100 * \frac{99.576}{99.115} = 100.465$$

C16bis : IRRÉGULIER CORRIGÉ DES EFFETS JOURS OUVRABLES ISSUS DE LA RÉGRESSION

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	99.720	100.817	98.860
1986	100.788	99.439	98.624	106.509	99.508	100.010	100.419	97.234	99.592	100.822	99.969	99.557
1987	94.585	101.051	101.106	99.928	100.645	99.189	100.312	98.843	100.360	99.464	100.452	100.325
1988	99.781	100.028	100.102	99.649	101.241	100.052	99.022	99.602	100.363	97.527	98.655	101.446
1989	100.565	100.465	97.151	102.535	98.770	101.478	99.834	99.080	99.922	99.647	100.680	100.517
1990	99.976	99.355	100.313	99.776	100.346	98.863	100.387	101.033	99.769	100.728	100.012	98.610
1991	100.712	99.941	99.378	100.547	98.995	100.739	102.189	100.220	99.542	99.954	100.494	99.604
1992	99.833	100.406	99.837	100.244	100.172	99.039	99.415	100.680	100.799	100.999	99.892	98.131
1993	98.962	102.519	100.511	98.119	100.710	100.147	100.081	100.537	99.381	99.613	99.486	99.911
1994	100.818	99.920	98.931	99.895	101.474	100.248	100.408	102.064	99.885	99.308	99.558	101.086
1995	100.204	99.511	99.712	.	.	.	.	.	.	.	.	.

4.2.15 Tableau C17 : Poids préliminaires pour la correction de l'irrégulier

4.2.15.1 Description et mode de calcul

A partir de l'estimation de la composante irrégulière C16bis, ou du tableau C13 si aucune correction pour jours ouvrables n'est demandée, on cherche à identifier et à corriger les points atypiques. Pour cela, on utilise l'algorithme de détection des points atypiques et de calcul de poids correctifs détaillé aux tableaux B4 et B9 (voir respectivement 4.1.4 et 4.1.9). Comme on dispose déjà d'une estimation de l'irrégulier, seules les étapes 4 et 5 sont appliquées.

4.2.15.2 Remarques

Les remarques faites aux tableaux B4 et B9 concernant le calcul des écart-types mobiles et des poids correctifs sont aussi valables ici.

4.2.15.3 Exemple

- ***Calcul d'un écart-type mobile***

L'écart-type correspondant à l'année 1989 sera calculé à partir des données des années 1987 à 1991 (deux années avant, deux années après) selon la formule [31]:

$$\sigma_{89} = \sqrt{\frac{1}{60}\sum_{i=1}^{60}(x_i - 100)^2} = 1.4629$$

Ceux des années 1988, 1990, 1991 et 1992 sont calculés selon le même principe. Pour X11-ARIMA88 et X12-ARIMA, l'écart-type de 1987 est calculé à partir de l'ensemble des observations disponibles de 1985 à 1990, soit 63 observations. Les résultats du calcul de X11-ARIMA88 et X12-ARIMA figurent dans le tableau C17a, colonne Écart-type 1.

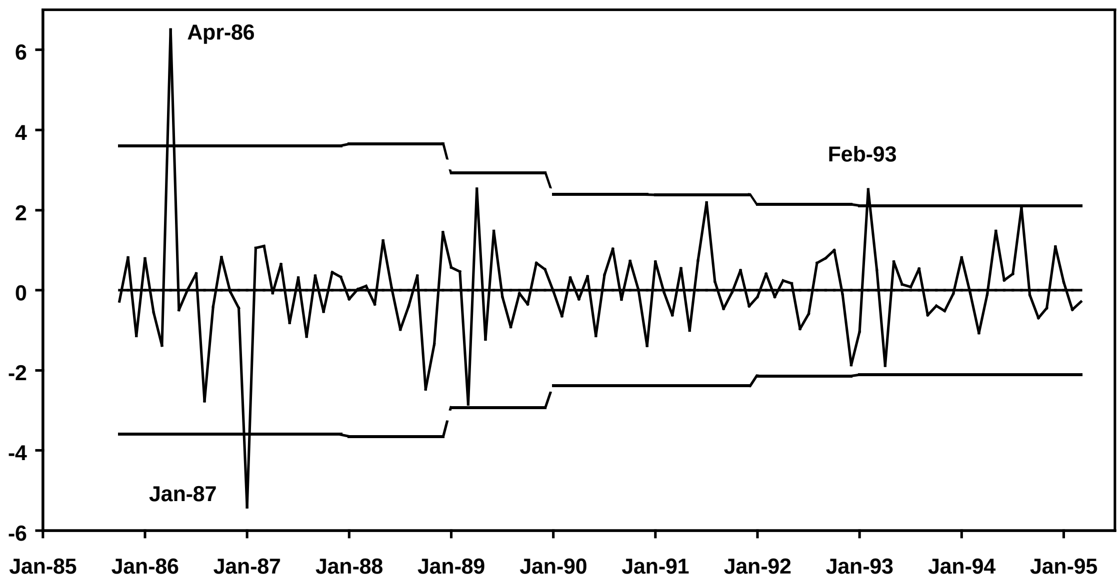
C17a : VALEURS DES ÉCART-TYPES MOBILES SUR CINQ ANS

ANNÉE	ÉCART-TYPE 1	ÉCART-TYPE 2
1985	1.4389	0.9815
1986	1.4389	0.9815
1987	1.4389	0.9815
1988	1.4629	0.9889
1989	1.1712	0.9476
1990	0.9538	0.9538
1991	0.9526	0.9030
1992	0.8592	0.8021
1993	0.8420	0.7861
1994	0.8420	0.7861
1995	0.8420	0.7861

Ce premier calcul sert à repérer d'éventuels points atypiques. Comme le montre la figure 17, les valeurs de avril 1986, janvier 1987 et février 1993 sont détectées comme très atypiques.

[31] où la moyenne théorique est considérée égale à 100 pour tenir compte du fait que les valeurs de l'irrégulier ont été elles-mêmes multipliées par 100

Figure 17 : Écart à la moyenne de l'irrégulier et "limites supérieures de confiance" (tableau C17)



On a en effet :

$$\left|APR86-100\right| = \left|106.509-100\right| \ = 6.509 \ > \ 2.5*\sigma_{86} = 2.5*1.4389 = 3.597 \ ,$$

$$\left|JAN87-100\right| = \left|94.585-100\right| \ = 5.415 \ > \ 2.5*\sigma_{87} = 2.5*1.4389 = 3.597 \ \text{ et}$$

$$\left|FEB93-100\right| = \left|102.519-100\right| \ = 2.519 \ > \ 2.5*\sigma_{93} = 2.5*0.8420 = 2.105$$

Ces points sont donc éliminés du second calcul de l'écart-type mobile qui conduit aux résultats de la colonne "Écart-type 2" du tableau C17a.

- ***Détection et correction des valeurs atypiques***

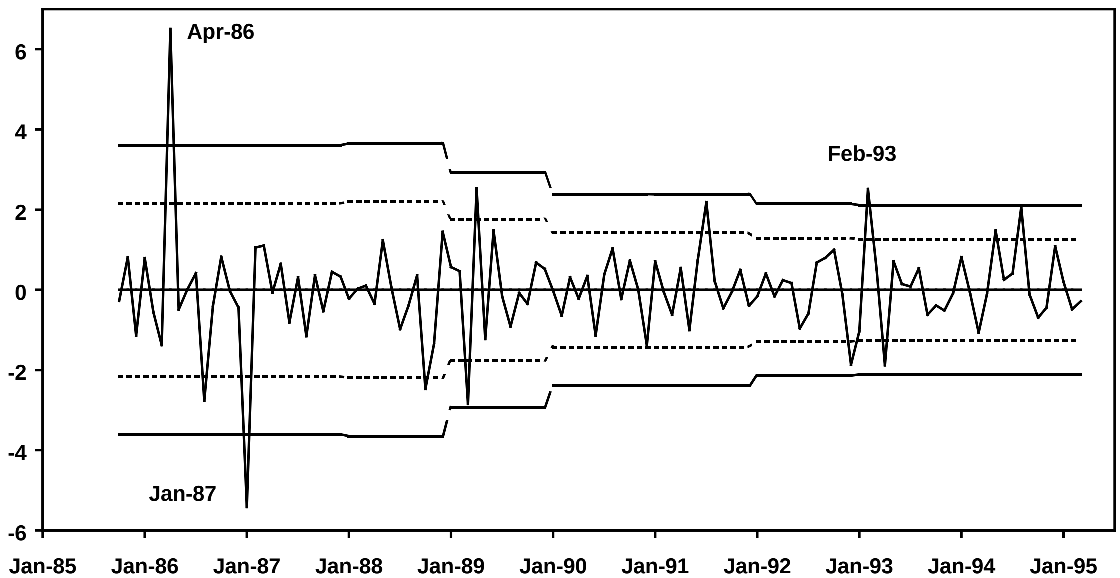
Les valeurs de l'irrégulier sont alors situées par rapport aux limites de confiance supérieures et inférieures calculées à partir des nouvelles estimations des écart-types. Toutes les valeurs situées au delà des limites de confiance inférieures (voir figure 18) sont considérées comme atypiques et vont donc être corrigées, à des degrés divers, et les poids associés à chacune de ces valeurs figurent dans le tableau C17.

C17 : POIDS PRÉLIMINAIRES POUR LA CORRECTION DE L'IRRÉGULIER

BORNES RETENUES : DE 1.5 À 2.5 SIGMA

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.000	100.000	100.000
1986	100.000	100.000	100.000	0.000	100.000	100.000	100.000	0.000	100.000	100.000	100.000	100.000
1987	0.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	0.000	100.000	100.000
1989	100.000	100.000	0.000	0.000	100.000	94.034	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1991	100.000	100.000	100.000	100.000	100.000	100.000	7.552	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	16.963
1993	100.000	0.000	100.000	10.773	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1994	100.000	100.000	100.000	100.000	62.449	100.000	100.000	0.000	100.000	100.000	100.000	100.000
1995	100.000	100.000	100.000	.	.	.	.	.	.	.	.	.

Figure 18 : Écart à la moyenne de l'irrégulier et "limites de confiance" (tableau C17)



Les valeurs jugées précédemment très atypiques le restent et sont affectées d'un poids nul. Ainsi :

$$\left|APR86-100\right| = \left|106.509-100\right| = 6.509 > 2.5*\sigma_{86} = 2.5*0.9815 = 2.454$$

Pour juillet 1991, situé entre les deux limites de confiance et donc jugé moyennement atypique on a :

$$\left|JUL91-100\right| = \left|102.189-100\right| = 2.189$$

$$1.5*\sigma_{91} = 1.5*0.9030 = 1.3545 < 2.189 < 2.5*\sigma_{91} = 2.5*0.9030 = 2.2575$$

On va attribuer à cette valeur jugée moyennement atypique, un poids proportionnel à l'écart à la moyenne constaté de, aux arrondis près :

$$poids(JUL91) = \frac{2.2575-2.189}{2.2575-1.3545} = 0.075$$

4.2.16 Tableau C18 : Coefficients pour jours ouvrables combinés (issus de l'ajustement *a priori* et de la régression pour jours ouvrables)

4.2.16.1 Description et mode de calcul

Comme au tableau B18, si on a fourni au préalable des coefficients journaliers de correction a priori des effets de jours ouvrables (schéma multiplicatif seulement) et si on a aussi demandé une régression pour jours ouvrables, le tableau C18 présente le résultat combiné de ces deux corrections, simple addition des deux effets, où les poids combinés proviennent du tableau C15.

4.2.16.2 Remarque

Dans X12-ARIMA, d'autres effets peuvent être estimés au moment de la régression sur la composante irrégulière (Pâques, Labor Day, Thanksgiving …..). Dans ce cas le tableau C18 prend en compte l'ensemble des effets.

4.2.16.3 Exemple

Dans notre cas, le tableau C18 est identique au tableau C16.

C18: COEFFICIENTS COMBINÉS DE CORRECTION POUR EFFET DE JOURS OUVRABLES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	102.198	98.646	99.895
1986	101.662	99.115	97.557	101.084	99.839	99.083	102.198	97.504	101.116	101.662	97.167	101.347
1987	99.839	99.115	99.895	101.463	97.504	101.116	101.662	97.557	101.084	99.839	99.083	102.198
1988	97.504	102.982	102.198	98.646	99.895	101.463	97.504	101.347	101.441	97.557	101.084	99.839
1989	99.895	99.115	101.662	97.167	101.347	101.441	97.557	102.198	98.646	99.895	101.463	97.504
1990	101.347	99.115	99.839	99.083	102.198	98.646	99.895	101.662	97.167	101.347	101.441	97.557
1991	102.198	99.115	97.504	101.116	101.662	97.167	101.347	99.839	99.083	102.198	98.646	99.895
1992	101.662	100.947	99.895	101.463	97.504	101.116	101.662	97.557	101.084	99.839	99.083	102.198
1993	97.504	99.115	101.347	101.441	97.557	101.084	99.839	99.895	101.463	97.504	101.116	101.662
1994	97.557	99.115	102.198	98.646	99.895	101.463	97.504	101.347	101.441	97.557	101.084	99.839
1995	99.895	99.115	101.662	.	.	.	.	.	.	.	.	.

4.2.17 Tableau C19 : Série brute corrigée des effets de jours ouvrables

4.2.17.1 Description et mode de calcul

La série du tableau B1, ou du tableau A1 si aucun ajustement préalable n'est demandé, est corrigée par les effets de jours ouvrables estimés précédemment (tableau C18). On a donc :

C19 = B1 _op C18

4.2.17.2 Exemple

C19: SÉRIE BRUTE CORRIGÉE DES EFFETS DE JOURS OUVRABLES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	113.212	111.307	100.706
1986	104.858	99.581	106.502	108.326	97.858	104.660	97.556	67.382	104.039	115.186	111.458	103.012
1987	100.663	104.121	113.019	105.555	102.560	107.105	100.136	70.421	107.535	117.089	115.761	107.634
1988	110.457	107.009	116.147	109.584	107.513	113.046	103.791	74.990	112.972	120.853	119.999	114.886
1989	118.024	113.202	118.236	118.044	109.031	118.591	108.245	77.692	115.768	126.834	124.971	115.585
1990	119.490	113.504	123.800	117.174	113.114	118.404	111.918	81.939	117.941	130.246	125.294	113.575
1991	120.648	113.807	122.354	118.082	111.448	120.103	113.767	81.732	117.477	129.553	126.513	115.922
1992	121.482	115.803	124.131	118.269	112.611	117.390	110.268	82.004	118.021	129.209	123.230	111.353
1993	116.611	114.110	121.069	112.578	110.602	115.845	108.275	79.784	113.144	124.098	120.357	112.924
1994	119.213	112.496	121.333	116.984	114.120	119.255	112.303	84.265	118.887	129.566	126.331	120.194
1995	124.231	117.338	128.072	.	.	.	.	.	.	.	.	.

On a donc par exemple : $APR86 \ = \ 100 * (109.500 \, / \, 101.084) \ = 108.326$

4.2.18 Tableau C20 : Valeurs de correction des points atypiques de l'irrégulier

4.2.18.1 Description et mode de calcul

Les valeurs de la composante irrégulière C16bis, ou C13 si on ne demande pas de régression pour jours ouvrables, détectées comme atypiques lors de la constitution du tableau C17, et pour lesquelles on a donc calculé un poids, sont corrigées, comme au tableau B20, de la façon suivante :

*C20 = C16bis _op [_xbar + C17 * (C16bis – _xbar)]*

4.2.18.2 Exemple

La valeur de mai 1994, détectée comme atypique et affectée d'un poids égal à 0.62449 sera donc corrigée de la façon suivante, puisque nous avons ici un schéma multiplicatif :

$$MAY94 = 100 * \frac{1.01474}{1 + 0.62449 * (1.01474 - 1)} = 100.549$$

C20:VALEURS DE CORRECTION POUR LES POINTS JUGÉS ATYPIQUES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.000	100.000	100.000
1986	100.000	100.000	100.000	106.509	100.000	100.000	100.000	97.234	100.000	100.000	100.000	100.000
1987	94.585	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	97.527	100.000	100.000
1989	100.000	100.000	97.151	102.535	100.000	100.087	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1991	100.000	100.000	100.000	100.000	100.000	100.000	102.021	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	98.443
1993	100.000	102.519	100.000	98.319	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1994	100.000	100.000	100.000	100.000	100.549	100.000	100.000	102.064	100.000	100.000	100.000	100.000
1995	100.000	100.000	100.000	.	.	.	.	.	.	.	.	.

4.3 PARTIE D : Estimation finale des différentes composantes

4.3.1 Tableau D1 : Série brute ajustée pour tenir compte des ajustements *a priori,* de ceux liés à la correction pour jours ouvrables et des points atypiques détectés

4.3.1.1 Description et mode de calcul

Ce tableau présente la série brute corrigée des divers effets mis en évidence lors de la partie C, points jugés atypiques et effets liés aux jours ouvrables, et ajustée *a priori* par les éléments de la partie A.
Il est donc calculé à partir du tableau C19 qui prend en compte les effets dus aux jours ouvrables, ou du tableau B1 si la régression pour jours ouvrables n'est pas demandée, et du tableau C20 qui précise les corrections à apporter aux points jugés atypiques.
On a donc : *D1 = C19 _op C20*

4.3.1.2 Exemple

D1: SÉRIE ORIGINALE AJUSTÉE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	113.212	111.307	100.706
1986	104.858	99.581	106.502	101.706	97.858	104.660	97.556	69.299	104.039	115.186	111.458	103.012
1987	106.425	104.121	113.019	105.555	102.560	107.105	100.136	70.421	107.535	117.089	115.761	107.634
1988	110.457	107.009	116.147	109.584	107.513	113.046	103.791	74.990	112.972	123.917	119.999	114.886
1989	118.024	113.202	121.703	115.126	109.031	118.488	108.245	77.692	115.768	126.834	124.971	115.585
1990	119.490	113.504	123.800	117.174	113.114	118.404	111.918	81.939	117.941	130.246	125.294	113.575
1991	120.648	113.807	122.354	118.082	111.448	120.103	111.514	81.732	117.477	129.553	126.513	115.922
1992	121.482	115.803	124.131	118.269	112.611	117.390	110.268	82.004	118.021	129.209	123.230	113.114
1993	116.611	111.306	121.069	114.503	110.602	115.845	108.275	79.784	113.144	124.098	120.357	112.924
1994	119.213	112.496	121.333	116.984	113.498	119.255	112.303	82.561	118.887	129.566	126.331	120.194
1995	124.231	117.338	128.072	.	.	.	.	.	.	.	.	.

Ainsi, la valeur du mois d'avril 1986, jugée atypique, devient :
$APR86 = 100*(108.326/106.509) = 101.706$

4.3.2 Tableau D2 : Estimation préliminaire de la Tendance-Cycle

4.3.2.1 Description et mode de calcul

Une nouvelle estimation de la composante Tendance-Cycle est obtenue, comme en B2 et C2, en appliquant aux données du tableau D1 une moyenne mobile centrée simple d'ordre 12.

4.3.2.2 Remarques

- X11-ARIMA88 et X12-ARIMA proposent aussi une moyenne mobile centrée sur 24 termes (Cholette, 1981).
- Les 6 premiers et les 6 derniers points de la série ne sont pas imputés à ce stade du calcul.

4.3.2.3 Exemple

D2: TENDANCE-CYCLE - MOYENNE MOBILE SIMPLE SUR 12 TERMES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	101.023	101.111	101.213	101.375	101.629	102.090	102.522	102.878	103.176
1987	103.385	103.540	103.732	103.957	104.216	104.588	104.948	105.236	105.487	105.785	106.160	106.614
1988	107.013	107.356	107.773	108.284	108.745	109.224	109.841	110.415	110.904	111.366	111.661	111.951
1989	112.363	112.661	112.890	113.128	113.457	113.693	113.784	113.857	113.957	114.130	114.385	114.552
1990	114.702	115.032	115.299	115.532	115.687	115.617	115.582	115.642	115.595	115.572	115.541	115.542
1991	115.596	115.571	115.543	115.495	115.516	115.665	115.798	115.915	116.073	116.154	116.211	116.146
1992	115.981	115.941	115.975	115.983	115.832	115.578	115.258	114.868	114.553	114.268	114.028	113.879
1993	113.732	113.557	113.261	112.845	112.512	112.384	112.485	112.643	112.703	112.818	113.042	113.305
1994	113.615	113.898	114.253	114.720	115.197	115.749	116.261	116.672	117.154	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

Ainsi la valeur d'avril 1986 s'obtient à partir des valeurs du tableau D1 de octobre 1985 à octobre 1986 (6 mois avant et 6 mois après) :

$$
\begin{aligned}
APR86 = & \frac{113.212}{24} + \\
& \frac{111.307 + 100.706 + 104.858 + 99.581 + 106.502 + 101.706 + 97.858 + 104.660 + 97.556 + 69.299 + 104.039}{12} \\
& + \frac{115.186}{24} \\
& = 101.023
\end{aligned}
$$

4.3.3 Tableau D4 : Estimation préliminaire de la composante Saisonnier-Irrégulier modifiée

4.3.3.1 Description et mode de calcul

La composante Tendance-Cycle est enlevée de la série analysée pour obtenir une estimation de la composante Saisonnier-Irrégulier.
On a donc : *D4 = D1 _op D2*

4.3.3.2 Remarque

Là encore, pas d'estimation pour les 6 valeurs de début et les 6 valeurs de fin de la série.

4.3.3.3 Exemple

D4: COMPOSANTE SAISONNIER-IRRÉGULIER MODIFIÉE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	100.677	96.783	103.405	96.233	68.188	101.909	112.353	108.339	99.841
1987	102.940	100.562	108.953	101.538	98.411	102.407	95.415	66.917	101.941	110.686	109.045	100.957
1988	103.218	99.677	107.770	101.201	98.867	103.499	94.492	67.917	101.865	111.269	107.468	102.622
1989	105.039	100.480	107.806	101.766	96.099	104.217	95.132	68.237	101.589	111.131	109.255	100.902
1990	104.175	98.672	107.373	101.422	97.776	102.410	96.830	70.855	102.030	112.696	108.442	98.297
1991	104.371	98.474	105.895	102.240	96.478	103.836	96.301	70.510	101.210	111.535	108.865	99.807
1992	104.743	99.881	107.033	101.971	97.219	101.568	95.671	71.390	103.028	113.075	108.070	99.328
1993	102.531	98.018	106.894	101.470	98.303	103.079	96.257	70.829	100.391	109.998	106.471	99.664
1994	104.927	98.769	106.197	101.974	98.525	103.029	96.596	70.764	101.479	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

La valeur de avril 1986 s'obtient donc simplement : $APR86 = 100*(101.706/101.023) = 100.677$

4.3.4 Tableau D5 : Estimation de la composante saisonnière

4.3.4.1 Description et mode de calcul

Cette estimation est obtenue à partir des valeurs de la composante Saisonnier-Irrégulier du tableau D4. Encore une fois, on procède en deux étapes, comme en B5 et C5 :

- ***Étape 1 : estimation de la composante saisonnière*** avec une moyenne 3x3.

- ***Étape 2 : normalisation des coefficients saisonniers*** par une moyenne mobile centrée sur 12 termes

4.3.4.2 Exemple

L'estimation est faite à partir de la composante Saisonnier-Irrégulier modifiée du tableau D4.

- ***Étape 1 : estimation de la composante saisonnière***.

Les données du tableau précédent sont lissées colonne par colonne (mois par mois), avec une moyenne mobile 3x3, pour aboutir au tableau D5a.
Le facteur saisonnier du mois d'avril 1988 sera donc estimé par :

$$APR88 = \frac{100.677 + 101.538 * 2 + 101.201 * 3 + 101.766 * 2 + 101.422}{9} = 101.367$$

Cette moyenne mobile symétrique peut s'appliquer pour estimer les valeurs des coefficients saisonniers des années 1988 à 1992. Pour le début de la série (années 1986 et 1987) et la fin de la série (années 1993 et 1994), on utilise des moyennes asymétriques prédéfinies. Ainsi, par exemple, en avril 1987, on utilise un point dans le passé, le point courant et deux points dans le futur :

$$APR87 = \frac{100.677 * 7 + 101.538 * 10 + 101.201 * 7 + 101.766 * 3}{27} = 101.252$$

D5a: FACTEURS SAISONNIERS PROVISOIRES (mm 3x3)

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	101.124	97.832	103.016	95.577	67.620	101.914	111.473	108.465	100.811
1987	103.442	100.186	108.259	101.252	97.850	103.150	95.356	67.652	101.874	111.319	108.476	101.093
1988	103.724	100.003	108.042	101.367	97.798	103.284	95.292	68.122	101.844	111.388	108.420	101.080
1989	104.135	99.686	107.617	101.591	97.386	103.413	95.529	68.853	101.745	111.505	108.610	100.590
1990	104.367	99.276	107.147	101.717	97.174	103.156	95.947	69.929	101.842	111.974	108.567	99.807
1991	104.280	99.003	106.800	101.860	97.092	102.973	96.144	70.565	101.747	111.920	108.372	99.471
1992	104.126	98.897	106.694	101.859	97.502	102.664	96.173	70.941	101.755	111.836	107.903	99.425
1993	103.930	98.746	106.638	101.816	97.877	102.758	96.198	70.922	101.448	111.536	107.566	99.553
1994	103.917	98.669	106.636	101.768	98.193	102.779	96.287	70.906	101.323	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

- ***Étape 2 : normalisation des coefficients saisonniers*.**

D5b: MOYENNE MOBILE CENTRÉE SUR 12 MOIS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.			
1986				99.982	99.982	99.982	99.982	99.982	99.982	99.982	99.988	99.994
1987	99.991	99.983	99.983	99.974	99.968	99.981	100.004	100.008	99.992	99.987	99.990	99.994
1988	99.996	100.013	100.032	100.033	100.034	100.031	100.048	100.051	100.021	100.012	100.004	99.993
1989	100.008	100.048	100.074	100.075	100.088	100.075	100.065	100.057	100.021	100.006	100.003	99.983
1990	99.990	100.052	100.101	100.125	100.142	100.108	100.072	100.057	100.031	100.022	100.025	100.014
1991	100.014	100.049	100.072	100.065	100.055	100.033	100.012	100.002	99.993	99.988	100.005	100.010
1992	99.998	100.015	100.031	100.028	100.005	99.983	99.973	99.959	99.950	99.946	99.960	99.979
1993	99.984	99.984	99.971	99.946	99.919	99.910	99.915	99.911	99.908	99.906	99.917	99.931
1994	99.936	99.939	99.933	99.933	99.933	99.933	99.933	99.933	99.933			
1995				.	.	.	.	.	.	.	.	.

Au tableau D5a, on applique une moyenne mobile centrée sur 12 mois. Le premier terme calculable est donc celui de octobre 1986 et le dernier, celui de mars 1994. Ainsi :

$$\begin{aligned} APR90 = & \frac{111.505}{24} \\ & + \frac{108.610+100.590+104.367+99.276+107.147+101.717+97.174+103.156+95.947+69.929+101.842}{12} \\ & + \frac{111.974}{24} \\ & = 100.125 \end{aligned}$$

Les six premières valeurs, de avril à septembre 1986, non calculables avec cette moyenne mobile symétrique, seront prises égales à la première valeur calculable, celle de octobre 1986 (99.982). On procède de même pour la fin de série : la valeur calculée pour mars 1994 (99.933) est répétée pour les six mois suivants. Les coefficients saisonniers normalisés sont alors obtenus en divisant le tableau D5a par le tableau D5b, pour obtenir :

D5: FACTEURS SAISONNIERS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	111.493	108.478	100.817
1986	103.452	100.203	108.278	101.143	97.850	103.034	95.594	67.632	101.932	111.493	108.478	100.817
1987	103.452	100.203	108.278	101.278	97.881	103.170	95.352	67.646	101.882	111.333	108.487	101.100
1988	103.728	99.989	108.008	101.334	97.765	103.252	95.247	68.087	101.823	111.374	108.416	101.088
1989	104.127	99.638	107.537	101.515	97.300	103.335	95.467	68.814	101.724	111.498	108.608	100.607
1990	104.378	99.224	107.039	101.590	97.036	103.045	95.878	69.890	101.811	111.949	108.540	99.793
1991	104.265	98.955	106.723	101.794	97.038	102.940	96.132	70.564	101.755	111.933	108.366	99.461
1992	104.128	98.882	106.661	101.831	97.498	102.681	96.199	70.970	101.806	111.896	107.947	99.445
1993	103.947	98.762	106.669	101.871	97.956	102.851	96.279	70.985	101.541	111.641	107.655	99.622
1994	103.984	98.729	106.707	101.836	98.259	102.848	96.351	70.954	101.391	111.641	107.655	99.622
1995	103.984	98.729	106.707	.	.	.	.	.	.	.	.	.

Avec par exemple : $APR86 = 100*(101.124/99.982) = 101.143$

Les valeurs manquantes d'octobre 1985 à mars 1986 sont obtenues en dupliquant la première valeur calculée pour le mois considéré. De même, pour les valeurs d'octobre 1994 à mars 1995, on duplique la dernière valeur calculée pour le mois considéré.

4.3.5 Tableau D6 : Estimation de la série corrigée des variations saisonnières

4.3.5.1 Description et mode de calcul

Cette estimation s'obtient simplement en retirant à la série du tableau D1, l'estimation de la composante saisonnière du tableau D5 : *D6 = D1 _op D5*

4.3.5.2 Exemple

D6: SÉRIE DÉSAISONNALISÉE PROVISOIRE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	101.542	102.608	99.890
1986	101.359	99.379	98.360	100.557	100.008	101.577	102.052	102.465	102.067	103.312	102.746	102.178
1987	102.874	103.910	104.379	104.223	104.780	103.814	105.017	104.101	105.548	105.171	106.705	106.463
1988	106.487	107.020	107.536	108.142	109.971	109.485	108.970	110.138	110.950	111.262	110.685	113.649
1989	113.347	113.613	113.173	113.408	112.056	114.664	113.385	112.902	113.806	113.754	115.067	114.888
1990	114.479	114.392	115.659	115.340	116.569	114.905	116.730	117.240	115.843	116.344	115.436	113.811
1991	115.714	115.010	114.646	116.001	114.850	116.673	116.001	115.827	115.451	115.742	116.746	116.550
1992	116.665	117.112	116.379	116.143	115.501	114.325	114.625	115.547	115.927	115.472	114.158	113.745
1993	112.183	112.702	113.499	112.400	112.910	112.634	112.459	112.396	111.427	111.158	111.798	113.352
1994	114.646	113.943	113.707	114.875	115.509	115.953	116.556	116.359	117.256	116.055	117.348	120.650
1995	119.472	118.848	120.022	.	.	.	.	.	.	.	.	.

Avec par exemple : $APR86 = 100*(101.706/101.143) = 100.557$

4.3.6 Tableau D7 : Estimation de la composante Tendance-Cycle

4.3.6.1 Description et mode de calcul

Ce tableau présente une estimation de la composante Tendance-Cycle réalisée à partir de la série désaisonnalisée du tableau précédent. La méthodologie est la même que celle suivie pour construire le tableau C7 (voir 4.2.6).

4.3.6.2 Exemple

- ***Étape 1 : choix de la moyenne mobile, calcul du ratio*** $\bar{I}/\bar{C}$

On lisse tout d'abord le tableau D6 par une moyenne mobile de Henderson sur 13 termes dont les coefficients figurent dans le tableau 3 (voir 3.3.1).
On a par exemple pour avril 1990 :

$$
\begin{aligned}
APR90= \; & 113.754*(-0.01935)+115.067*(-0.02786)+114.888*(0.00000)+114.479*(0.06549) \\
& +114.392*(0.14736)+115.659*(0.21434)+115.340*(0.24006)+116.569*(0.21434) \\
& +114.905*(0.14736)+116.730*(0.06549)+117.240*(0.00000)+115.843*(-0.02786) \\
& +116.344*(-0.01935) \qquad = 115.507
\end{aligned}
$$

A cette étape du calcul, on ne se préoccupe pas d'estimer les 6 points non calculables en début et en fin de série. On en déduit une estimation de la tendance (tableau D7a) et, par rapport avec le tableau D6, de la composante irrégulière (tableau D7b):

D7a : TENDANCE-CYCLE (Moyenne Mobile de Henderson sur 13 termes)

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	99.924	100.448	101.139	101.793	102.277	102.510	102.619	102.727	102.919
1987	103.227	103.605	103.961	104.237	104.392	104.458	104.568	104.811	105.176	105.595	105.986	106.358
1988	106.769	107.255	107.783	108.342	108.897	109.373	109.753	110.132	110.620	111.245	111.976	112.620
1989	113.055	113.305	113.399	113.377	113.299	113.273	113.379	113.577	113.823	114.082	114.322	114.577
1990	114.827	115.007	115.221	115.507	115.844	116.190	116.387	116.378	116.190	115.852	115.462	115.118
1991	114.922	114.960	115.161	115.427	115.657	115.784	115.845	115.893	115.954	116.098	116.330	116.583
1992	116.752	116.675	116.343	115.875	115.469	115.236	115.197	115.241	115.170	114.874	114.359	113.735
1993	113.185	112.848	112.740	112.756	112.737	112.564	112.240	111.920	111.807	111.956	112.336	112.886
1994	113.478	114.036	114.508	114.946	115.391	115.783	116.130	116.484	116.891	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

D7b : IRRÉGULIER

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	100.633	99.562	100.433	100.254	100.183	99.568	100.676	100.018	99.280
1987	99.659	100.295	100.402	99.987	100.372	99.383	100.430	99.323	100.354	99.598	100.678	100.099
1988	99.737	99.781	99.771	99.816	100.986	100.102	99.286	100.005	100.298	100.015	98.847	100.914
1989	100.258	100.272	99.800	100.027	98.903	101.228	100.005	99.406	99.985	99.713	100.652	100.271
1990	99.697	99.466	100.380	99.855	100.626	98.894	100.294	100.740	99.702	100.425	99.978	98.865
1991	100.689	100.043	99.553	100.497	99.301	100.767	100.135	99.943	99.567	99.693	100.358	99.972
1992	99.926	100.375	100.031	100.231	100.028	99.209	99.503	100.265	100.658	100.520	99.825	100.008
1993	99.115	99.871	100.673	99.684	100.153	100.062	100.195	100.426	99.660	99.286	99.521	100.413
1994	101.029	99.919	99.300	99.939	100.102	100.147	100.367	99.893	100.313	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

Le schéma étant multiplicatif, on calcule les taux de croissance moyens.

D7c : CROISSANCE, EN VALEUR ABSOLUE, DE LA TENDANCE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC	TOTAL
1985	.	.	.	.	.	.	.	.	.	.	.	.	0.000
1986	.	.	.	.	0.524	0.688	0.646	0.476	0.227	0.106	0.106	0.187	2.961
1987	0.299	0.367	0.343	0.265	0.149	0.064	0.105	0.232	0.348	0.399	0.370	0.350	3.291
1988	0.386	0.456	0.492	0.518	0.513	0.437	0.347	0.345	0.443	0.566	0.657	0.575	5.735
1989	0.387	0.221	0.083	0.019	0.069	0.023	0.093	0.175	0.217	0.227	0.210	0.224	1.948
1990	0.218	0.156	0.186	0.249	0.291	0.299	0.169	0.008	0.161	0.291	0.337	0.298	2.664
1991	0.170	0.033	0.175	0.231	0.199	0.110	0.053	0.041	0.053	0.124	0.199	0.217	1.606
1992	0.146	0.066	0.285	0.401	0.351	0.201	0.034	0.038	0.062	0.257	0.449	0.545	2.835
1993	0.484	0.298	0.095	0.014	0.017	0.154	0.288	0.285	0.100	0.133	0.339	0.489	2.697
1994	0.525	0.491	0.414	0.382	0.388	0.340	0.299	0.305	0.350	.	.	.	3.494
1995	.	.	.	.	.	.	.	.	.	.	.	.	0.000

$$\overline{C} = \frac{2.961+3.291+5.735+1.948+2.664+1.606+2.835+2.697+3.494}{101} = 0.2696$$

D7d : CROISSANCE, EN VALEUR ABSOLUE, DE L'IRRÉGULIER

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC	TOTAL
1985	.	.	.	.	.	.	.	.	.	.	.	.	0.000
1986	.	.	.	.	1.065	0.875	0.178	0.071	0.615	1.113	0.653	0.739	5.308
1987	0.382	0.638	0.107	0.414	0.385	0.985	1.053	1.102	1.038	0.754	1.085	0.575	8.518
1988	0.362	0.044	0.010	0.045	1.173	0.876	0.815	0.724	0.293	0.283	1.168	2.092	7.884
1989	0.650	0.014	0.471	0.227	1.124	2.351	1.208	0.599	0.582	0.272	0.942	0.378	8.817
1990	0.573	0.232	0.920	0.523	0.772	1.721	1.416	0.445	1.031	0.725	0.445	1.113	9.916
1991	1.845	0.641	0.490	0.948	1.190	1.476	0.628	0.191	0.377	0.127	0.668	0.385	8.965
1992	0.047	0.450	0.343	0.200	0.203	0.819	0.297	0.766	0.392	0.137	0.692	0.184	4.527
1993	0.893	0.762	0.804	0.983	0.471	0.091	0.133	0.230	0.762	0.375	0.236	0.896	6.637
1994	0.613	1.099	0.619	0.643	0.164	0.044	0.220	0.473	0.420	.	.	.	4.295
1995	.	.	.	.	.	.	.	.	.	.	.	.	0.000

$$\overline{I} = \frac{5.308+8.518+7.884+8.817+9.916+8.965+4.527+6.637+4.295}{101} = 0.64224$$

donc $\dfrac{\overline{I}}{\overline{C}} = \dfrac{0.64224}{0.2696} = 2.3822$

- ***Étape 2 : lissage de la série CVS par une moyenne mobile de Henderson***

Le ratio étant supérieur à 1 et inférieur à 3.5, on choisit une moyenne mobile de Henderson sur 13 termes dont les coefficients et ceux des moyennes mobiles asymétriques associées figurent dans le tableau 3 (voir 3.3.1). L'estimation de la tendance pour octobre 1985 se fait en utilisant le point courant et six points dans le futur auxquels on applique les coefficients de la moyenne mobile H6_0 du tableau 3.

Oct85 = 101.542(0.42113)+102.608*(0.35315)+99.890*(0.24390)+101.359*(0.11977)*
+99.379(0.01202)+98.360*(-0.05811)+100.557*(-0.09186) =101.743*

Ce qui conduit au tableau D7 :

D7: TENDANCE-CYCLE
valeur de I/C: 2.3822; Une moyenne mobile de HENDERSON sur 13 termes a été choisie

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	101.743	101.327	100.829
1986	100.322	99.896	99.730	99.924	100.448	101.139	101.793	102.277	102.510	102.619	102.727	102.919
1987	103.227	103.605	103.961	104.237	104.392	104.458	104.568	104.811	105.176	105.595	105.986	106.358
1988	106.769	107.255	107.783	108.342	108.897	109.373	109.753	110.132	110.620	111.245	111.976	112.620
1989	113.055	113.305	113.399	113.377	113.299	113.273	113.379	113.577	113.823	114.082	114.322	114.577
1990	114.827	115.007	115.221	115.507	115.844	116.190	116.387	116.378	116.190	115.852	115.462	115.118
1991	114.922	114.960	115.161	115.427	115.657	115.784	115.845	115.893	115.954	116.098	116.330	116.583
1992	116.752	116.675	116.343	115.875	115.469	115.236	115.197	115.241	115.170	114.874	114.359	113.735
1993	113.185	112.848	112.740	112.756	112.737	112.564	112.240	111.920	111.807	111.956	112.336	112.886
1994	113.478	114.036	114.508	114.946	115.391	115.783	116.130	116.484	116.891	117.414	118.042	118.680
1995	119.243	119.691	120.001	.	.	.	.	.	.	.	.	.

4.3.7 Tableau D8 : Estimation de la composante Saisonnier-Irrégulier non modifiée

4.3.7.1 Description et mode de calcul

La composante Tendance-Cycle est enlevée de la série analysée, C19 ou B1 si l'option de correction pour jours ouvrables n'a pas été choisie, pour obtenir une estimation de la composante Saisonnier-Irrégulier. On a donc : *D8 = C19 _op D7* et les valeurs extrêmes sont présentes dans cette série.
Plusieurs tests, paramétrique et non-paramétrique, sur la présence de saisonnalité sont édités.

1. Deux tests de « *saisonnalité stable* » sont proposés, l'un paramétrique et l'autre non-paramétrique, pour trancher entre les deux hypothèses :

   $$\begin{cases} H_0 : m_1 = m_2 = ... = m_k \\ H_1 : m_i \neq m_j \text{ pour au moins un couple } (i, j) \end{cases}$$

   où $m_1, m_2, ..., m_k$ sont les facteurs saisonniers stables pour les 12 mois ou les 4 trimestres.
   - Le premier, dit test de saisonnalité stable, est un test paramétrique basé sur un modèle d'analyse de la variance à un facteur en tout point similaire à celui édité après le tableau B1 (voir 4.1.3.1).
   - Le second est un test non-paramétrique de Kruskal-Wallis : les données du tableau D8 sont supposées provenir de *k* échantillons indépendants $A_1, A_2, ..., A_k$ (4 si la série est trimestrielle, 12 si elle est mensuelle) de tailles respectives $n_1, n_2, ..., n_k$. Le test est fondé sur la statistique $H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{S_i^2}{n_i} - 3(N+1)$ où S_i est la somme des rangs des observations de l'échantillon A_i dans la série des $N = \sum_{i=1}^{k} n_i$ observations. Cette quantité, sous l'hypothèse nulle, suit une loi du Chi-deux à *k-1* degrés de liberté.

2. Vient ensuite un test dit de « saisonnalité évolutive » qui est un test basé sur un modèle d'analyse de la variance à deux facteurs (le mois ou le trimestre, et l'année) proposé par Higginson (1975).
   Ce test repose sur la modélisation des valeurs de la composante Saisonnier-Irrégulier du tableau D8 suivante :
   $\left| SI_{ij} - \_xbar \right| = X_{ij} = m_i + b_j + e_{ij}$, où
   - m_i désigne l'effet du mois, ou trimestre *i (i=1 …k)*
   - b_j désigne l'effet de l'année *j (j=1….N)*
   - e_{ij} représente l'effet résiduel, réalisation de lois indépendantes et identiquement distribuées (*i.i.d*), de moyenne nulle.

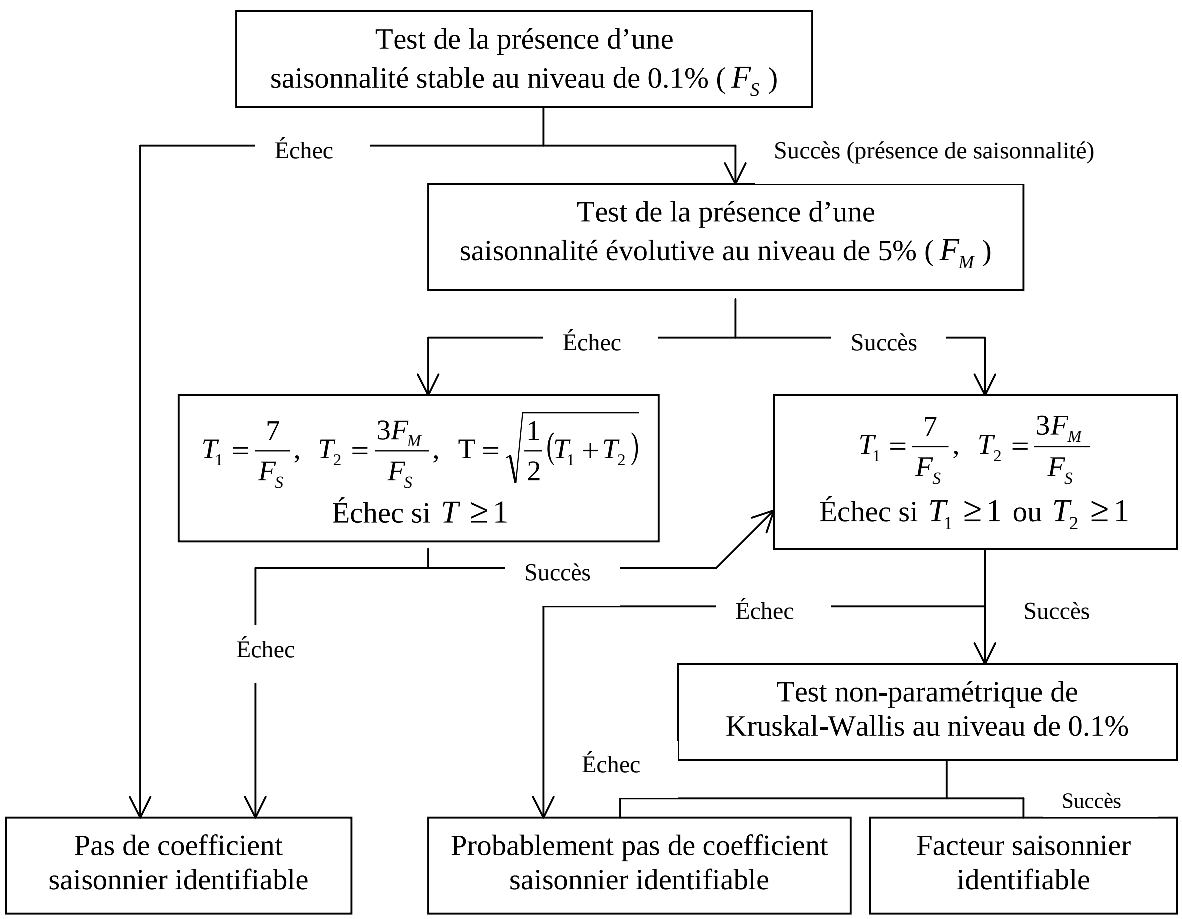
   Le test est basé sur la décomposition $S^2 = S_A^2 + S_B^2 + S_R^2$ où :
   - S^2 est la somme des carrés totale : $S^2 = \sum_{i=1}^{k} \sum_{j=1}^{N} \left(X_{ij} - \overline{X}_{..} \right)^2$ avec $\overline{X}_{..} = \frac{1}{kN} \sum_{i=1}^{k} \sum_{j=1}^{N} X_{ij}$
   - S_A^2 est la somme des carrés « inter mois » : $S_A^2 = N \sum_{i=1}^{k} \left(\overline{X}_{i.} - \overline{X}_{..} \right)^2$ avec $\overline{X}_{i.} = \frac{1}{N} \sum_{j=1}^{N} X_{ij}$

- S_B^2 est la somme des carrés « inter année » : $S_B^2 = k\sum_{j=1}^{N}\left(\overline{X}_{.j} - \overline{X}_{..}\right)^2$ avec $\overline{X}_{.j} = \frac{1}{k}\sum_{i=1}^{k} X_{ij}$
- S_R^2 est la somme des carrés « résiduelle » : $S_R^2 = \sum_{i=1}^{k}\sum_{j=1}^{N}\left(X_{ij} - \overline{X}_{i.} - \overline{X}_{.j} + \overline{X}_{..}\right)^2$.

L'hypothèse nulle $(H_0): b_1 = b_2 = ... = b_k$, c'est à dire que la saisonnalité n'évolue pas au cours des années, peut être testée grâce à la statistique $F_M = \dfrac{S_B^2/(N-1)}{S_R^2/(N-1)(k-1)}$ qui suit , sous H_0, une loi de Fisher-Snedecor à *(N-1)* et *(k-1)(N-1)* degrés de liberté.

3. Enfin, ces tests sont complétés par un test dit de « présence d'une saisonnalité identifiable » construit à partir des valeurs des statistiques de Fisher des tests paramétrique de « *saisonnalité stable* » (statistique F_S) et de « *saisonnalité évolutive* » (statistique F_M) évoqués ci-dessus. Ce test a été élaboré, à partir de considérations théoriques et pratiques, par Lothian et Morry (1978). La valeur de la statistique de test T est : $T_1 = \dfrac{7}{F_S}$, $T_2 = \dfrac{3F_M}{F_S}$, $\mathrm{T} = \sqrt{\dfrac{1}{2}(T_1 + T_2)}$

La version finale du test combiné s'interprète alors selon le diagramme suivant :



4.3.7.2 Exemple

D8: SÉRIE BRUTE DÉBARRASSÉE DE SA TENDANCE (estimation Saisonnier Irrégulier)

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	111.272	109.850	99.878
1986	104.521	99.685	106.791	108.408	97.421	103.481	95.838	65.882	101.491	112.247	108.498	100.090
1987	97.516	100.498	108.713	101.265	98.245	102.534	95.762	67.188	102.243	110.885	109.223	101.200
1988	103.455	99.770	107.760	101.147	98.729	103.358	94.567	68.091	102.126	108.636	107.165	102.012
1989	104.396	99.909	104.265	104.116	96.233	104.695	95.472	68.405	101.709	111.178	109.316	100.880
1990	104.061	98.694	107.446	101.443	97.643	101.905	96.160	70.407	101.507	112.424	108.516	98.660
1991	104.983	98.997	106.246	102.300	96.361	103.730	98.207	70.524	101.314	111.589	108.754	99.434
1992	104.051	99.253	106.694	102.066	97.525	101.869	95.721	71.158	102.476	112.478	107.757	97.905
1993	103.027	101.119	107.388	99.842	98.106	102.914	96.468	71.287	101.196	110.845	107.140	100.034
1994	105.054	98.649	105.960	101.774	98.899	102.998	96.705	72.340	101.708	110.349	107.022	101.276
1995	104.183	98.034	106.726	.	.	.	.	.	.	.	.	.

La valeur d'avril 1986 s'obtient donc simplement : $APR86 = 100*(108.326/99.924) = 108.408$

TEST DE LA PRÉSENCE D'UNE SAISONNALITÉ STABLE

	Somme des Carrés	D.D.L	Moyenne Quadratique	F-Value	PROB>F
Entre les Mois	11264.919	11	1024.084	498.194	0.000
Résidu	209.670	102	2.056		
TOTAL	11474.589	113			

TEST NON-PARAMÉTRIQUE DE LA PRÉSENCE D'UNE SAISONNALITÉ STABLE

Statistique de Kruskal-Wallis	D.D.L	PROB>F
104.780	11	0.000

TEST DE LA PRÉSENCE D'UNE SAISONNALITÉ ÉVOLUTIVE

	Somme des Carrés	D.D.L	Moyenne Quadratique	F-Value	PROB>F
Entre les Années	20.628	8	2.578	1.724	0.104
Résidu	131.614	88	1.496		

Nous ne détaillerons pas les calculs complexes conduisant aux valeurs des différentes statistiques de test. Par contre, explicitons le test combiné :

1. La statistique F_S est suffisamment grande pour rejeter l'hypothèse d'égalité des coefficients saisonniers.
2. La statistique F_M montre que la série ne comporte aucune composante saisonnière évolutive : le test accepte l'hypothèse H_0.
3. On a $T_1 = \dfrac{7}{F_S} = \dfrac{7}{498.194} = 0.014$ et $T_2 = \dfrac{3F_M}{F_S} = \dfrac{3*1.724}{498.194} = 0.010$ et donc ces deux statistiques sont plus petites que 1.
4. Le test non-paramétrique de Kruskal-Wallis confirme l'existence d'une saisonnalité

Le test combiné statuera donc sur la présence d'une saisonnalité identifiable.

4.3.8 Tableau D9 : Valeurs de remplacement pour les points atypiques de la composante Saisonnier-Irrégulier.

4.3.8.1 Description et mode de calcul

Le tableau D8 nous donne une estimation de la composante Saisonnier-Irrégulier "brute" puisqu'elle est issue de la série C19 corrigée des effets pour jours ouvrables (ou de la série B1 si aucune correction n'a été demandée).
On peut utiliser, à la place de la série C19 (ou B1), la série D1 corrigée à la fois des valeurs atypiques et des effets de jours ouvrables. La comparaison de ces deux estimations de la composante Saisonnier-Irrégulier (corrigée ou non des valeurs atypiques) nous permet de repérer les valeurs de remplacement utilisées pour les points jugés atypiques. Ces valeurs sont alors consignées dans le tableau D9.
On compare donc les séries *C19 _op D7* et *D1 _op D7*.

4.3.8.2 Exemple

On enlève à la série D1 l'estimation de la tendance obtenue en D7.

D9bis: VALEURS FINALES DES S-I corrigés

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	111.272	109.850	99.878
1986	104.521	99.685	106.791	101.783	97.421	103.481	95.838	67.756	101.491	112.247	108.498	100.090
1987	103.098	100.498	108.713	101.265	98.245	102.534	95.762	67.188	102.243	110.885	109.223	101.200
1988	103.455	99.770	107.760	101.147	98.729	103.358	94.567	68.091	102.126	111.390	107.165	102.012
1989	104.396	99.909	107.322	101.542	96.233	104.604	95.472	68.405	101.709	111.178	109.316	100.880
1990	104.061	98.694	107.446	101.443	97.643	101.905	96.160	70.407	101.507	112.424	108.516	98.660
1991	104.983	98.997	106.246	102.300	96.361	103.730	96.261	70.524	101.314	111.589	108.754	99.434
1992	104.051	99.253	106.694	102.066	97.525	101.869	95.721	71.158	102.476	112.478	107.757	99.453
1993	103.027	98.634	107.388	101.549	98.106	102.914	96.468	71.287	101.196	110.845	107.140	100.034
1994	105.054	98.649	105.960	101.774	98.359	102.998	96.705	70.878	101.708	110.349	107.022	101.276
1995	104.183	98.034	106.726	.	.	.	.	.	.	.	.	.

Avec par exemple : $APR86 = 100*(101.706 / 99.924) = 101.783$.

Cette table est comparée au tableau D8 ; les valeurs qui ne coïncident pas correspondent aux valeurs de remplacement des points jugés atypiques et sont éditées dans le tableau D9

D9: VALEURS FINALES DE REMPLACEMENT DES S-I EXTRÊMES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	101.783	.	.	.	67.756	.	.	.	.
1987	103.098	.	.	.	.	.	.	.	.	.	.	.
1988	.	.	.	.	.	.	.	.	.	111.390	.	.
1989	.	.	107.322	101.542	.	104.604	.	.	.	.	.	.
1990	.	.	.	.	.	.	.	.	.	.	.	.
1991	.	.	.	.	.	.	96.261	.	.	.	.	.
1992	.	.	.	.	.	.	.	.	.	.	.	99.453
1993	.	98.634	.	101.549	.	.	.	.	.	.	.	.
1994	.	.	.	.	98.359	.	.	70.878	.	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

4.3.9 Tableau D9A : Calcul des "Ratios de Saisonnalité Mobile" (RSM).

4.3.9.1 Description et mode de calcul

A partir du tableau non édité D9bis précédent, on va essayer de mesurer l'importance de la composante irrégulière par rapport à la composante saisonnière et ce, pour chaque mois. L'idée est que nous allons avoir à lisser cette composante pour estimer la composante saisonnière ; selon l'importance de l'irrégulier dans la composante Saisonnier-Irrégulier, nous serons amenés à utiliser des moyennes mobiles plus ou moins longues. Nous avons déjà rencontré cette idée dans le programme d'extraction de la Tendance-Cycle où le ratio $\bar{I}/\bar{C}$ servait à choisir la longueur de la moyenne mobile de Henderson. Le programme procède en plusieurs étapes.

- ***Étape 1 : estimation des composantes irrégulière et saisonnière***

Une estimation de la composante saisonnière est obtenue en lissant, mois par mois donc colonne par colonne, le tableau D9bis par une moyenne mobile simple sur 7 termes, donc de coefficients $\left\{\frac{1}{7},\frac{1}{7},\frac{1}{7},\frac{1}{7},\frac{1}{7},\frac{1}{7},\frac{1}{7}\right\}$. Pour ne pas perdre 3 points au début et à la fin de chaque colonne, chacune d'entre elle va être fictivement complétée. Supposons que la colonne correspondant au mois soit composée de N valeurs $\{x_1, x_2, x_3, \ldots, x_{N-1}, x_N\}$. Elle va être transformée en une série $\{x_{-2}, x_{-1}, x_0, x_1, x_2, x_3, \ldots, x_{N-1}, x_N, x_{N+1}, x_{N+2}, x_{N+3}\}$ avec :

$x_{-2} = x_{-1} = x_0 = \frac{x_1 + x_2 + x_3}{3}$ et $x_{N+1} = x_{N+2} = x_{N+3} = \frac{x_N + x_{N-1} + x_{N-2}}{3}$

On a donc alors les estimations souhaitées : $S = M_7(D9bis)$ et $I = D9bis$ _op S

- ***Étape 2 : Calcul des "Ratios de Saisonnalité Mobile"***

Pour chaque mois i on s'intéresse aux moyennes des évolutions annuelles de chaque composante en calculant :

$\bar{S}_i = \frac{1}{N_i - 1}\sum_{t=2}^{N_i}\left|S_{i,t} \;\_op\; S_{i,t-1} - \_xbar\right|$, $\bar{I}_i = \frac{1}{N_i - 1}\sum_{t=2}^{N_i}\left|I_{i,t} \;\_op\; I_{i,t-1} - \_xbar\right|$, où N_i désigne le nombre de mois i dans les données, et le ratio de saisonnalité mobile du mois i : $RSM_i = \bar{I}_i / \bar{S}_i$. Ces ratios sont publiés dans le tableau D9A

- ***Étape 3 : Calcul du "Ratio de Saisonnalité Mobile" global***

Un Ratio de Saisonnalité Mobile global est calculé de la façon suivante :

$$RSM = \frac{\sum_i N_i * \bar{I}_i}{\sum_i N_i * \bar{S}_i}$$

4.3.9.2 Remarques

- Ces Ratios servent à mesurer l'importance du "bruit" dans la composante Saisonnier-Irrégulier. L'idée de base est bien d'avoir, pour chaque mois, un indicateur permettant de choisir une moyenne mobile appropriée pour enlever le bruit et obtenir une bonne estimation du coefficient saisonnier. Plus le ratio est élevé, plus la série est "chaotique" et plus l'ordre de la moyenne mobile à utiliser doit être important. Pour la suite, le programme par défaut choisit la même moyenne mobile pour lisser chaque série mensuelle mais on a ici des éléments permettant de choisir une moyenne mobile pour chaque mois.

- En fait ce n'est pas tout à fait aussi simple et une correction est apportée au niveau mensuel au calcul des $\bar{S}_i$ et $\bar{I}_i$. Ces quantités sont multipliées respectivement par des paramètres CS_i et FIS_i dont la valeur dépend du nombre d'années disponibles pour chaque mois : le calcul théorique et la justification de ces constantes se trouve dans Lothian (1978). On trouvera ci-après les valeurs de ces constantes selon le nombre d'années *nyear* de la colonne *i* :

nyear	CS	FIS
4	3	$\dfrac{90}{2\sqrt{842}+21\sqrt{2}}$
5	$\dfrac{3\sqrt{2}}{1+\sqrt{3}}$	$\dfrac{60}{\sqrt{894}+2\sqrt{211}}$
6	$\dfrac{5\sqrt{6}}{8+\sqrt{2}}$	$\dfrac{25\sqrt{3}}{2\sqrt{298}+\sqrt{67}}$
7 et plus	$\dfrac{\sqrt{3}*nyear}{6\sqrt{2}+(nyear-6)\sqrt{3}}$	$\dfrac{5\sqrt{6}*nyear}{6\sqrt{149}+5\sqrt{6}(nyear-6)}$

4.3.9.3 Exemple

A partir du tableau D9bis, nous allons détailler le calcul pour le mois de mai par exemple. La composante Saisonnier-Irrégulier pour ce mois est celle de la colonne 1 du tableau suivant :

ANNÉE	Colonne 1	Colonne 2	MM7
		98.132	
		98.132	
1985	.	98.132	.
1986	97.421	97.421	97.861
1987	98.245	98.245	97.791
1988	98.729	98.729	97.538
1989	96.233	96.233	97.451
1990	97.643	97.643	97.549
1991	96.361	96.361	97.565
1992	97.525	97.525	97.461
1993	98.106	98.106	97.713
1994	98.359	98.359	97.763
1995	.	97.997	.
		97.997	
		97.997	

La moyenne des 3 premières années est : $\dfrac{97.421+98.245+98.729}{3}=98.132$

Celle des 3 dernières années : $\dfrac{97.525+98.106+98.359}{3}=97.997$

La série à traiter est donc celle de la colonne 2 du tableau précédent. En lissant par une moyenne mobile d'ordre 7, on a, par exemple pour le point de l'année 1986 :

$$MAY86 = \frac{98.132 + 98.132 + 98.132 + 97.421 + 98.245 + 98.729 + 96.233}{7} = 97.861$$

Ces calculs conduisent à l'estimation des facteurs saisonniers du tableau suivant.

FACTEURS SAISONNIERS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	111.457	108.901	100.621
1986	103.792	99.974	107.693	101.419	97.861	103.335	95.401	67.782	101.919	111.415	108.919	100.691
1987	103.845	99.789	107.649	101.425	97.791	103.161	95.511	68.172	101.855	111.552	108.823	100.444
1988	104.029	99.648	107.433	101.554	97.538	103.248	95.636	68.579	101.763	111.569	108.760	100.308
1989	104.081	99.544	107.282	101.650	97.451	103.068	95.683	69.076	101.838	111.742	108.461	100.247
1990	103.867	99.394	107.367	101.616	97.549	102.988	95.773	69.580	101.796	111.541	108.267	100.239
1991	104.147	99.130	106.974	101.689	97.565	103.054	95.908	70.107	101.719	111.465	107.953	100.250
1992	104.251	98.882	106.826	101.782	97.461	102.945	96.155	70.538	101.672	111.441	107.973	99.999
1993	104.207	98.672	106.736	101.818	97.713	102.658	96.273	70.924	101.684	111.448	107.686	99.909
1994	104.210	98.635	106.628	101.868	97.763	102.756	96.293	71.024	101.724	111.276	107.513	100.137
1995	104.083	98.555	106.692	.	.	.	.	.	.	.	.	.

En enlevant cette estimation du tableau D9bis, on obtient une estimation de la composante irrégulière. Par exemple : $APR86 = 100 * (101.783 / 101.419) = 100.359$

IRRÉGULIER

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	99.834	100.871	99.261
1986	100.703	99.711	99.162	100.359	99.551	100.141	100.458	99.961	99.581	100.746	99.614	99.403
1987	99.281	100.711	100.989	99.842	100.465	99.391	100.263	98.557	100.381	99.402	100.368	100.753
1988	99.448	100.122	100.304	99.599	101.221	100.106	98.883	99.289	100.357	99.840	98.534	101.699
1989	100.303	100.367	100.038	99.895	98.750	101.490	99.779	99.029	99.873	99.496	100.788	100.631
1990	100.187	99.296	100.073	99.830	100.097	98.949	100.404	101.189	99.717	100.792	100.230	98.425
1991	100.803	99.867	99.320	100.601	98.765	100.656	100.369	100.594	99.601	100.111	100.742	99.186
1992	99.808	100.375	99.877	100.279	100.066	98.955	99.549	100.879	100.791	100.931	99.800	99.455
1993	98.868	99.962	100.611	99.736	100.403	100.250	100.202	100.511	99.520	99.459	99.493	100.125
1994	100.809	100.014	99.374	99.907	100.610	100.236	100.428	99.794	99.983	99.167	99.543	101.137
1995	100.097	99.471	100.032	.	.	.	.	.	.	.	.	.

Colonne par colonne, on calcule ensuite les variations annuelles moyennes de l'irrégulier et de la composante saisonnière selon les formules correspondant au schéma multiplicatif :

$$\bar{S}_i = \frac{1}{N-1}\sum_{t=2}^{N}\left|S_{i,t}/S_{i,t-1} - 1\right| = \frac{1}{N-1}\sum_{t=2}^{N}\left|\frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1}}\right| \quad \text{et} \quad \bar{I}_i = \frac{1}{N-1}\sum_{t=2}^{N}\left|I_{i,t}/I_{i,t-1} - 1\right| = \frac{1}{N-1}\sum_{t=2}^{N}\left|\frac{I_{i,t} - I_{i,t-1}}{I_{i,t-1}}\right|$$

Les accroissements annuels, en pourcentage sont :

TAUX DE CROISSANCE DES FACTEURS SAISONNIERS EN VALEUR ABSOLUE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1986	.	.	.	.	.	.	.	.	.	0.0371	0.0165	0.0696
1987	0.0509	0.1844	0.0410	0.0063	0.0713	0.1685	0.1155	0.5752	0.0625	0.1226	0.0884	0.2454
1988	0.1776	0.1413	0.2002	0.1270	0.2588	0.0839	0.1305	0.5963	0.0898	0.0155	0.0572	0.1360
1989	0.0493	0.1049	0.1410	0.0939	0.0889	0.1737	0.0496	0.7249	0.0733	0.1544	0.2748	0.0604
1990	0.2051	0.1508	0.0795	0.0329	0.1004	0.0785	0.0941	0.7302	0.0415	0.1792	0.1790	0.0081
1991	0.2689	0.2658	0.3663	0.0715	0.0167	0.0645	0.1407	0.7575	0.0751	0.0686	0.2904	0.0108
1992	0.0999	0.2502	0.1381	0.0912	0.1072	0.1059	0.2577	0.6148	0.0468	0.0213	0.0186	0.2505
1993	0.0422	0.2124	0.0844	0.0356	0.2585	0.2789	0.1227	0.5473	0.0119	0.0059	0.2659	0.0893
1994	0.0037	0.0369	0.1010	0.0495	0.0517	0.0959	0.0204	0.1411	0.0401	0.1538	0.1605	0.2280
1995	0.1226	0.0809	0.0596	.	.	.	.	.	.	.	.	.

Avec par exemple pour février 1988 : $FEB88 = 100 * \left|\frac{99.648 - 99.789}{99.789}\right| = 0.1413$

CROISSANCE DE L'IRRÉGULIER EN VALEUR ABSOLUE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1986	.	.	.	.	.	.	.	.	.	0.9131	1.2463	0.1430
1987	1.4115	1.0026	1.8419	0.5154	0.9177	0.7482	0.1942	1.4050	0.8034	1.3343	0.7568	1.3578
1988	0.1679	0.5841	0.6776	0.2434	0.7531	0.7193	1.3761	0.7425	0.0241	0.4405	1.8274	0.9392
1989	0.8598	0.2446	0.2659	0.2969	2.4415	1.3817	0.9062	0.2614	0.4821	0.3446	2.2876	1.0502
1990	0.1157	1.0674	0.0357	0.0648	1.3638	2.5036	0.6266	2.1809	0.1564	1.3025	0.5536	2.1928
1991	0.6148	0.5749	0.7530	0.7724	1.3301	1.7250	0.0354	0.5876	0.1158	0.6751	0.5114	0.7735
1992	0.9865	0.5093	0.5606	0.3196	1.3171	1.6898	0.8170	0.2832	1.1944	0.8188	0.9352	0.2711
1993	0.9421	0.4118	0.7348	0.5417	0.3364	1.3088	0.6563	0.3646	1.2607	1.4584	0.3080	0.6735
1994	1.9633	0.0523	1.2294	0.1714	0.2061	0.0140	0.2256	0.7141	0.4654	0.2936	0.0508	1.0115
1995	0.7066	0.5430	0.6625	.	.	.	.	.	.	.	.	.

Avec par exemple pour janvier 1988 : $JAN88 = 100 * \left| \frac{99.448 - 99.281}{99.281} \right| = 0.168$

Pour calculer le Ratio de Saisonnalité Mobile du mois d'avril par exemple, on a, avec 8 années d'observations :

$$CS_{APR} = \frac{8*1.732051}{8.485281 + (8\text{-}6)*1.732051} = 1.1596 \text{ et } FIS_{APR} = \frac{8*12.247449}{73.239334 + (8\text{-}6)*12.247449} = 1.0025$$

$$\bar{I}_{APR} = \frac{0.5154 + 0.2434 + 0.2969 + 0.0648 + 0.7724 + 0.3196 + 0.5417 + 0.1714}{8} * FIS_{APR} \approx 0.3657 * 1.0025 = 0.3666$$

et

$$\bar{S}_{APR} = \frac{0.0063 + 0.1270 + 0.0939 + 0.0329 + 0.0715 + 0.0912 + 0.0356 + 0.0495}{8} * CS_{APR} = 0.0635 * 1.1596 = 0.0736$$

d'où finalement : $RSM_{APR} = \frac{0.367}{0.075} \approx 4.98$

Les résultats complets sont consignés dans le tableau D9A :

D9A: VARIATIONS SUR UN AN DES COMPOSANTES IRRÉGULIÈRE ET SAISONNIÈRE ET RAPPORT DE SAISONNALITÉ ÉVOLUTIVE

	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
I	0.8651	0.5557	0.7529	0.3666	1.0859	1.2645	0.6062	0.8195	0.5642	0.8442	0.9440	0.9368
S	0.1292	0.1807	0.1533	0.0736	0.1382	0.1522	0.1350	0.6794	0.0639	0.0960	0.1711	0.1390
RATIO	6.697	3.075	4.911	4.979	7.858	8.310	4.491	1.206	8.826	8.790	5.518	6.739

4.3.10 Tableau D10 : Estimation finale des facteurs saisonniers

4.3.10.1 Description et mode de calcul

- ***Étape 1 : Calcul du Ratio de Saisonnalité Mobile global***

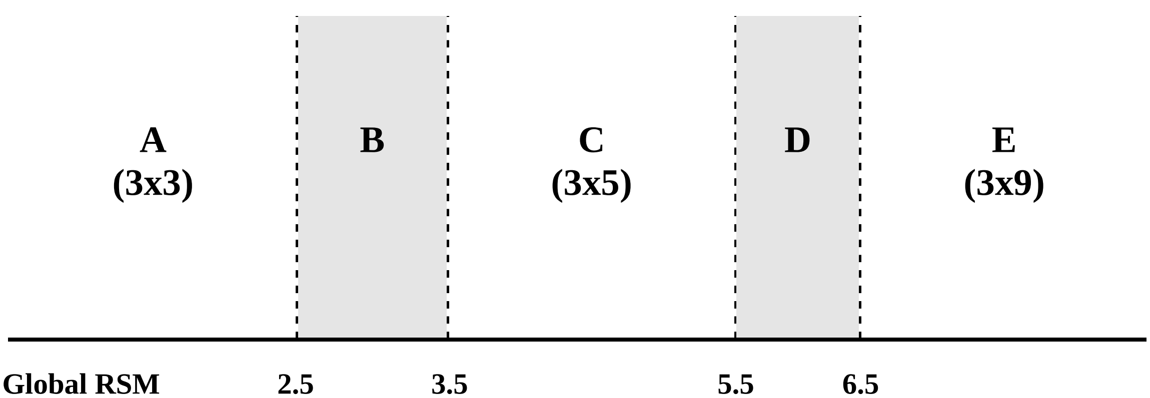
Pour calculer un ratio de saisonnalité mobile global, le programme utilise l'ensemble des données disponibles jusqu'à la dernière année complète, ceci pour éviter en cours d'année, de devoir changer le filtre utilisé. Il calcule, pour chaque mois, les valeurs moyennes des évolutions des composantes saisonnière et irrégulière selon la méthode décrite précédemment. Le ratio global s'en déduit en pondérant ces quantités par le nombre de mois de chaque type :

$$RSM = \frac{\sum_i N_i * \bar{I}_i}{\sum_i N_i * \bar{S}_i}$$

- ***Étape 2 : choix d'une moyenne mobile et estimation de la composante saisonnière***

Selon la valeur du ratio, le programme va sélectionner automatiquement une moyenne mobile qui sera appliquée, colonne par colonne (*i.e.* mois par mois) à la composante Saisonnier-Irrégulier du tableau D8 modifiée, pour les points atypiques, par les valeurs du tableau D9.

Figure 19 : Critères de sélection de la moyenne mobile saisonnière



Le choix de la moyenne mobile repose, dans X11 ARIMA, sur la stratégie suivante :

- Un Ratio de Saisonnalité Mobile global est calculé de la manière indiquée au tableau D9A.
- Si ce ratio tombe dans la zone A (RSM < 2.5), une moyenne mobile 3x3 est utilisée, s'il est dans la zone C (3.5 < RSM < 5.5), une moyenne mobile 3x5 est sélectionnée et s'il est dans la zone E (RSM > 6.5), c'est une moyenne mobile 3x9 qui sera utilisée.
- Si le ratio RSM tombe dans la zone B ou D, on enlève une année d'observations et on recalcule le RSM. S'il retombe dans les zones B ou D, on recommence ... et ainsi de suite en enlevant au maximum 5 années d'observations. Si cela ne marche pas, c'est à dire que l'on tombe encore dans les zones B ou D, on choisit une 3x5.

La moyenne mobile symétrique choisie porte selon le cas sur 5 (3x3), 7 (3x5) ou 11 (3x9) termes et ne permet donc pas d'estimer les valeurs des coefficients saisonniers des 2 (ou 3 ou 5) premières et 2 (ou 3 ou 5) dernières années. Celles-ci sont alors calculées grâce à des moyennes mobiles asymétriques *ad hoc*. On obtient alors une série de coefficients provisoires *fspro*

- ***Étape 3 : Normalisation des coefficients saisonniers***

Ces coefficients saisonniers provisoires sont alors normalisés à l'aide d'une moyenne mobile centrée sur 12 termes.

- ***Étape 4 : Prévision des coefficients saisonniers***

Une prévision sur un an des facteurs saisonniers est alors calculée par simple projection linéaire à partir des deux derniers coefficients saisonniers d'un mois donné. Si N est la dernière année disponible pour un mois donné, on a donc :

$$fs_{N+1} = fs_N + \frac{1}{2}\left(fs_N - fs_{N-1}\right) = \frac{3 * fs_N - fs_{N-1}}{2}$$

4.3.10.2 Remarques

- Un problème pourrait intervenir dans l'application de ces moyennes mobiles asymétriques si on ne dispose pas d'assez d'années d'observations. Supposons que nous n'ayons, pour un mois donné, que 5 années d'observations pour une moyenne mobile 3x5. Le point central ne peut être estimé puisqu'on ne dispose ni de 3 points dans le futur, ni de 3 points dans le passé ce qui serait nécessaire pour pouvoir utiliser une moyenne mobile asymétrique. Dans ce cas, il sera estimé par la moyenne simple des 5 observations disponibles. A chaque fois qu'on rencontre un tel problème d'estimation, avec par exemple une 3x9 et moins de 11 années d'observations, on utilise la moyenne simple des observations disponibles.
- Il est toujours possible de fixer soit même la moyenne mobile à utiliser. Dans ce cas, X11-ARIMA88 permet de choisir entre une moyenne mobile simple sur 3 termes, une 3x3, une 3x5, une 3x9 et une saisonnalité stable. X12-ARIMA propose en plus une 3x15.

4.3.10.3 Exemple

L’analogue du tableau D9A calculé sur les données disponibles jusqu’en décembre 1994 (la dernière année complète disponible), est le suivant :

	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
I	0.88256	0.54414	0.76503	0.36663	1.08594	1.26447	0.60619	0.81945	0.56420	0.84420	0.94401	0.93682
S	0.12803	0.16754	0.16818	0.07363	0.13820	0.15217	0.13498	0.67941	0.06392	0.09604	0.17107	0.13902
RATIO	6.89366	3.24787	4.54879	4.97930	7.85791	8.30979	4.49105	1.20612	8.82600	8.79035	5.51829	6.73894

Comme on peut le constater, seules les valeurs des 3 premiers mois, ceux dont les dernières valeurs ont été exclues, sont modifiées. Le ratio global s’en déduit aisément :

$$RSM = \frac{\sum_i N_i * \bar{I}_i}{\sum_i N_i * \bar{S}_i}$$

$$= \frac{8*(0.88256+0.54414+0.76503+0.36663+1.08594+1.26447+0.60619+0.81945+0.56420)+9*(0.84420+0.94401+0.93682)}{8*(0.12803+0.16754+0.16818+0.07363+0.13820+0.15217+0.13498+0.67941+0.06392)+9*(0.09604+0.17107+0.13902)}$$

$$= \frac{79.71415}{17.30365} = 4.6068$$

Le Ratio de Saisonnalité Mobile global calculé est de 4.61. Il est donc directement dans la zone C et on va choisir une moyenne mobile 3x5 (voir 3.3.3, tableau 4) pour lisser la composante Saisonnier-Irrégulier corrigée du tableau D9bis (voir 4.3.8.2) et obtenir ainsi le tableau D10bis.
Le facteur saisonnier du mois d'avril 1989 sera donc estimé par :

$$APR89 = \frac{101.783 + 2*101.265 + 3*101.147 + 3*101.542 + 3*101.443 + 2*102.300 + 102.066}{15} = 101.558$$

Pour le début de la série (années 1986 à 1988) et la fin de la série (années 1992 à 1994), on utilise des moyennes asymétriques prédéfinies, par exemple :

$$APR87 = \frac{15*101.783 + 15*101.265 + 15*101.147 + 11*101.542 + 4*101.443}{60} = 101.428$$

(un point dans le passé, le point courant et trois points dans le futur)

D10bis: FACTEURS SAISONNIERS PROVISOIRES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	111.456	108.887	100.633
1986	103.797	99.973	107.690	101.420	97.847	103.346	95.401	67.787	101.917	111.434	108.827	100.720
1987	103.845	99.885	107.655	101.428	97.751	103.314	95.456	67.994	101.879	111.489	108.679	100.725
1988	103.924	99.750	107.584	101.470	97.585	103.299	95.538	68.384	101.829	111.504	108.650	100.606
1989	104.031	99.537	107.399	101.558	97.465	103.165	95.614	69.003	101.807	111.643	108.480	100.364
1990	104.097	99.332	107.204	101.673	97.305	103.108	95.766	69.666	101.749	111.669	108.398	100.072
1991	104.176	99.089	106.953	101.769	97.357	102.927	95.972	70.308	101.702	111.684	108.145	99.850
1992	104.199	98.886	106.790	101.843	97.453	102.855	96.189	70.723	101.651	111.518	107.934	99.838
1993	104.191	98.734	106.681	101.865	97.673	102.756	96.282	70.954	101.686	111.371	107.652	99.998
1994	104.141	98.626	106.662	101.872	97.751	102.764	96.292	71.020	101.721	111.279	107.524	100.131
1995	104.082	98.561	106.692	.	.	.	.	.	.	.	.	.

Ces facteurs sont alors normalisés en appliquant une moyenne mobile centrée d'ordre 12 qui conduit au tableau D10ter suivant :

D10ter: FACTEURS DE NORMALISATION (mm12)

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.012	100.012	100.012
1986	100.012	100.012	100.012	100.012	100.009	100.010	100.015	100.014	100.009	100.007	100.004	99.998
1987	99.999	100.010	100.017	100.018	100.014	100.008	100.012	100.009	100.001	100.000	99.994	99.987
1988	99.990	100.009	100.023	100.022	100.021	100.015	100.015	100.010	99.994	99.990	99.988	99.978
1989	99.975	100.004	100.029	100.034	100.033	100.016	100.008	100.003	99.986	99.983	99.981	99.972
1990	99.976	100.009	100.035	100.033	100.031	100.015	100.006	100.000	99.979	99.973	99.979	99.973
1991	99.974	100.010	100.035	100.033	100.023	100.004	99.995	99.988	99.973	99.969	99.976	99.977
1992	99.983	100.009	100.025	100.016	100.000	99.991	99.990	99.983	99.972	99.968	99.979	99.984
1993	99.983	99.997	100.008	100.003	99.985	99.980	99.985	99.978	99.973	99.972	99.976	99.980
1994	99.980	99.984	99.988	99.985	99.976	99.976	99.979	99.974	99.973	99.973	99.973	99.973
1995	99.973	99.973	99.973	.	.	.	.	.	.	.	.	.

On a par exemple :

$$\begin{aligned} APR86 = & \frac{111.456}{24} \\ & + \frac{108.887 + 100.633 + 103.797 + 99.973 + 107.690 + 101.420 + 97.847 + 103.346 + 95.401 + 67.787 + 101.917}{12} \\ & + \frac{111.434}{24} \\ & = 100.012 \end{aligned}$$

Il reste maintenant à corriger le tableau D10bis par le tableau D10ter pour obtenir l'estimation finale des coefficients saisonniers du tableau D10 :

D10: FACTEURS SAISONNIERS FINAUX

UNE MOYENNE MOBILE 3x5 A ÉTÉ CHOISIE. RAPPORT I/S = 4.949

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	111.443	108.874	100.621
1986	103.785	99.961	107.677	101.408	97.839	103.336	95.387	67.778	101.908	111.426	108.823	100.721
1987	103.846	99.874	107.636	101.410	97.738	103.306	95.445	67.987	101.878	111.490	108.686	100.739
1988	103.935	99.741	107.558	101.448	97.564	103.283	95.524	68.377	101.836	111.516	108.662	100.628
1989	104.057	99.532	107.368	101.524	97.433	103.149	95.606	69.001	101.821	111.663	108.501	100.393
1990	104.122	99.323	107.167	101.639	97.275	103.092	95.760	69.666	101.770	111.700	108.421	100.099
1991	104.202	99.080	106.916	101.735	97.334	102.923	95.976	70.317	101.730	111.719	108.171	99.873
1992	104.217	98.876	106.764	101.827	97.454	102.865	96.199	70.735	101.679	111.554	107.957	99.854
1993	104.208	98.737	106.672	101.862	97.688	102.776	96.297	70.969	101.714	111.402	107.678	100.018
1994	104.161	98.642	106.675	101.887	97.775	102.789	96.312	71.038	101.749	111.309	107.553	100.158
1995	104.111	98.588	106.721	.	.	.	.	.	.	.	.	.

Avec par exemple : $APR86 = 100*(101.420/100.012) = 101.408$

On peut enfin prévoir le facteur saisonnier du mois d'avril 1995 de la façon suivante :

$$APR95 = \frac{3*101.887 - 101.862}{2} = 101.899$$

Ce qui conduit au tableau D10A suivant :

D10A: PRÉVISION DES FACTEURS SAISONNIERS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1995	.	.	.	101.899	97.818	102.795	96.320	71.073	101.766	111.262	107.490	100.229
1996	104.085	98.561	106.743	.	.	.	.	.	.	.	.	.

4.3.11 Tableau D11 : Série corrigée des variations saisonnières et des effets de jours ouvrables

4.3.11.1 Description et mode de calcul

La série de départ ou, si une correction pour jours ouvrables a été demandée, la série du tableau C19 est corrigée des facteurs saisonniers du tableau D10, pour obtenir la série désaisonnalisée finale (tableau D11) : *D11 = C19 _op D10*.

Un test de Fisher est effectué pour vérifier qu'il n'existe pas de saisonnalité résiduelle dans la série D11. Pour cela, la tendance est d'abord enlevée par différenciation, sur 3 mois pour une série mensuelle ou sur un trimestre pour une série trimestrielle. Le test est alors fait sur la série complète mais aussi sur les trois dernières années, soit sur les 36 dernières observations pour une série mensuelle et sur les 12 dernières pour une série trimestrielle. Ce test de Fisher correspond a un test d'analyse de la variance à un facteur tout à fait similaire à celui présenté aux tableaux B3 et D8.
Seuls les statistiques de Fisher et le résultat du test aux niveaux de 1% et 5% sont édités. Un message prévient aussi de la difficulté d'interpréter ce test lorsque la série évolue fortement dans les dernières années d'observation.

4.3.11.2 Exemple

D11: SÉRIE DÉSAISONNALISÉE FINALE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	101.587	102.236	100.085
1986	101.034	99.620	98.909	106.822	100.020	101.281	102.274	99.415	102.091	103.374	102.421	102.275
1987	96.935	104.252	105.001	104.088	104.934	103.677	104.916	103.579	105.552	105.022	106.510	106.845
1988	106.275	107.287	107.985	108.020	110.198	109.453	108.654	109.671	110.936	108.373	110.433	114.168
1989	113.423	113.734	110.122	116.272	111.904	114.971	113.220	112.595	113.697	113.586	115.180	115.133
1990	114.759	114.279	115.521	115.285	116.283	114.852	116.874	117.617	115.890	116.603	115.563	113.463
1991	115.783	114.864	114.439	116.068	114.501	116.692	118.537	116.234	115.479	115.963	116.956	116.069
1992	116.566	117.119	116.267	116.147	115.553	114.120	114.624	115.931	116.072	115.827	114.147	111.515
1993	111.902	115.569	113.496	110.520	113.221	112.715	112.439	112.420	111.238	111.397	111.774	112.903
1994	114.450	114.044	113.741	114.818	116.718	116.020	116.603	118.619	116.844	116.402	117.460	120.004
1995	119.326	119.019	120.007	.	.	.	.	.	.	.	.	.

Avec par exemple : $APR86 = 100*(108.326/101.408) = 106.822$.

La série est alors différenciée sur 3 mois pour enlever la tendance avant d'effectuer le test de saisonnalité résiduelle. Ainsi, on a : $APR86 = 106.822 - 101.034 = 5.788$.

SÉRIE DÉSAISONNALISÉE DIFFÉRENCIÉE POUR TEST DE SAISONNALITÉ RÉSIDUELLE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	-0.553	-2.616	-1.176	5.788	0.400	2.372	-4.548	-0.604	0.810	1.100	3.005	0.184
1987	-6.440	1.832	2.726	7.154	0.682	-1.324	0.827	-1.355	1.875	0.107	2.931	1.293
1988	1.253	0.776	1.140	1.745	2.911	1.467	0.634	-0.527	1.483	-0.281	0.763	3.233
1989	5.050	3.300	-4.046	2.849	-1.830	4.849	-3.052	0.691	-1.274	0.366	2.585	1.436
1990	1.173	-0.901	0.388	0.525	2.004	-0.669	1.589	1.334	1.038	-0.271	-2.053	-2.427
1991	-0.820	-0.699	0.976	0.286	-0.364	2.252	2.468	1.733	-1.212	-2.574	0.722	0.590
1992	0.603	0.163	0.198	-0.419	-1.566	-2.147	-1.523	0.377	1.951	1.202	-1.784	-4.556
1993	-3.925	1.423	1.981	-1.381	-2.349	-0.781	1.919	-0.800	-1.477	-1.042	-0.646	1.665
1994	3.054	2.270	0.837	0.368	2.673	2.279	1.785	1.901	0.824	-0.201	-1.159	3.160
1995	2.924	1.559	0.003	.	.	.	.	.	.	.	.	.

Les résultats des tests de Fisher sont alors édités :

```
TEST SUR LA PRÉSENCE DE SAISONNALITÉ RÉSIDUELLE

Pas  de  saisonnalité  résiduelle  détectée  dans  la  série  complète  au   F=0.52
niveau de 1%
                                                                            F=0.38
Pas de saisonnalité résiduelle détectée dans les 3 dernières années au
niveau de 1%
Pas de saisonnalité résiduelle détectée dans les 3 dernières années au
niveau de 5%

Remarque : Les résultats de ce test sur les 3 dernières années peuvent être
sujet à caution si le niveau de la série désaisonnalisée évolue fortement.
```

4.3.12 Tableau D11A : Série désaisonnalisée finale dont les totaux sont révisés

4.3.12.1 Description et mode de calcul

L'utilisateur peut demander d'ajuster la série désaisonnalisée du tableau D11, de façon à ce que les totaux annuels de la série brute (éventuellement ajustée *a priori*) et de la série ainsi ajustée D11A soient les mêmes (à des fins de publication par exemple). Dans ce cas, les différences annuelles observées entre D11 et A1 sont réparties sur les valeurs de la série désaisonnalisée de telle sorte que les évolutions mensuelles (ou trimestrielles) soient les plus semblables possible sur les séries D11 et D11A.
En principe, cet ajustement n'est pas souhaitable lorsque la saisonnalité évolue sensiblement ou lorsqu'on a demandé une correction pour jours ouvrables.
On trouvera tous les détails théoriques de l'ajustement dans Cholette (1978) ou Cholette et Dagum (1994).
Supposons que nos séries - originale Y, corrigée des variations saisonnières Z et ajustée X - comprennent T points répartis sur I années de J périodes (12 mois ou 4 trimestres). La série inconnue X doit évoluer le plus parallèlement possible à la série connue Z, sous la contrainte que ses sommes annuelles soient égales à celles de la série Y.
Si Δ désigne l'opérateur différence première, le problème peut donc s'écrire :

$$\begin{cases} \underset{x}{Min} \ \sum_{t=2}^{T} \left[\Delta(x_t - z_t) \right]^2 \\ \sum_{j=1}^{J} x_{ij} = y_{i.} \quad \forall i = 1, \ldots., I \end{cases}$$

Ou encore, en définissant les matrices suivantes :

$$\underset{(T-1,T)}{D} = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix}, \quad \underset{(J,1)}{e} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix} \quad \text{et} \quad \underset{(T,I)}{B} = \begin{bmatrix} e & 0 & \cdots & 0 & 0 \\ 0 & e & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & e & 0 \\ 0 & 0 & \cdots & 0 & e \end{bmatrix}$$

$$\begin{cases} \underset{X}{Min} \ (X - Z)'D'D(X - Z) = (X - Z)'A(X - Z) \\ Y - B'X = 0 \end{cases}$$

En faisant intervenir un vecteur Λ de I multiplicateurs de Lagrange, la solution s'écrit :

$$\begin{bmatrix} \hat{X} \\ \hat{\Lambda} \end{bmatrix} = \begin{bmatrix} A & B \\ B' & 0 \end{bmatrix}^{-1} \begin{bmatrix} A & 0 \\ B' & I_I \end{bmatrix} \begin{bmatrix} Z \\ Y - B'Z \end{bmatrix} = W \begin{bmatrix} Z \\ Y - B'Z \end{bmatrix} = W \begin{bmatrix} Z \\ R \end{bmatrix}$$

où I_k désigne la matrice unité d'ordre k.

Comme on peut le voir, la matrice *W*, de dimensions (*T*+*I*,*T*+*I*), ne dépend pas des données, *Y* et *Z*, et on peut montrer qu'elle a toujours la forme $W = \begin{bmatrix} I_T & W_X \\ 0 & W_\Lambda \end{bmatrix}$ où W_X est une matrice de dimensions (*T*,*I*).
On a donc finalement :

$$\forall t = 1,\ldots,T \quad \hat{x}_t = z_t + \sum_{i=1}^{I} w_{X(t,i)} * r_i$$

En d'autres termes, le coefficient correcteur apporté à la série z_t (notre série D11) est une moyenne pondérée des différences entre les totaux annuels des séries y_t (notre série A1) et z_t. Cette moyenne porte sur les *I* années d'observations.

4.3.12.2 Remarques

- Le calcul ci-dessus est rendu un peu complexe par le fait que la matrice $A = D'D$ n'est pas inversible. On trouve une solution élégante, basée sur une régression par moindres carrés généralisés dans Cholette et Dagum (1994) et Bournay et Laroque (1979).
- Les coefficients correcteurs dépendant de toutes les différences de totaux annuels observées, cela signifie que toute nouvelle année complète de données entraîne une révision de toute la série D11A. Cela n'est pas souhaitable et c'est pourquoi les programmes X11-ARIMA88 et X12-ARIMA incorporent les valeurs de la matrice de poids W_x calculée dans le cas de 5 années d'observations : matrice de dimensions (60,5) dans le cas mensuel et (20,5) dans le cas trimestriel. C'est cette matrice qui est utilisée, quelle que soit la longueur de la série pour calculer la série ajustée D11A. Dans ce cas, voir Tableau 5, la première année complète de données est associée aux 12 (pour le cas mensuel) premières lignes de la matrice W_x, la seconde année est associée avec les 12 lignes suivantes. Les deux dernières années complètes sont associées de même aux 24 dernières lignes de la matrice. Toutes les autres années, de la troisième à l'antépénultième, sont associées aux 12 lignes centrales de la matrice W_x.
- Si la dernière année n'est pas complète, les estimations du tableau D11A pour les mois renseignés se font en appliquant le dernier facteur correctif calculé (c'est à dire celui du mois de décembre de la dernière année complète), ce qui correspond d'ailleurs à la solution du problème de minimisation exposé ci-dessus dans le cas d'années incomplètes.
- X11-ARIMA88 et X12-ARIMA n'envisagent que le cas d'un ajustement additif pas nécessairement compatible avec un schéma de composition multiplicatif. Ainsi, pour ce type de schéma, il peut arriver que le tableau D11A présente des valeurs négatives. Cholette et Dagum (1994) donnent la solution proportionnelle à appliquer dans le cas multiplicatif, solution non disponible dans les versions actuelles des logiciels.

4.3.12.3 Exemple

Nous devons réconcilier les séries D11 et B1. Les valeurs des totaux annuels de ces deux séries et de leurs différences (résidus R) sont :

ANNÉE	B1	D11	R
1986	1220.5	1219.53575	0.96425
1987	1252.8	1251.31218	1.48782
1988	1312.5	1311.45328	1.04672
1989	1361.2	1363.83663	-2.63663
1990	1385.3	1386.98833	-1.68833
1991	1391.1	1391.58606	-0.48606
1992	1389.2	1383.88885	5.31115
1993	1348.8	1349.59492	-0.79492
1994	1391.8	1395.72244	-3.92244

Comme on peut le voir, les différences entre les totaux annuels sont assez faibles. Cela tient au fait que nous avons, régulièrement dans le processus de désaisonnalisation, *normalisé*[32] les estimations des coefficients saisonniers.
Pour les données de la première année complète (1986), les matrices de poids et de résidus utilisées sont :

Poids	1	2	3	4	5	Résidu	Corrections
Jan	0.10539496	-.02790048	0.00737759	-.00191944	0.00038069		0.0723
Feb	0.10446930	-.02672983	0.00706804	-.00183890	0.00036472	0.96425	0.0726
Mar	0.10261797	-.02438853	0.00644894	-.00167783	0.00033277	1.48782	0.0733
Apr	0.09984099	-.02087658	0.00552030	-.00143622	0.00028485	1.04672	0.0743
May	0.09613833	-.01619398	0.00428210	-.00111408	0.00022096	-2.63663	0.0757
June	0.09151002	-.01034074	0.00273435	-.00071140	0.00014110	-1.68833	0.0774
July	0.08595605	-.00331684	0.00087706	-.00022818	0.00004526		0.0794
Aug	0.07947641	0.00487771	-.00128979	0.00033557	-.00006655		0.0818
Sep	0.07207110	0.01424290	-.00376618	0.00097985	-.00019434		0.0845
Oct	0.06374014	0.02477874	-.00655213	0.00170467	-.00033810		0.0875
Nov	0.05448351	0.03648524	-.00964762	0.00251003	-.00049783		0.0909
Dec	0.04430122	0.04936238	-.01305266	0.00339593	-.00067353		0.0947

Les corrections apportées se calculent par le produit de ces deux matrices et on a, par exemple, pour Février 1986 :

$$FEB86 = (0.10446930 * 0.96425) + (-.02672983 * 1.48782) + (0.00706804 * 1.04672) + (-.00183890 * -2.63663) + (0.00036472 * -1.68833) = 0.0726$$

Ce qui conduit à la valeur ajusté :

$$D11A(FEB86) = D11(FEB86) + 0.0726 = 99.620 + 0.0726 = 99.693$$

[32] C'est à dire fait en sorte que leur somme soit à peu près égale à 0, ou à 12 selon le schéma, sur toute période d'un an.

Pour la quatrième année de données complètes (1989), les matrices de poids et de résidus utilisées sont :

Poids	1	2	3	4	5	Résidu	Corrections
Jan	-.00879152	0.04432649	0.06039494	-.01571303	0.00311645		-0.1009
Feb	-.00616123	0.03106468	0.07290679	-.01805856	0.00358166		-0.1401
Mar	-.00384049	0.01936360	0.08291627	-.01884336	0.00373731	1.48782	-0.1741
Apr	-.00182930	0.00922326	0.09042338	-.01806743	0.00358341	1.04672	-0.2027
May	-.00012766	0.00064366	0.09542812	-.01573076	0.00311997	-2.63663	-0.2261
June	0.00126443	-.00637522	0.09793049	-.01183335	0.00234698	-1.68833	-0.2442
July	0.00234698	-.01183335	0.09793049	-.00637522	0.00126443	-0.48606	-0.2570
Aug	0.00311997	-.01573076	0.09542812	0.00064366	-.00012766		-0.2645
Sep	0.00358341	-.01806743	0.09042338	0.00922326	-.00182930		-0.2667
Oct	0.00373731	-.01884336	0.08291627	0.01936361	-.00384049		-0.2636
Nov	0.00358166	-.01805856	0.07290679	0.03106468	-.00616123		-0.2553
Dec	0.00311645	-.01571303	0.06039494	0.04432649	-.00879152		-0.2416

On a, par exemple, pour Août 1989 :

$$AUG89 = (0.00311997 * 1.48782) + (-0.01573076 * 1.04672) + (0.09542812 * -2.63663)$$
$$+ (0.00064366 * -1.68833) + (-0.00012766 * -0.48606) = -0.2645$$

Ce qui conduit à la valeur ajustée suivante :

$$D11A(AUG89) = D11(AUG89) - 0.2645 = 112.595 - 0.2645 = 112.331$$

Pour la dernière année de données complètes (1994), les matrices de poids et de résidus utilisées sont :

Poids	1	2	3	4	5	Résidu	Corrections
Jan	-.00067353	0.00339593	-.01305266	0.04936238	0.04430122		-0.2828
Feb	-.00049783	0.00251003	-.00964762	0.03648524	0.05448351		-0.2943
Mar	-.00033810	0.00170467	-.00655213	0.02477874	0.06374014	-1.68833	-0.3048
Apr	-.00019434	0.00097985	-.00376618	0.01424290	0.07207110	-0.48606	-0.3142
May	-.00006655	0.00033557	-.00128979	0.00487771	0.07947641	5.31115	-0.3225
June	0.00004526	-.00022818	0.00087706	-.00331684	0.08595605	-0.79492	-0.3298
July	0.00014110	-.00071140	0.00273435	-.01034074	0.09151002	-3.92244	-0.3361
Aug	0.00022096	-.00111408	0.00428210	-.01619398	0.09613833		-0.3413
Sep	0.00028485	-.00143622	0.00552030	-.02087658	0.09984099		-0.3455
Oct	0.00033277	-.00167783	0.00644894	-.02438853	0.10261797		-0.3486
Nov	0.00036472	-.00183890	0.00706804	-.02672983	0.10446930		-0.3507
Dec	0.00038069	-.00191944	0.00737759	-.02790048	0.10539496		-0.3518

On a, par exemple, pour décembre 1994 :

$$DEC94 = (0.00038069 * -1.68833) + (-0.00191944 * -0.48606) + (0.00737759 * 5.31115)$$
$$+ (-0.02790048 * -0.79492) + (0.10539496 * -3.92244) = -0.3518$$

Ce qui conduit à la valeur ajustée suivante :

$$D11A(DEC94) = D11(DEC94) - 0.3518 = 120.004 - 0.3518 = 119.652$$

Ce dernier facteur correctif va nous servir à estimer les valeurs de la dernière année incomplète 1995. Ainsi, on aura :

$$D11A(JAN95) = D11(JAN95) - 0.3518 = 119.326 - 0.3518 = 118.974$$

Ce qui conduit au tableau final D11A suivant :

D11A: SÉRIE DÉSAISONNALISÉE FINALE AJUSTÉE SUR LES TOTAUX ANNUELS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	101.106	99.693	98.982	106.896	100.095	101.358	102.353	99.497	102.175	103.462	102.512	102.369
1987	97.033	104.355	105.108	104.200	105.050	103.798	105.041	103.710	105.688	105.163	106.656	106.996
1988	106.432	107.445	108.141	108.168	110.335	109.574	108.756	109.749	110.986	108.391	110.415	114.109
1989	113.322	113.594	109.948	116.070	111.678	114.727	112.963	112.331	113.430	113.323	114.925	114.891
1990	114.556	114.095	115.355	115.133	116.144	114.723	116.752	117.499	115.774	116.487	115.444	113.337
1991	115.616	114.694	114.273	115.915	114.368	116.587	118.468	116.210	115.507	116.051	117.112	116.301
1992	116.880	117.502	116.705	116.626	116.060	114.642	115.146	116.440	116.554	116.270	114.536	111.838
1993	112.143	115.736	113.594	110.554	113.197	112.639	112.316	112.257	111.040	111.168	111.522	112.633
1994	114.167	113.750	113.436	114.504	116.395	115.690	116.267	118.278	116.498	116.053	117.109	119.652
1995	118.974	118.667	119.655	.	.	.	.	.	.	.	.	.

Tableau 5 : Matrice des poids utilisée par X11-ARIMA88 et X12-ARIMA dans le cas mensuel.

	1	2	3	4	5
Jan	0.10539496	-0.02790048	0.00737759	-0.00191944	0.00038069
Feb	0.10446930	-0.02672983	0.00706804	-0.00183890	0.00036472
Mar	0.10261797	-0.02438853	0.00644894	-0.00167783	0.00033277
Apr	0.09984099	-0.02087658	0.00552030	-0.00143622	0.00028485
May	0.09613833	-0.01619398	0.00428210	-0.00111408	0.00022096
June	0.09151002	-0.01034074	0.00273435	-0.00071140	0.00014110
July	0.08595605	-0.00331684	0.00087706	-0.00022818	0.00004526
Aug	0.07947641	0.00487771	-0.00128979	0.00033557	-0.00006655
Sep	0.07207110	0.01424290	-0.00376618	0.00097985	-0.00019434
Oct	0.06374014	0.02477874	-0.00655213	0.00170467	-0.00033810
Nov	0.05448351	0.03648524	-0.00964762	0.00251003	-0.00049783
Dec	0.04430122	0.04936238	-0.01305266	0.00339593	-0.00067353
Jan	0.03319327	0.06341017	-0.01676725	0.00436236	-0.00086521
Feb	0.02325596	0.07505210	-0.01892111	0.00492273	-0.00097635
Mar	0.01448931	0.08428817	-0.01951423	0.00507704	-0.00100696
Apr	0.00689330	0.09111838	-0.01854662	0.00482530	-0.00095703
May	0.00046795	0.09554273	-0.01601827	0.00416749	-0.00082656
June	-0.00478676	0.09756121	-0.01192919	0.00310363	-0.00061556
July	-0.00887082	0.09717384	-0.00627938	0.00163371	-0.00032402
Aug	-0.01178422	0.09438060	0.00093117	-0.00024226	0.00004805
Sep	-0.01352698	0.08918150	0.00970246	-0.00252430	0.00050066
Oct	-0.01409909	0.08157654	0.02003448	-0.00521239	0.00103380
Nov	-0.01350055	0.07156572	0.03192723	-0.00830655	0.00164748
Dec	-0.01173136	0.05914904	0.04538072	-0.01180676	0.00234170
Jan	-0.00879152	0.04432649	0.06039494	-0.01571303	0.00311645
Feb	-0.00616123	0.03106468	0.07290679	-0.01805856	0.00358166
Mar	-0.00384049	0.01936360	0.08291627	-0.01884336	0.00373731
Apr	-0.00182930	0.00922326	0.09042338	-0.01806743	0.00358341
May	-0.00012766	0.00064366	0.09542812	-0.01573076	0.00311997
June	0.00126443	-0.00637522	0.09793049	-0.01183335	0.00234698
July	0.00234698	-0.01183335	0.09793049	-0.00637522	0.00126443
Aug	0.00311997	-0.01573076	0.09542812	0.00064366	-0.00012766
Sep	0.00358341	-0.01806743	0.09042338	0.00922326	-0.00182930
Oct	0.00373731	-0.01884336	0.08291627	0.01936361	-0.00384049
Nov	0.00358166	-0.01805856	0.07290679	0.03106468	-0.00616123
Dec	0.00311645	-0.01571303	0.06039494	0.04432649	-0.00879152
Jan	0.00234170	-0.01180676	0.04538072	0.05914904	-0.01173136
Feb	0.00164748	-0.00830655	0.03192723	0.07156572	-0.01350055
Mar	0.00103380	-0.00521239	0.02003448	0.08157654	-0.01409909
Apr	0.00050066	-0.00252430	0.00970246	0.08918150	-0.01352698
May	0.00004805	-0.00024226	0.00093117	0.09438060	-0.01178422
June	-0.00032402	0.00163371	-0.00627938	0.09717384	-0.00887082
July	-0.00061556	0.00310363	-0.01192919	0.09756121	-0.00478676
Aug	-0.00082656	0.00416749	-0.01601827	0.09554273	0.00046795
Sep	-0.00095703	0.00482530	-0.01854662	0.09111838	0.00689330
Oct	-0.00100696	0.00507704	-0.01951423	0.08428817	0.01448931
Nov	-0.00097635	0.00492273	-0.01892111	0.07505210	0.02325596
Dec	-0.00086521	0.00436236	-0.01676725	0.06341017	0.03319327
Jan	-0.00067353	0.00339593	-0.01305266	0.04936238	0.04430122
Feb	-0.00049783	0.00251003	-0.00964762	0.03648524	0.05448351
Mar	-0.00033810	0.00170467	-0.00655213	0.02477874	0.06374014
Apr	-0.00019434	0.00097985	-0.00376618	0.01424290	0.07207110
May	-0.00006655	0.00033557	-0.00128979	0.00487771	0.07947641
June	0.00004526	-0.00022818	0.00087706	-0.00331684	0.08595605
July	0.00014110	-0.00071140	0.00273435	-0.01034074	0.09151002
Aug	0.00022096	-0.00111408	0.00428210	-0.01619398	0.09613833
Sep	0.00028485	-0.00143622	0.00552030	-0.02087658	0.09984099
Oct	0.00033277	-0.00167783	0.00644894	-0.02438853	0.10261797
Nov	0.00036472	-0.00183890	0.00706804	-0.02672983	0.10446930
Dec	0.00038069	-0.00191944	0.00737759	-0.02790048	0.10539496

4.3.13 Tableau D12 : Estimation finale de la Tendance-Cycle

4.3.13.1 Description et mode de calcul

La série D1, où ont été enlevés les points atypiques et les effets de jours ouvrables, est corrigée par les facteurs saisonniers du tableau D10 pour obtenir une série désaisonnalisée modifiée. C'est celle-ci qui est lissée, avec une moyenne mobile de Henderson, à cette étape pour obtenir une estimation finale de la Tendance-Cycle. La méthodologie est la même que celle suivie pour construire le tableau C7 (voir 4.2.6). On a donc *CVS = D1 _op D10*.

- ***Étape 1 : choix de la moyenne mobile, calcul du ratio*** $\bar{I}/\bar{C}$
- ***Étape 2 : lissage de la série CVS par une moyenne mobile de Henderson***

4.3.13.2 Exemple

On calcule d'abord une série CVS modifiée.

D11bis: SÉRIE DÉSAISONNALISÉE MODIFIÉE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	101.587	102.236	100.085
1986	101.034	99.620	98.909	100.294	100.020	101.281	102.274	102.244	102.091	103.374	102.421	102.275
1987	102.484	104.252	105.001	104.088	104.934	103.677	104.916	103.579	105.552	105.022	106.510	106.845
1988	106.275	107.287	107.985	108.020	110.198	109.453	108.654	109.671	110.936	111.120	110.433	114.168
1989	113.423	113.734	113.351	113.398	111.904	114.871	113.220	112.595	113.697	113.586	115.180	115.133
1990	114.759	114.279	115.521	115.285	116.283	114.852	116.874	117.617	115.890	116.603	115.563	113.463
1991	115.783	114.864	114.439	116.068	114.501	116.692	116.189	116.234	115.479	115.963	116.956	116.069
1992	116.566	117.119	116.267	116.147	115.553	114.120	114.624	115.931	116.072	115.827	114.147	113.279
1993	111.902	112.730	113.496	112.410	113.221	112.715	112.439	112.420	111.238	111.397	111.774	112.903
1994	114.450	114.044	113.741	114.818	116.081	116.020	116.603	116.220	116.844	116.402	117.460	120.004
1995	119.326	119.019	120.007	.	.	.	.	.	.	.	.	.

Avec, par exemple : $APR86 = 100*(101.706/101.408) = 100.294$

- ***Étape 1 : choix de la moyenne mobile, calcul du ratio*** $\bar{I}/\bar{C}$

On lisse tout d'abord le tableau D11bis par une moyenne mobile de Henderson sur 13 termes dont les coefficients figurent dans le tableau 3 (voir 3.3.1).

Le premier terme calculable est donc celui de avril 1986, et on a, aux arrondis près :

APR86= 101.587(-0.01935)+102.236*(-0.02786)+100.085*(0.00000)*
+101.034(0.06549)+99.620*(0.14736)+98.909*(0.21434)*
+100.294(0.24006)+100.020*(0.21434)+101.281*(0.14736)*
+102.274(0.06549)+102.244*(0.00000)+102.091*(-0.02786)*
+103.374(-0.01935) =99.974*

A cette étape du calcul, on ne se préoccupe pas d'estimer les 6 points non calculables en début et en fin de série. On en déduit une estimation de la tendance (tableau D12a) et, par rapport avec le tableau D11bis, de la composante irrégulière (tableau D12b):

D12a : TENDANCE-CYCLE (Moyenne Mobile de Henderson sur 13 termes)

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	99.974	100.452	101.097	101.732	102.206	102.428	102.530	102.646	102.889
1987	103.273	103.736	104.129	104.379	104.447	104.388	104.399	104.597	104.981	105.466	105.942	106.409
1988	106.900	107.438	107.964	108.469	108.927	109.284	109.565	109.900	110.422	111.138	111.983	112.724
1989	113.206	113.457	113.517	113.439	113.287	113.193	113.251	113.435	113.720	114.050	114.367	114.672
1990	114.915	115.024	115.159	115.400	115.752	116.179	116.475	116.536	116.363	115.983	115.517	115.089
1991	114.825	114.818	115.012	115.309	115.604	115.821	115.971	116.074	116.128	116.209	116.346	116.516
1992	116.647	116.562	116.246	115.807	115.447	115.279	115.308	115.399	115.318	114.956	114.338	113.620
1993	113.033	112.734	112.717	112.815	112.839	112.665	112.313	111.950	111.784	111.883	112.219	112.753
1994	113.367	113.993	114.565	115.078	115.541	115.902	116.190	116.476	116.818	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

D12b : IRRÉGULIER

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	100.320	99.569	100.182	100.533	100.037	99.670	100.824	99.781	99.403
1987	99.235	100.497	100.837	99.721	100.466	99.319	100.495	99.027	100.544	99.579	100.536	100.410
1988	99.416	99.859	100.020	99.587	101.167	100.154	99.169	99.792	100.465	99.984	98.616	101.281
1989	100.192	100.244	99.854	99.964	98.779	101.483	99.973	99.260	99.980	99.593	100.711	100.402
1990	99.865	99.352	100.315	99.900	100.459	98.858	100.342	100.927	99.593	100.535	100.040	98.587
1991	100.834	100.041	99.502	100.659	99.046	100.752	100.188	100.138	99.441	99.789	100.525	99.617
1992	99.931	100.478	100.018	100.294	100.092	98.995	99.407	100.461	100.653	100.758	99.833	99.700
1993	98.999	99.996	100.691	99.641	100.338	100.045	100.112	100.420	99.512	99.565	99.603	100.134
1994	100.955	100.045	99.280	99.774	100.467	100.102	100.356	99.781	100.022	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

Le schéma étant multiplicatif, on calcule les taux de croissance :

D12c : CROISSANCE, EN VALEUR ABSOLUE, DE LA TENDANCE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC	TOTAL
1985	.	.	.	.	.	.	.	.	.	.	.	.	0.000
1986	.	.	.	.	0.479	0.642	0.628	0.466	0.217	0.099	0.113	0.236	2.881
1987	0.374	0.448	0.379	0.240	0.065	0.057	0.010	0.190	0.367	0.462	0.452	0.440	3.484
1988	0.461	0.504	0.489	0.467	0.423	0.327	0.257	0.306	0.475	0.648	0.760	0.662	5.780
1989	0.428	0.221	0.053	0.069	0.133	0.084	0.051	0.163	0.251	0.290	0.278	0.267	2.288
1990	0.211	0.095	0.117	0.210	0.304	0.369	0.255	0.052	0.148	0.327	0.401	0.371	2.862
1991	0.229	0.007	0.169	0.259	0.255	0.188	0.130	0.089	0.046	0.070	0.118	0.146	1.705
1992	0.112	0.073	0.271	0.378	0.311	0.146	0.025	0.079	0.070	0.315	0.538	0.627	2.944
1993	0.517	0.264	0.015	0.087	0.021	0.154	0.312	0.323	0.148	0.088	0.301	0.475	2.706
1994	0.545	0.552	0.502	0.447	0.403	0.312	0.249	0.246	0.294	.	.	.	3.550
1995	.	.	.	.	.	.	.	.	.	.	.	.	0.000

$$\overline{C} = \frac{2.881 + 3.484 + 5.780 + 2.288 + 2.862 + 1.705 + 2.944 + 2.706 + 3.550}{101} = 0.2792$$

D12d : CROISSANCE, EN VALEUR ABSOLUE, DE L'IRRÉGULIER

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC	TOTAL
1985	.	.	.	.	.	.	.	.	.	.	.	.	0.000
1986	.	.	.	.	0.749	0.615	0.350	0.493	0.366	1.157	1.035	0.378	5.144
1987	0.169	1.272	0.338	1.107	0.747	1.142	1.184	1.461	1.532	0.959	0.961	0.126	10.998
1988	0.990	0.446	0.161	0.433	1.587	1.001	0.984	0.628	0.675	0.478	1.368	2.702	11.453
1989	1.076	0.053	0.389	0.110	1.186	2.738	1.488	0.713	0.725	0.386	1.122	0.307	10.293
1990	0.535	0.514	0.969	0.414	0.560	1.594	1.502	0.583	1.322	0.945	0.492	1.452	10.882
1991	2.279	0.786	0.538	1.162	1.602	1.722	0.560	0.050	0.695	0.349	0.738	0.903	11.385
1992	0.315	0.547	0.458	0.276	0.201	1.096	0.416	1.060	0.192	0.104	0.918	0.134	5.716
1993	0.703	1.007	0.695	1.043	0.700	0.293	0.068	0.308	0.905	0.054	0.038	0.532	6.345
1994	0.821	0.902	0.764	0.498	0.694	0.364	0.254	0.573	0.241	.	.	.	5.110
1995	.	.	.	.	.	.	.	.	.	.	.	.	0.000

$$\bar{I} = \frac{5.144 + 10.998 + 11.453 + 10.293 + 10.882 + 11.385 + 5.716 + 6.345 + 5.110}{101} = 0.7656$$

donc $\dfrac{\bar{I}}{\overline{C}} = \dfrac{0.7656}{0.2792} = 2.742$

- ***Étape 2 : lissage de la série CVS par une moyenne mobile de Henderson***

Le ratio étant supérieur à 1 et inférieur à 3.5, on choisit une moyenne mobile de Henderson sur 13 termes dont les coefficients et ceux des moyennes mobiles asymétriques associées figurent dans le tableau 3 (voir 3.3.1).. L'estimation de la tendance pour octobre 1985 se fait en utilisant le point courant et six points dans le futur auxquels on applique les coefficients de la moyenne mobile H6_0 du tableau 3.

$$OCT85 = 101.587*(0.42113)+102.236*(0.35315)+100.085*(0.24390)+101.034*(0.11977) +99.620*(0.01202)+98.909*(-0.05811)+100.294*(-0.09186) \quad =101.634$$

Ce qui conduit au tableau D12 :

D12: TENDANCE-CYCLE
valeur de I/C: 2.7420
UNE MOYENNE MOBILE DE HENDERSON SUR 13 TERMES A ÉTÉ CHOISIE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	101.634	101.254	100.809
1986	100.356	99.967	99.809	99.974	100.452	101.097	101.732	102.206	102.428	102.530	102.646	102.889
1987	103.273	103.736	104.129	104.379	104.447	104.388	104.399	104.597	104.981	105.466	105.942	106.409
1988	106.900	107.438	107.964	108.469	108.927	109.284	109.565	109.900	110.422	111.138	111.983	112.724
1989	113.206	113.457	113.517	113.439	113.287	113.193	113.251	113.435	113.720	114.050	114.367	114.672
1990	114.915	115.024	115.159	115.400	115.752	116.179	116.475	116.536	116.363	115.983	115.517	115.089
1991	114.825	114.818	115.012	115.309	115.604	115.821	115.971	116.074	116.128	116.209	116.346	116.516
1992	116.647	116.562	116.246	115.807	115.447	115.279	115.308	115.399	115.318	114.956	114.338	113.620
1993	113.033	112.734	112.717	112.815	112.839	112.665	112.313	111.950	111.784	111.883	112.219	112.753
1994	113.367	113.993	114.565	115.078	115.541	115.902	116.190	116.476	116.818	117.300	117.921	118.567
1995	119.144	119.619	119.961	.	.	.	.	.	.	.	.	.

4.3.14 Tableau D13 : Estimation finale de la composante irrégulière

4.3.14.1 Description et mode de calcul

Cette estimation finale de la composante irrégulière s'obtient en enlevant la composante tendancielle (tableau D12) de l'estimation de la série désaisonnalisée obtenue en D11 : *D13 = D11 _op D12*.

4.3.14.2 Exemple

D13: COMPOSANTE IRRÉGULIÈRE FINALE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	99.954	100.970	99.281
1986	100.676	99.653	99.099	106.850	99.569	100.182	100.533	97.269	99.670	100.824	99.781	99.403
1987	93.862	100.497	100.837	99.721	100.466	99.319	100.495	99.027	100.544	99.579	100.536	100.410
1988	99.416	99.859	100.020	99.587	101.167	100.154	99.169	99.792	100.465	97.512	98.616	101.281
1989	100.192	100.244	97.009	102.498	98.779	101.571	99.973	99.260	99.980	99.593	100.711	100.402
1990	99.865	99.352	100.315	99.900	100.459	98.858	100.342	100.927	99.593	100.535	100.040	98.587
1991	100.834	100.041	99.502	100.659	99.046	100.752	102.212	100.138	99.441	99.789	100.525	99.617
1992	99.931	100.478	100.018	100.294	100.092	98.995	99.407	100.461	100.653	100.758	99.833	98.148
1993	98.999	102.515	100.691	97.966	100.338	100.045	100.112	100.420	99.512	99.565	99.603	100.134
1994	100.955	100.045	99.280	99.774	101.018	100.102	100.356	101.840	100.022	99.235	99.609	101.212
1995	100.153	99.499	100.038	.	.	.	.	.	.	.	.	.

Avec, par exemple : $APR86 = 100*(106.822/99.974) = 106.850$

4.3.15 Tableau D16 : Estimation des différents effets saisonnier et de calendrier

4.3.15.1 Description et mode de calcul

Cette estimation finale de la composante saisonnière et des effets de calendrier s'obtient en enlevant la série désaisonnalisée (tableau D11) des données brutes (tableau A1 ou A3 si des ajustements permanents ont été demandés) : *D16 = A1 _op D11*.

4.3.15.2 Exemple

D16: COMPOSANTE SAISONNIÈRE ET EFFETS DE CALENDRIER

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	113.892	107.399	100.515
1986	105.509	99.077	105.046	102.507	97.681	102.388	97.483	66.086	103.046	113.278	105.740	102.078
1987	103.678	98.991	107.523	102.894	95.298	104.459	97.030	66.326	102.982	111.310	107.689	102.953
1988	101.341	102.715	109.922	100.074	97.461	104.794	93.140	69.298	103.303	108.791	109.840	100.466
1989	103.947	98.651	109.152	98.648	98.746	104.635	93.270	70.518	100.443	111.545	110.089	97.887
1990	105.525	98.444	106.993	100.707	99.413	101.696	95.659	70.823	98.887	113.205	109.983	97.653
1991	106.493	98.203	104.248	102.870	98.951	100.007	97.269	70.203	100.797	114.174	106.706	99.768
1992	105.949	99.813	106.651	103.317	95.021	104.013	97.798	69.007	102.781	111.373	106.968	102.049
1993	101.607	97.863	108.109	103.330	95.301	103.890	96.141	70.895	103.202	108.621	108.880	101.680
1994	101.616	97.769	109.020	100.507	97.672	104.293	93.908	71.995	103.215	108.589	108.718	99.997
1995	104.001	97.715	108.494	.	.	.	.	.	.	.	.	.

On a alors par exemple : $APR86 = 100*(109.500/106.822) = 102.507$

4.3.16 Tableau D18 : Effets de calendrier combinés.

4.3.16.1 Description et mode de calcul

Le tableau D18 ne figure que dans les sorties de X12-ARIMA. Ce logiciel propose un traitement des effets de calendrier beaucoup plus sophistiqué que celui de X11-ARIMA. Ces effets peuvent être estimés soit à partir d'une estimation de la composante irrégulière (instruction « X11regression »), soit à partir de la série brute avant toute procédure de désaisonnalisation (instruction « Regression »).
Le tableau D18 présente les effets de calendrier combinés, effets de jours ouvrables et effets liés aux jours fériés, estimés au préalable dans la partie Reg-ARIMA de X12-ARIMA ou par une régression sur la composante irrégulière.

4.3.16.2 Exemple

Dans notre exemple, ce tableau serait égal au tableau C18.

D18: COEFFICIENTS COMBINÉS DE CORRECTION POUR EFFETS DE CALENDRIER

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	102.198	98.646	99.895
1986	101.662	99.115	97.557	101.084	99.839	99.083	102.198	97.504	101.116	101.662	97.167	101.347
1987	99.839	99.115	99.895	101.463	97.504	101.116	101.662	97.557	101.084	99.839	99.083	102.198
1988	97.504	102.982	102.198	98.646	99.895	101.463	97.504	101.347	101.441	97.557	101.084	99.839
1989	99.895	99.115	101.662	97.167	101.347	101.441	97.557	102.198	98.646	99.895	101.463	97.504
1990	101.347	99.115	99.839	99.083	102.198	98.646	99.895	101.662	97.167	101.347	101.441	97.557
1991	102.198	99.115	97.504	101.116	101.662	97.167	101.347	99.839	99.083	102.198	98.646	99.895
1992	101.662	100.947	99.895	101.463	97.504	101.116	101.662	97.557	101.084	99.839	99.083	102.198
1993	97.504	99.115	101.347	101.441	97.557	101.084	99.839	99.895	101.463	97.504	101.116	101.662
1994	97.557	99.115	102.198	98.646	99.895	101.463	97.504	101.347	101.441	97.557	101.084	99.839
1995	99.895	99.115	101.662	.	.	.	.	.	.	.	.	.

4.4 PARTIE E : Composantes corrigées des points les plus atypiques

4.4.1 Tableau E1 : Série brute corrigée des points les plus atypiques

4.4.1.1 Description et mode de calcul

Ce tableau présente la série brute corrigée des points atypiques ayant reçu un poids nul lors de l'étape C17. Lorsque, à une date donnée, un point a été jugé très atypique, la valeur de la série brute est remplacée par la composée de la valeur de la composante tendancielle (tableau D12), de la saisonnalité (tableau D10) et le cas échéant des ajustements *a priori* et des corrections des effets de calendrier.
On a donc : *E1 = D12 _invop D10 _invop C16*.

4.4.1.2 Exemple

E1 : SÉRIE BRUTE CORRIGÉE DES VALEURS ATYPIQUES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	115.700	109.800	100.600
1986	106.600	98.700	103.900	102.480	97.700	103.700	99.700	67.544	105.200	117.100	108.300	104.400
1987	107.072	103.200	112.900	107.100	100.000	108.300	101.800	68.700	108.700	116.900	114.700	110.000
1988	107.700	110.200	118.700	108.100	107.400	114.700	101.200	76.000	114.600	120.908	121.300	114.700
1989	117.900	112.200	123.906	111.905	110.500	120.300	105.600	79.400	114.200	126.700	126.800	112.700
1990	121.100	112.500	123.600	116.100	115.600	116.800	111.800	83.300	114.600	132.000	127.100	110.800
1991	123.300	112.800	119.300	119.400	113.300	116.700	115.300	81.600	116.400	132.400	124.800	115.800
1992	123.500	116.900	124.000	120.000	109.800	118.700	112.100	80.000	119.300	129.000	122.100	113.800
1993	113.700	110.326	122.700	114.200	107.900	117.100	108.100	79.700	114.800	121.000	121.700	114.800
1994	116.300	111.500	124.000	115.400	114.000	121.000	109.500	83.857	120.600	126.400	127.700	120.000
1995	124.100	116.300	130.200	.	.	.	.	.	.	.	.	.

On a ainsi : $APR86 = 99.974 * 1.014 * 1.0108 = 102.480$

4.4.2 Tableau E2 : Série désaisonnalisée corrigée des points les plus atypiques

4.4.2.1 Description et mode de calcul

Ce tableau présente la série désaisonnalisée du tableau D11 corrigée des points atypiques ayant reçu un poids nul lors de l'étape C17. Lorsque, à une date donnée, un point a été jugé très atypique, la valeur de la série désaisonnalisée est remplacée par la valeur de la composante tendancielle (tableau D12).

4.4.2.2 Exemple

E2 : SÉRIE DÉSAISONNALISÉE CORRIGÉE DES VALEURS ATYPIQUES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	101.587	102.236	100.085
1986	101.034	99.620	98.909	99.974	100.020	101.281	102.274	102.206	102.091	103.374	102.421	102.275
1987	103.273	104.252	105.001	104.088	104.934	103.677	104.916	103.579	105.552	105.022	106.510	106.845
1988	106.275	107.287	107.985	108.020	110.198	109.453	108.654	109.671	110.936	111.138	110.433	114.168
1989	113.423	113.734	113.517	113.439	111.904	114.971	113.220	112.595	113.697	113.586	115.180	115.133
1990	114.759	114.279	115.521	115.285	116.283	114.852	116.874	117.617	115.890	116.603	115.563	113.463
1991	115.783	114.864	114.439	116.068	114.501	116.692	118.537	116.234	115.479	115.963	116.956	116.069
1992	116.566	117.119	116.267	116.147	115.553	114.120	114.624	115.931	116.072	115.827	114.147	111.515
1993	111.902	112.734	113.496	110.520	113.221	112.715	112.439	112.420	111.238	111.397	111.774	112.903
1994	114.450	114.044	113.741	114.818	116.718	116.020	116.603	116.476	116.844	116.402	117.460	120.004
1995	119.326	119.019	120.007	.	.	.	.	.	.	.	.	.

En avril 1986 un point atypique a été détecté. La valeur pour cette date du tableau D11 (106.822) est remplacée par la valeur correspondante du tableau D12 (99.974).

4.4.3 Tableau E3 : Composante irrégulière finale corrigée des points les plus atypiques

4.4.3.1 Description et mode de calcul

Ce tableau présente la composante irrégulière du tableau D13 corrigée des points atypiques ayant reçu un poids nul lors de l'étape C17. Lorsque, à une date donnée, un point a été jugé très atypique, la valeur de la composante est remplacée par la moyenne théorique (1 ou 0 selon le schéma).

4.4.3.2 Exemple

E3 : COMPOSANTE IRRÉGULIÈRE CORRIGÉE DES VALEURS ATYPIQUES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	99.954	100.970	99.281
1986	100.676	99.653	99.099	100.000	99.569	100.182	100.533	100.000	99.670	100.824	99.781	99.403
1987	100.000	100.497	100.837	99.721	100.466	99.319	100.495	99.027	100.544	99.579	100.536	100.410
1988	99.416	99.859	100.020	99.587	101.167	100.154	99.169	99.792	100.465	100.000	98.616	101.281
1989	100.192	100.244	100.000	100.000	98.779	101.571	99.973	99.260	99.980	99.593	100.711	100.402
1990	99.865	99.352	100.315	99.900	100.459	98.858	100.342	100.927	99.593	100.535	100.040	98.587
1991	100.834	100.041	99.502	100.659	99.046	100.752	102.212	100.138	99.441	99.789	100.525	99.617
1992	99.931	100.478	100.018	100.294	100.092	98.995	99.407	100.461	100.653	100.758	99.833	98.148
1993	98.999	100.000	100.691	97.966	100.338	100.045	100.112	100.420	99.512	99.565	99.603	100.134
1994	100.955	100.045	99.280	99.774	101.018	100.102	100.356	100.000	100.022	99.235	99.609	101.212
1995	100.153	99.499	100.038	.	.	.	.	.	.	.	.	.

4.4.4 Tableau E4 : Comparaison des totaux annuels entre Série brute et Série désaisonnalisée

4.4.4.1 Description et mode de calcul

Ce tableau compare les totaux annuels des séries brute et désaisonnalisée, et ce pour deux couples de séries : la série brute A1 et la série désaisonnalisée finale D11 d'une part et les séries correspondantes corrigées des points atypiques (E1 et E2). Pour chacun de ces couples et pour chaque année n, on calcule par exemple $A1_n$ *_op* $D11_n$.

4.4.4.2 Exemple

E4 : COMPARAISON DES TOTAUX ANNUELS

	A1 et D11	E1 et E2
1986	100.079	99.987
1987	100.119	100.137
1988	100.080	100.098
1989	99.807	99.832
1990	99.878	99.878
1991	99.965	99.965
1992	100.384	100.384
1993	99.941	99.945
1994	99.719	99.762

Ainsi, pour l'année 1987, on a : $\frac{1252.8}{1251.312} * 100 = 100.119$ et $\frac{1259.372}{1257.651} = 100.137$

4.4.5 Tableau E5 : évolution de la série brute

4.4.5.1 Description et mode de calcul

Ce tableau présente l'évolution de la série brute. Pour une série mensuelle par exemple, il donne les accroissements mensuels lorsque le schéma de composition est additif ou le taux d'accroissement mensuel dans le cas d'un schéma multiplicatif. Pour une date t donnée, on a donc : $E5_t = A1_t$ _op $A1_{t-1}$ - _xbar

4.4.5.2 Exemple

E5 : ÉVOLUTION DE LA SÉRIE ORIGINALE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	-5.099	-8.379
1986	5.964	-7.411	5.268	5.390	-10.776	6.141	-3.857	-34.102	60.122	11.312	-7.515	-3.601
1987	-3.736	2.687	9.399	-5.137	-6.629	8.300	-6.002	-32.515	58.224	7.544	-1.882	-4.098
1988	-2.091	2.321	7.713	-8.930	-0.648	6.797	-11.770	-24.901	50.789	2.880	2.884	-5.441
1989	2.790	-4.835	7.130	-4.576	-3.662	8.869	-12.219	-24.811	43.829	10.946	0.079	-11.120
1990	7.453	-7.102	9.867	-6.068	-0.431	1.038	-4.281	-25.492	37.575	15.183	-3.712	-12.825
1991	11.282	-8.516	5.762	0.084	-5.109	3.001	-1.200	-29.228	42.647	13.746	-5.740	-7.212
1992	6.649	-5.344	6.074	-3.226	-8.500	8.106	-5.560	-28.635	49.125	8.131	-5.349	-6.798
1993	-0.088	-0.528	8.488	-6.927	-5.517	8.526	-7.686	-26.272	44.040	5.401	0.579	-5.670
1994	1.307	-4.127	11.211	-6.935	-1.213	6.140	-9.504	-22.009	41.218	4.809	1.028	-6.030
1995	3.417	-6.285	11.952	.	.	.	.	.	.	.	.	.

Ainsi, le taux d'accroissement de mars 1986 à avril1986 est : $APR86 = 100 * \left(\frac{109.5}{103.9} - 1 \right) = 5.390$

4.4.6 Tableau E6 : évolution de la série désaisonnalisée finale

4.4.6.1 Description et mode de calcul

Ce tableau présente l'évolution de la série désaisonnalisée finale (tableau D11) calculée de la même façon que ci-dessus. Pour une date t donnée, on a donc :
$E6_t = D11_t$ _op $D11_{t-1}$ - _xbar

4.4.6.2 Exemple

E6 : ÉVOLUTION DE LA SÉRIE DÉSAISONNALISÉE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	0.638	-2.104
1986	0.948	-1.400	-0.714	8.000	-6.368	1.261	0.980	-2.795	2.691	1.257	-0.923	-0.143
1987	-5.221	7.549	0.718	-0.869	0.813	-1.198	1.194	-1.274	1.905	-0.502	1.417	0.314
1988	-0.533	0.952	0.651	0.032	2.016	-0.677	-0.729	0.936	1.153	-2.310	1.901	3.382
1989	-0.653	0.274	-3.176	5.585	-3.757	2.741	-1.523	-0.552	0.978	-0.097	1.403	-0.041
1990	-0.324	-0.419	1.087	-0.205	0.866	-1.230	1.760	0.635	-1.468	0.615	-0.892	-1.818
1991	2.045	-0.793	-0.370	1.424	-1.351	1.913	1.581	-1.943	-0.649	0.419	0.857	-0.759
1992	0.428	0.474	-0.727	-0.103	-0.511	-1.240	0.442	1.140	0.122	-0.211	-1.450	-2.305
1993	0.346	3.278	-1.794	-2.622	2.443	-0.446	-0.245	-0.016	-1.052	0.142	0.339	1.010
1994	1.370	-0.355	-0.266	0.947	1.655	-0.598	0.503	1.728	-1.497	-0.378	0.909	2.166
1995	-0.565	-0.257	0.830	.	.	.	.	.	.	.	.	.

Ainsi, le taux d'accroissement de mars 1986 à avril1986 est : $APR86 = 100 * \left(\frac{106.822}{98.909} - 1 \right) = 8.000$

4.4.7 Tableau E7 : évolution de la Tendance-Cycle finale

4.4.7.1 Description et mode de calcul

Ce tableau présente l'évolution de la composante Tendance-Cycle finale (tableau D12) calculée de la même façon que ci-dessus. Pour une date t donnée, on a donc :
$E7_t = D12_t$ _op $D12_{t-1}$ - _xbar

4.4.7.2 Remarque

Ce tableau n'est édité que par X12-ARIMA

4.4.7.3 Exemple

E7 : ÉVOLUTION DE LA TENDANCE-CYCLE FINALE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	-0.375	-0.439
1986	-0.450	-0.388	-0.158	0.166	0.479	0.642	0.628	0.466	0.217	0.099	0.113	0.236
1987	0.374	0.448	0.379	0.240	0.065	-0.057	0.010	0.190	0.367	0.462	0.452	0.440
1988	0.461	0.504	0.489	0.467	0.423	0.327	0.257	0.306	0.475	0.648	0.760	0.662
1989	0.428	0.221	0.053	-0.069	-0.133	-0.084	0.051	0.163	0.251	0.290	0.278	0.267
1990	0.211	0.095	0.117	0.210	0.304	0.369	0.255	0.052	-0.148	-0.327	-0.401	-0.371
1991	-0.229	-0.007	0.169	0.259	0.255	0.188	0.130	0.089	0.046	0.070	0.118	0.146
1992	0.112	-0.073	-0.271	-0.378	-0.311	-0.146	0.025	0.079	-0.070	-0.315	-0.538	-0.627
1993	-0.517	-0.264	-0.015	0.087	0.021	-0.154	-0.312	-0.323	-0.148	0.088	0.301	0.475
1994	0.545	0.552	0.502	0.447	0.403	0.312	0.249	0.246	0.294	0.412	0.529	0.548
1995	0.487	0.398	0.286	.	.	.	.	.	.	.	.	.

Ainsi, le taux d'accroissement de mars 1986 à avril1986 est : $APR86 = 100 * \left(\frac{99.974}{99.809} - 1 \right) \approx 0.166$

4.4.8 Tableau E11 : Estimation robuste de la série désaisonnalisée finale

4.4.8.1 Description et mode de calcul

Ce tableau présente une estimation robuste de la série désaisonnalisée finale. Il est équivalent au tableau E2 sauf pour les points jugés très atypiques, c'est à dire ceux qui ont reçu un poids égal à 0 au tableau C17. La valeur de E2 aux dates correspondantes est remplacée par *D12 + (A1 - E1).*

4.4.8.2 Remarque

Ce tableau n'est édité que par X12-ARIMA

4.4.8.3 Exemple

E11 : ESTIMATION ROBUSTE DE LA SÉRIE DÉSAISONNALISÉE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	101.587	102.236	100.085
1986	101.034	99.620	98.909	106.994	100.020	101.281	102.274	100.362	102.091	103.374	102.421	102.275
1987	96.701	104.252	105.001	104.088	104.934	103.677	104.916	103.579	105.552	105.022	106.510	106.845
1988	106.275	107.287	107.985	108.020	110.198	109.453	108.654	109.671	110.936	108.130	110.433	114.168
1989	113.423	113.734	109.811	116.234	111.904	114.971	113.220	112.595	113.697	113.586	115.180	115.133
1990	114.759	114.279	115.521	115.285	116.283	114.852	116.874	117.617	115.890	116.603	115.563	113.463
1991	115.783	114.864	114.439	116.068	114.501	116.692	118.537	116.234	115.479	115.963	116.956	116.069
1992	116.566	117.119	116.267	116.147	115.553	114.120	114.624	115.931	116.072	115.827	114.147	111.515
1993	111.902	115.509	113.496	110.520	113.221	112.715	112.439	112.420	111.238	111.397	111.774	112.903
1994	114.450	114.044	113.741	114.818	116.718	116.020	116.603	118.019	116.844	116.402	117.460	120.004
1995	119.326	119.019	120.007	.	.	.	.	.	.	.	.	.

Ainsi, pour avril 1986, on a : $APR86 = 99.974 + (109.500 - 102.480) = 106.994$

4.5 PARTIE F : Mesures de qualité de la désaisonnalisation

4.5.1 Tableau F1 : Lissage de la série désaisonnalisée par une moyenne mobile MCD

4.5.1.1 Description et mode de calcul

Ce tableau présente un lissage de la série désaisonnalisée finale (tableau D11) par une moyenne mobile simple dont le calcul de l'ordre (dit "MCD" : Month for Cyclical Dominance) est expliqué au tableau F2E.

4.5.1.2 Remarques

- Si le paramètre MCD calculé est pair, on utilise alors une moyenne mobile simple centrée, c'est dire une 2xMCD.
- Quelque soit l'ordre de la moyenne mobile, on ne se préoccupe pas de l'estimation des extrémités de la série.

4.5.1.3 Exemple

Le paramètre "MCD" est ici égal à 5 (voir tableau F2). Le tableau F1 s'obtient donc simplement, à partir du tableau D11, par une moyenne mobile simple d'ordre 5.

F1 : LISSAGE DE LA SÉRIE DÉSAISONNALISÉE PAR UNE MOYENNE MOBILE MCD

MCD = 5

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	.	100.912
1986	100.377	101.294	101.281	101.330	101.861	101.962	101.016	101.687	101.915	101.915	101.419	101.851
1987	102.177	102.510	103.042	104.391	104.523	104.239	104.532	104.549	105.116	105.502	106.041	106.388
1988	106.981	107.283	107.953	108.589	108.862	109.199	109.782	109.417	109.613	110.716	111.467	112.026
1989	112.376	113.544	113.091	113.401	113.298	113.792	113.277	113.614	113.656	114.038	114.471	114.587
1990	114.974	114.995	115.225	115.244	115.763	116.182	116.303	116.367	116.509	115.827	115.460	115.255
1991	114.822	114.924	115.131	115.313	116.047	116.406	116.289	116.581	116.634	116.140	116.207	116.535
1992	116.596	116.434	116.331	115.841	115.343	115.275	115.260	115.315	115.320	114.698	113.892	113.792
1993	113.326	112.601	112.942	113.104	112.478	112.263	112.407	112.042	111.854	111.947	112.353	112.914
1994	113.383	113.991	114.754	115.068	115.580	116.556	116.961	116.898	117.186	117.866	118.007	118.442
1995	119.163	.	.	.	.	.	.	.	.	.	.	.

Le premier terme calculable est celui de décembre 1985 et on a :

$$DEC85 = \frac{101.587 + 102.236 + 100.085 + 101.034 + 99.620}{5} = 100.912$$

4.5.2 Tableau F2A : Évolutions, en valeur absolue, des principales composantes

4.5.2.1 Description et mode de calcul

Ce tableau présente les évolutions moyennes, sur plusieurs périodes (de 1 à 12 mois pour une série mensuelle, de 1 à 4 trimestres pour une série trimestrielle), de quelques séries.
Prenons l'exemple de la série brute du tableau A1. Pour un délai *d* donné (de 1 à 12 mois par exemple), on calcule : $\overline{O}_d = \frac{1}{n-d}\sum_{t=d+1}^{n}\left|A1_t \ \_op \ \mathrm{A1}_{t-d} - \_xbar\right|$, c'est à dire, pour un schéma multiplicatif, la moyenne des taux de croissance absolus sur *d* mois. Ce calcul est fait, pour chaque délai *d*, pour les 10 séries suivantes :

Tableau	Symbole	Série
A1 ou B1	$\overline{O}_d$	La série originale
D11	$\overline{CI}_d$	La série désaisonnalisée finale
D13	$\bar{I}_d$	La composante irrégulière finale
D12	$\overline{C}_d$	La Tendance-Cycle finale
D10	$\overline{S}_d$	Les facteurs saisonniers finals
A2	$\overline{P}_d$	Les coefficients de l'ajustement préalable
C18	$\overline{\overline{TD}}_d$	Les coefficients finals pour jours ouvrables
F1	$\overline{MCD}_d$	La série désaisonnalisée finale lissée par une moyenne mobile MCD
E1	$\overline{O}_d^M$	La série originale corrigée des valeurs extrêmes
E2	$\overline{CI}_d^M$	La série désaisonnalisée finale corrigée des valeurs extrêmes
E3	$\bar{I}_d^M$	La composante irrégulière finale corrigée des valeurs extrêmes

4.5.2.2 Exemple

F2A : ÉVOLUTIONS EN VALEUR ABSOLUE POUR PLUSIEURS DÉLAIS

Délai	A1 (O)	D11 (CI)	D13 (I)	D12 (C)	D10 (S)	A2 (P)	C18 (TD)	F1 (MCD)	E1 (Mod. O)	E2 (Mod. CI)	E3 (Mod. I)
1	11.03	1.34	1.29	0.29	10.73	0.00	2.46	0.34	11.02	0.90	0.86
2	11.84	1.43	1.26	0.57	11.25	0.00	2.16	0.58	11.76	1.06	0.83
3	11.54	1.55	1.21	0.83	11.47	0.00	1.26	0.78	11.46	1.23	0.79
4	11.95	1.70	1.19	1.07	11.37	0.00	2.45	1.00	11.99	1.43	0.78
5	11.22	1.72	1.08	1.30	10.69	0.00	1.93	1.23	11.37	1.57	0.74
6	12.04	1.91	1.14	1.50	12.03	0.00	1.51	1.44	12.34	1.71	0.66
7	11.74	2.07	1.12	1.70	10.91	0.00	2.35	1.64	11.93	1.90	0.75
8	12.05	2.21	1.22	1.89	11.39	0.00	1.86	1.85	12.00	2.06	0.82
9	11.85	2.44	1.17	2.07	10.68	0.00	1.17	2.03	11.81	2.22	0.74
10	12.09	2.52	1.14	2.26	10.92	0.00	2.53	2.22	12.08	2.40	0.75
11	11.04	2.65	1.10	2.44	10.32	0.00	1.84	2.40	11.24	2.60	0.75
12	3.35	2.96	1.25	2.60	0.14	0.00	1.50	2.58	3.23	2.85	0.88

Le schéma étant ici multiplicatif, ce sont les taux de croissance qui sont utilisé pour chaque délai. La valeur 11.03 obtenue pour la moyenne des taux de croissance mensuels absolus de la série originale peut se retrouver en prenant la moyenne des valeurs absolues des données du tableau E5.

4.5.3 Tableau F2B : Contributions relatives des composantes aux évolutions de la série brute

4.5.3.1 Description et mode de calcul

Ce tableau présente pour un délai *d* donné, la contribution relative de chaque composante à l'évolution de la série brute. Les composantes étant indépendantes, on doit avoir, au moins approximativement :

$$\overline{O}_d^2 \approx \overline{I}_d^2 + \overline{C}_d^2 + \overline{S}_d^2 + \overline{P}_d^2 + \overline{TD}_d^2$$

Les deux membres de cette équation n'étant pas rigoureusement égaux, on va calculer en fait :

$$\overline{I}_d^2 + \overline{C}_d^2 + \overline{S}_d^2 + \overline{P}_d^2 + \overline{TD}_d^2 = \overline{O}_d'^2$$

On calcule ensuite les rapports $100 * \overline{I}_d^2 / \overline{O}_d'^2, \cdots, 100 * \overline{TD}_d^2 / \overline{O}_d'^2$ afin d'obtenir la contribution relative de chaque composante. On calcule enfin le rapport $100 * \overline{O}_d'^2 / \overline{O}_d^2$ pour mesurer la qualité de l'approximation.

4.5.3.2 Remarques

- X12 utilise dans ce calcul la série E3, à la place de D13, comme estimation de la composante irrégulière.
- X12 calcule la quantité $\overline{O}_d'^2$ à partir de la série E1 et non de la série originale comme X11-ARIMA88.

4.5.3.3 Exemple

F2B : CONTRIBUTIONS DES COMPOSANTES À L'ÉVOLUTION DE LA SÉRIE BRUTE

Délai	D13 (I)	D12 (C)	D10 (S)	A2 (P)	C18 (TD)	TOTAL	RATIO x100
1	1.36	0.07	93.65	0.00	4.92	100.00	101.17
2	1.20	0.24	95.04	0.00	3.52	100.00	94.97
3	1.08	0.51	97.23	0.00	1.17	100.00	101.72
4	1.02	0.83	93.81	0.00	4.34	100.00	96.57
5	0.97	1.39	94.56	0.00	3.08	100.00	96.01
6	0.86	1.50	96.13	0.00	1.51	100.00	103.85
7	0.97	2.23	92.49	0.00	4.30	100.00	93.30
8	1.07	2.58	93.84	0.00	2.51	100.00	95.20
9	1.12	3.54	94.20	0.00	1.14	100.00	86.32
10	0.99	3.86	90.30	0.00	4.86	100.00	90.30
11	1.03	5.08	91.00	0.00	2.89	100.00	95.97
12	14.74	63.79	0.18	0.00	21.28	100.00	94.63

Ainsi, pour le délai 1, on a, aux erreurs d'arrondis près :

$$\overline{I}_d^2 + \overline{C}_d^2 + \overline{S}_d^2 + \overline{P}_d^2 + \overline{TD}_d^2 = \overline{O}_d'^2 = (1.29)^2 + (0.29)^2 + (10.73)^2 + (0.00)^2 + (2.46)^2 = 122.933$$

$$100 * \overline{O}_d'^2 / \overline{O}_d^2 = 100 * \frac{122.933}{(11.03)^2} \approx 101.17$$

et, par exemple, la contribution de la composante saisonnière est égale à : $100 * \frac{(10.73)^2}{122.933} \approx 93.65$

4.5.4 Tableau F2C : Moyennes et écart-types des évolutions en fonction du délai.

4.5.4.1 Description et mode de calcul

Pour chaque délai, on calcule la moyenne et l'écart-type des évolutions, en tenant compte du signe cette fois ci, de la série brute et des ses composantes ainsi que de la série moyenne MCD du tableau F1.

4.5.4.2 Exemple

F2C : MOYENNE ET ÉCART-TYPE DES ÉVOLUTIONS

Délai	A1 (O)		D13 (I)		D12 (C)		D10 (S)		D11 (CI)		F1 (MCD)	
	Avg	STD	Avg	STD	Avg	STD	Avg	STD	Avg	STD	Avg	STD
1	1.38	16.84	0.02	1.90	0.15	0.31	1.15	16.20	0.17	1.92	0.15	0.40
2	1.96	20.42	0.01	1.90	0.30	0.60	1.61	20.04	0.30	2.00	0.31	0.62
3	1.94	19.26	0.01	1.77	0.45	0.86	1.45	18.95	0.47	1.96	0.46	0.82
4	1.95	17.71	0.02	1.85	0.61	1.09	1.25	16.95	0.63	2.14	0.61	1.04
5	2.31	18.40	0.01	1.59	0.77	1.28	1.52	18.23	0.79	1.99	0.75	1.25
6	2.27	17.14	0.02	1.79	0.93	1.45	1.28	16.60	0.95	2.22	0.90	1.42
7	2.75	19.93	-0.05	1.66	1.09	1.62	1.61	18.99	1.04	2.28	1.04	1.60
8	2.35	17.87	-0.02	1.73	1.25	1.78	1.04	17.07	1.22	2.46	1.20	1.79
9	2.45	17.28	-0.02	1.86	1.40	1.95	1.02	16.52	1.37	2.69	1.34	1.97
10	2.96	19.42	-0.03	1.64	1.54	2.13	1.32	18.20	1.51	2.69	1.48	2.15
11	2.88	15.66	0.01	1.60	1.68	2.31	1.12	14.89	1.69	2.81	1.62	2.32
12	1.86	3.62	0.01	1.71	1.82	2.49	0.02	0.22	1.82	3.03	1.76	2.49

Ainsi, la moyenne des taux de croissance mensuels de la série brute (1.38) correspond à la moyenne du tableau E5.

4.5.5 Tableau F2D : Durées moyennes des phases de croissance et décroissance.

4.5.5.1 Description et mode de calcul

On calcule pour quelques séries la durée moyenne des phases de croissance et décroissance. Pour cela, on considère la série des évolutions (taux de croissance si le schéma est multiplicatif, accroissements s'il est additif). La durée d'une phase de croissance (décroissance) est le nombre de termes positifs (négatifs) qui se succèdent. Si un terme est nul, il est compté dans la phase en cours.

Les durées moyennes sont calculées pour les séries suivantes :

Tableau	Symbole	Série
D11	**CI**	La série désaisonnalisée finale
D13	*I*	La composante irrégulière finale
D12	*C*	La Tendance-Cycle finale
F1	*MCD*	La série désaisonnalisée finale lissée par une moyenne mobile MCD

4.5.5.2 Exemple

A titre d'exemple, considérons la Tendance-Cycle du tableau D12. La série des taux de croissance mensuels est :

TAUX DE CROISSANCE MENSUELS DE LA TENDANCE-CYCLE D12

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	.	-0.375	-0.439
1986	-0.450	-0.388	-0.158	0.166	0.479	0.642	0.628	0.466	0.217	0.099	0.113	0.236
1987	0.374	0.448	0.379	0.240	0.065	-0.057	0.010	0.190	0.367	0.462	0.452	0.440
1988	0.461	0.504	0.489	0.467	0.423	0.327	0.257	0.306	0.475	0.648	0.760	0.662
1989	0.428	0.221	0.053	-0.069	-0.133	-0.084	0.051	0.163	0.251	0.290	0.278	0.267
1990	0.211	0.095	0.117	0.210	0.304	0.369	0.255	0.052	-0.148	-0.327	-0.401	-0.371
1991	-0.229	-0.007	0.169	0.259	0.255	0.188	0.130	0.089	0.046	0.070	0.118	0.146
1992	0.112	-0.073	-0.271	-0.378	-0.311	-0.146	0.025	0.079	-0.070	-0.315	-0.538	-0.627
1993	-0.517	-0.264	-0.015	0.087	0.021	-0.154	-0.312	-0.323	-0.148	0.088	0.301	0.475
1994	0.545	0.552	0.502	0.447	0.403	0.312	0.249	0.246	0.294	0.412	0.529	0.548
1995	0.487	0.398	0.286	.	.	.	.	.	.	.	.	.

La série des durées des phases de croissance ou décroissance est :

{5, 14, 1, 21, 3, 14, 6, 11, 5, 2, 7, 2, 4, 18} soit une durée moyenne de $\frac{113}{14}=8.0714$

Le tableau F2D est donc :

F2D : DURÉE MOYENNE DES PHASES DE CROISSANCE OU DÉCROISSANCE

D11 (CI)	D13 (I)	D12 (C)	F1 (MCD)
1.6377	1.5067	8.071	3.2059

4.5.6 Tableau F2E : Calcul du ratio MCD ("Month for Cyclical Dominance").

4.5.6.1 Description et mode de calcul

A partir des données du tableau F2A, on calcule les rapports $\bar{I}_d / \bar{C}_d$ pour chaque valeur du délai *d*. La valeur MCD, utilisée notamment pour le tableau F1, est celle du premier délai *d* pour lequel $\bar{I}_d / \bar{C}_d < 1$.

4.5.6.2 Exemple

Les différentes valeurs du ratio se calculent aisément à partir des colonnes D13 et D12 du tableau F2A. Ainsi, pour le délai 1, on a, aux arrondis près, $\frac{\bar{I}_1}{\bar{C}_1} = \frac{1.29}{0.29} \approx 4.46$. D'où le tableau F2E :

F2E : RATIO I/C et MCD

1	2	3	4	5	6	7	8	9	10	11	12
4.46	2.22	1.45	1.11	0.84	0.76	0.66	0.65	0.56	0.51	0.45	0.48

La valeur du MCD est donc 5.

4.5.7 Tableau F2F : Contribution relative des composantes à la variance de la partie stationnaire de la série originale

4.5.7.1 Description et mode de calcul

Pour rendre stationnaire la série originale, on extrait une tendance linéaire si le schéma est additif ou une tendance exponentielle si le schéma est multiplicatif. Cette tendance n'est pas estimée sur la série brute (tableau A1) mais sur la composante Tendance-Cycle du tableau D12. Cette dernière est d'ailleurs aussi rendue stationnaire en enlevant la même tendance. Enfin, la contribution relative de chaque composante (irrégulier D13, Tendance-Cycle stationarisée, saisonnalité D10, coefficients d'ajustement a priori, coefficients pour jours ouvrables C18) est évaluée.

4.5.7.2 Remarques

- Dans le calcul des variances des composantes, on utilise les moyennes théoriques (_xbar) pour les composantes saisonnière, irrégulière, coefficients pour jours ouvrables et les moyennes empiriques pour la série brute, la composante Tendance-Cycle et les coefficients d'ajustement a priori.
- Comme estimation de la composante irrégulière X12-ARIMA utilise le tableau E3 et X11-ARIMA88 D13.

4.5.7.3 Exemple

Il serait fastidieux de décrire chaque étape de ce calcul assez long. On se contentera de résumer chaque étape du calcul. Le schéma étant multiplicatif, on se trouve dans le cas où la méthode employée est la plus complexe.

- On ajuste une droite, notée *ols*, par les moindres carrés ordinaires, au logarithme de la composante Tendance-Cycle du tableau D12, ce qui revient à ajuster une tendance exponentielle à la série D12.
- La série originale A1 et la Tendance-Cycle D12 sont alors rendues stationnaires en leur enlevant cette tendance exponentielle : *A1bis = A1 _op EXP(ols)* et *D12bis = D12 _op EXP(ols).*
- On dispose donc maintenant d'une série A1bis décomposée ici en 4 composantes indépendantes : D12bis, D13, D10 et C18. On va transformer ces variables par logarithme afin de se ramener à la somme de variables aléatoires indépendantes, ce qui permet d'utiliser l'égalité approximative :
  $$Var[LOG(A1bis)] \approx Var[LOG(D12bis)] + Var[LOG(D13)] + Var[LOG(D10)] + Var[LOG(C18)]$$
- Chacune de ces variances est alors calculée en utilisant les moyennes empiriques pour *LOG(A1bis)* et *LOG(D12bis)* et la moyenne théorique (ici 0 puisqu'on est en transformation logarithmique) pour *LOG(D10), LOG(C18)* et *LOG(D13).*
- Enfin, les contributions de la variance de chaque composante à la variance de la série originale stationnarisée sont éditées dans le tableau F2F :

F2F : CONTRIBUTION RELATIVE DES COMPOSANTES À LA VARIANCE DE LA SÉRIE BRUTE STATIONNARISÉE

I	C	S	TD	Total
1.09	5.36	91.50	1.91	99.86

4.5.8 Tableau F2G : Autocorrélogramme de la composante irrégulière

4.5.8.1 Description et mode de calcul

Ce tableau présente l'autocorrélogramme de la composante irrégulière D13 calculé pour des retards allant de 1 au nombre de périodes +2 (soit 14 pour une série mensuelle et 6 pour une série trimestrielle). Variance et autocorrélations sont calculées avec la moyenne théorique _xbar.

On a donc, avec N observations dans la série I : $Corr_k(I) = \dfrac{\sum_{t=k+1}^{N}(I_t - \_xbar)(I_{t-k} - \_xbar)}{\sum_{t=1}^{N}(I_t - \_xbar)^2}$

4.5.8.2 Exemple

Ici, la moyenne théorique est 1. Le calcul des autocorrélations jusqu'à l'ordre 14 donne :

F2G : AUTOCORRÉLOGRAMME DE LA COMPOSANTE IRRÉGULIÈRE

1	2	3	4	5	6	7	8	9	10	11	12	13	14
-0.15	-0.15	0.00	-0.10	0.21	0.00	0.00	-0.07	-0.26	0.05	0.08	-0.05	0.02	-0.08

4.5.9 Tableau F2H : Ratios $\frac{\bar{I}}{\bar{\bar{C}}}$ et $\frac{\bar{I}}{\bar{\bar{S}}}$

4.5.9.1 Description et mode de calcul

Ce tableau rappelle les valeurs des ratios $\frac{\bar{I}}{\bar{\bar{C}}}$ et $\frac{\bar{I}}{\bar{\bar{S}}}$ calculés respectivement aux tableaux D12 et D10.

4.5.9.2 Exemple

F2H : RATIOS

RATIO I/C FINAL POUR D12	2.74
RATIO I/S FINAL POUR D10	4.60

4.5.10 Tableau F2I : Rappel des résultats des tests d'existence de saisonnalité

4.5.10.1 Description et mode de calcul

Ce tableau rappelle les valeurs des tests d'existence de la saisonnalité effectués aux tableaux B1 (test de Fisher de saisonnalité stable) et D8 (test de Fisher de saisonnalité, test de Kruskal-Wallis de saisonnalité, test de Fisher de saisonnalité évolutive).

4.5.10.2 Remarque

X11-ARIMA88 imprime aussi le résultat du test d'existence d'un effet lié aux jours ouvrables (tableau C15).

4.5.10.3 Exemple

F2I : RAPPEL DES DIFFÉRENTES STATISTIQUES DE TEST

	STATISTIQUE	PROB > Stat (%)
Test F de saisonnalité stable (B1)	183.698	0.000
Test F d'effet pour jours ouvrables (C15)	68.245	0.000
Test F de saisonnalité stable (D8)	498.194	0.000
Test de Kruskal-Wallis de saisonnalité stable (D8)	104.780	0.000
Test F de saisonnalité évolutive (D8)	1.724	10.386

4.5.11 Tableau F3 : Statistiques de qualité de l'ajustement

4.5.11.1 Description et mode de calcul

Ce tableau présente onze statistiques permettant de juger la qualité de l'ajustement saisonnier. Leur calcul et leur justification sont détaillés dans Lothian & Morry (1978). Ces statistiques varient entre 0 et 3, mais seules les valeurs en dessous de 1 sont jugées acceptables. Un indicateur synthétique de la qualité de la désaisonnalisation est construit comme combinaison linéaire de ces 11 statistiques.

Statistique M1 :
Si dans l'évolution de la série la part de la composante irrégulière est trop forte, il sera difficile d'identifier et d'extraire une composante saisonnière. La statistique *M1* mesure la contribution de l'irrégulier à la variance totale à partir des résultats du tableau F2B. Lothian & Morry (1978) ont montré que, pour une série mensuelle, c'était le retard 3 qui permettait la meilleure comparaison des contributions respectives des composantes irrégulière et saisonnière.
La statistique *M1* tient aussi compte de la présence éventuelle de coefficients d'ajustement préalable et se définit, en reprenant les notations du tableau F2B par :

$$M1 = 10 * \frac{\bar{I}_3^2 / \bar{O}'^2_3}{1 - \bar{P}_3^2 / \bar{O}'^2_3}$$

La statistique *M1*, normalisée par le facteur 10, est jugée acceptable si elle ne dépasse pas 10%.

Statistique M2
Cette statistique, similaire à la statistique *M1*, est calculée à partir de la contribution de la composante irrégulière à la variance de la série brute stationnarisée.
La statistique *M2* tient aussi compte de la présence éventuelle de coefficients d'ajustement préalable et se calcule à partir des données du tableau F2F par :

$$M2 = 10 * \frac{Contr(I)}{1 - Contr(P)}$$

La statistique *M2*, normalisée par le facteur 10, est jugée acceptable si elle ne dépasse pas 10%.

Statistique M3
Pour obtenir une désaisonnalisation correcte, X11 estime successivement chacune des composantes. En particulier, dans l'extraction de la composante Tendance-Cycle, il est souhaitable que la contribution de l'irrégulier à l'évolution de l'estimation provisoire de la série désaisonnalisée ne soit pas trop importante. Dans le cas contraire, il sera difficile de séparer ces deux composantes. La statistique *M3* mesure cette contribution à partir du ratio I/C du tableau D12, repris dans le tableau F2H. On a :

$$M3 = \frac{1}{2} * \left(\frac{\bar{I}}{\bar{C}} - 1 \right)$$

Statistique M4
L'une des hypothèses de base conditionnant la validité des tests de Fisher fait au cours du traitement X11 est la nature aléatoire de la composante irrégulière. La statistique M4 teste la présence d'autocorrélation à partir de la durée moyenne des séquences (ADR) de l'irrégulier qui figure au tableau F2D, selon la formule :

$$M4 = \frac{\left| \dfrac{N-1}{ADR} - \dfrac{2(N-1)}{3} \right|}{2.577 * \sqrt{\dfrac{16N-29}{90}}}$$

Statistique M5

La valeur MCD calculée au tableau F2E mesure le nombre de mois nécessaires pour que les variations absolues de la composante Tendance-Cycle l'emportent sur celles de la composante irrégulière. Cette mesure, comme *M3*, permet donc de comparer l'importance des évolutions des composantes tendancielle et irrégulière.

Le MCD est le nombre de mois k tel que : $\frac{\bar{I}_k}{\bar{C}_k} \leq 1$ et $\frac{\bar{I}_{k-1}}{\bar{C}_{k-1}} > 1$. La valeur utilisée ici utilise une interpolation linéaire : $MCD' = (k-1) + \left(\frac{\bar{I}_{k-1}}{\bar{C}_{k-1}} - 1 \right) / \left(\frac{\bar{I}_{k-1}}{\bar{C}_{k-1}} - \frac{\bar{I}_k}{\bar{C}_k} \right)$. Il est généralement admis que, pour une série mensuelle, cette valeur ne doit pas dépasser 6, ce qui permet de définir la statistique *M5* comme :

$$M5 = \frac{MCD' - 0.5}{5}$$

Statistique M6

Pour extraire la composante saisonnière, X11 lisse une estimation du mouvement Saisonnier-Irrégulier par exemple à l'aide d'une moyenne mobile 3x5. L'expérience a montré que si les évolutions annuelles de la composante irrégulière étaient trop faibles en regard des évolutions annuelles de la composante saisonnière (faible ratio I/S), la moyenne 3x5 n'était pas assez flexible pour suivre le mouvement saisonnier. Lothian (1978) a montré que cette moyenne 3x5 fonctionne bien pour des valeurs du ratio I/S comprises entre 1.5 et 6.5. La statistique M6 est dérivée de ces valeurs et du ratio I/S du tableau D10, repris dans le tableau F2H. On a :

$$M6 = \frac{1}{2.5} * \left(\frac{\bar{I}}{\bar{S}} - 4 \right)$$

Statistique M7

M7 compare, à partir des statistiques des tests de Fisher du tableau D8, la part relative des saisonnalité stable (statistique F_s) et mobile (statistique F_M). Elle permet de voir si la saisonnalité est ou non identifiable par X11. On a :

$$M7 = \sqrt{\frac{1}{2}\left(\frac{7}{F_s} + \frac{3F_M}{F_S} \right)}$$

Les filtres saisonniers utilisés par X11 fonctionnent de façon optimale pour des saisonnalités constantes. Si ce mouvement saisonnier évolue au cours des ans, les estimations des coefficients saisonniers peuvent être plus ou moins erronées. On considère deux types de mouvements : celui dû à des variations quasi aléatoires de court terme et un autre dû à des évolutions de plus long terme. L'importance du premier type de mouvement peut être mesurée par la moyenne des évolutions absolues annuelles (statistiques *M8* et *M10*). La moyenne simple des évolutions annuelles donne au contraire une idée de l'importance d'un mouvement (linéaire) systématique (statistiques *M9* et *M11*).

Ces 4 dernières statistiques sont calculées à partir de facteurs saisonniers normalisés. À la composante saisonnière du tableau D10, on applique la transformation : $S'_t = \frac{S_t - \bar{S}}{\sigma(S)}$, où la moyenne utilisée est la moyenne théorique _xbar.

Statistique M8
L'ampleur des variations de la composante saisonnière est mesurée par la variation absolue moyenne :

$\left|\overline{\Delta S'}\right| = \frac{1}{J(N-1)} \sum_{j=1}^{J} \sum_{i=2}^{N} \left| S'_{Ji+j} - S'_{J(i-1)+j} \right|$ où N désigne le nombre d'années de la série et J la périodicité (4 ou 12). On a alors, si la limite de tolérance est de 10% : $M8 = 100 * \left|\overline{\Delta S'}\right| * \frac{1}{10}$.

Statistique M9
On a, si la limite de tolérance est de 10% :

$$M9 = 100 * \frac{\sum_{j=1}^{J} \left| S'_{J(N-1)+j} - S'_{j} \right|}{J*(N-1)} * \frac{1}{10}$$

Statistique M10
C'est l'équivalent de la statistique *M8* mesurée sur les dernières années :

$\left|\overline{\Delta S'}\right|_R = \frac{1}{3J} \sum_{j=1}^{J} \sum_{i=N-4}^{N-2} \left| S'_{Ji+j} - S'_{J(i-1)+j} \right|$ et $M10 = 100 * \left|\overline{\Delta S'}\right|_R * \frac{1}{10}$

Statistique M11
C'est l'équivalent de la statistique *M9* mesurée sur les dernières années :

$$M11 = 100 * \frac{\sum_{j=1}^{J} \left| S'_{J(N-2)+j} - S'_{J(N-5)+j} \right|}{3J} * \frac{1}{10}$$

Enfin, une statistique globale de qualité, combinaison linéaire des statistiques *M1* à *M11*, est calculée. On a :

$$Q = \frac{10M1 + 11M2 + 10M3 + 8M4 + 11M5 + 10M6 + 18M7 + 7M8 + 7M9 + 4M10 + 4M11}{100}.$$

4.5.11.2 Remarques

- X12-ARIMA utilise la série du tableau E3 comme estimation de la composante irrégulière, au lieu de D13 pour X11-ARIMA88 ce qui se répercute sur la définition des statistiques *M1* et *M2*.
- La statistique M6 n'a de sens que si une moyenne mobile 3x5 a été utilisée. Dans le cas contraire, son poids dans la statistique Q est nul.
- Les statistiques M8 à M11 ne sont calculables que si la série couvre au moins 6 années. Dans le cas contraire, le vecteur de poids utilisé pour calculer la statistique Q est (14, 15, 10, 8, 11, 10, 32, 0, 0, 0, 0).
- X12-ARIMA calcule aussi une statistique globale *Q2* qui ne prend pas en compte *M2*.

4.5.11.3 Exemple

Le tableau F3 édité pour notre exemple est :

F3 : STATISTIQUES DE QUALITÉ DE LA DÉSAISONNALISATION

M1 =	0.108
M2 =	0.109
M3 =	0.871
M4 =	0.029
M5 =	0.779
M6 =	0.241
M7 =	0.111
M8 =	0.126
M9 =	0.099
M10 =	0.163
M11 =	0.151

Statistique Q = 0.270

On a :

$$M1 = 10 * \frac{1.08/100}{1 - 0/100} = 0.108$$

$M2 = 10 * \frac{1.09/100}{1 - 0/100} = 0.109$ (compte tenu du fait que les résultats du tableau F2F sont exprimés en %).

$$M3 = \frac{1}{2}(2.742 - 1) = 0.871$$

$$M4 = \frac{\left| \frac{114-1}{1.5067} - \frac{2*(114-1)}{3} \right|}{2.577 * \sqrt{\frac{16*114-29}{90}}} \approx \frac{0.3350}{11.509} \approx 0.029$$

$$M5 = \frac{1}{5}\left[4 + \frac{1.14 - 1}{1.14 - 0.81} - 0.5 \right] \approx 0.78$$

$$M6 = \frac{1}{2.5}\left| 4.602 - 4 \right| = 0.241$$

$$M7 = \sqrt{\frac{1}{2}\left(\frac{7 + 3*1.724}{498.194} \right)} = 0.111$$

Le calcul des autres statistiques demande une normalisation des coefficients saisonniers du tableau D10. On calcule l'écart-type de ce tableau par la formule :

$$\sigma = \sqrt{\frac{1}{N}\sum_{t=1}^{t=N}(S_t - 1)^2} = 0.1001$$

On normalise ensuite le tableau D10 par la formule $S'_t = \frac{S_t - 1}{\sigma}$. Ainsi, pour la valeur de avril 1986, on a : $APR86 = \frac{1.01408 - 1}{0.1001} \approx 0.1407$.

On obtient alors le tableau suivant :

FACTEURS SAISONNIERS STANDARDISÉS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	1.1433	0.8866	0.0620
1986	0.3781	-0.0039	0.7670	0.1407	-0.2159	0.3333	-0.4609	-3.2194	0.1906	1.1416	0.8816	0.0721
1987	0.3843	-0.0125	0.7630	0.1408	-0.2261	0.3303	-0.4552	-3.1985	0.1877	1.1480	0.8678	0.0738
1988	0.3932	-0.0259	0.7552	0.1446	-0.2434	0.3280	-0.4472	-3.1596	0.1834	1.1506	0.8655	0.0628
1989	0.4053	-0.0467	0.7362	0.1523	-0.2565	0.3146	-0.4391	-3.0972	0.1820	1.1653	0.8494	0.0393
1990	0.4119	-0.0677	0.7160	0.1638	-0.2723	0.3089	-0.4237	-3.0308	0.1769	1.1690	0.8413	0.0099
1991	0.4199	-0.0920	0.6910	0.1733	-0.2664	0.2921	-0.4020	-2.9658	0.1729	1.1709	0.8164	-0.0127
1992	0.4213	-0.1123	0.6758	0.1825	-0.2544	0.2862	-0.3797	-2.9240	0.1678	1.1544	0.7951	-0.0146
1993	0.4205	-0.1262	0.6666	0.1860	-0.2310	0.2774	-0.3700	-2.9006	0.1712	1.1392	0.7672	0.0018
1994	0.4158	-0.1357	0.6670	0.1885	-0.2223	0.2786	-0.3685	-2.8937	0.1747	1.1299	0.7546	0.0158
1995	0.4107	-0.1411	0.6715	.	.	.	.	.	.	.	.	.

On en déduit les évolutions annuelles en retranchant à la ligne *t* la ligne *t-1*. Ainsi, l'évolution de la saisonnalité entre les mois d'avril 1986 et avril 1987 est proche de 0.
On obtient le tableau suivant :

ÉVOLUTION ANNUELLE DES FACTEURS SAISONNIERS STANDARDISÉS

	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	0.00168	0.00501	-0.01003
1986	-0.00612	0.00867	0.00406	-0.00018	0.01011	0.00302	-0.00577	-0.02089	0.00299	-0.00636	0.01377	-0.00173
1987	-0.00890	0.01333	0.00779	-0.00380	0.01736	0.00227	-0.00794	-0.03897	0.00425	-0.00259	0.00232	0.01103
1988	-0.01218	0.02085	0.01902	-0.00761	0.01306	0.01341	-0.00815	-0.06237	0.00141	-0.01470	0.01611	0.02350
1989	-0.00655	0.02095	0.02013	-0.01153	0.01582	0.00568	-0.01539	-0.06638	0.00514	-0.00370	0.00803	0.02938
1990	-0.00799	0.02429	0.02500	-0.00955	-0.00594	0.01687	-0.02166	-0.06503	0.00400	-0.00188	0.02491	0.02254
1991	-0.00145	0.02031	0.01527	-0.00922	-0.01193	0.00581	-0.02228	-0.04180	0.00506	0.01651	0.02138	0.00194
1992	0.00085	0.01391	0.00912	-0.00347	-0.02338	0.00884	-0.00971	-0.02342	-0.00341	0.01518	0.02789	-0.01640
1993	0.00471	0.00951	-0.00030	-0.00249	-0.00870	-0.00121	-0.00155	-0.00688	-0.00350	0.00926	0.01255	-0.01402
1994	0.00506	0.00538	-0.00453	.	.	.	.	.	.	.	.	.

La statistique *M8* se déduit à partir de la moyenne des valeurs absolues de ce tableau; c'est en fait 10 fois cette moyenne, soit 0.126.
De même, la statistique M10 se déduit à partir de la moyenne des valeurs absolues des données de ce tableau, de avril 1990 à mars 1993. M10 est égale à 10 fois cette moyenne, soit 0.163.

Pour le calcul de M9, on peut par exemple faire la moyenne, multipliée par 10, des valeurs absolues des moyennes mensuelles de ce tableau. Ainsi, la moyenne de la colonne du mois d'avril est :

$$Avril = \frac{-0.00018 - 0.00380 - 0.00761 - 0.01153 - 0.00955 - 0.00922 - 0.00347 - 0.00249}{8} \approx -0.00598$$

De même, pour les autres mois, on aurait :

JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
-0.00362	0.01525	0.01062	-0.00598	0.00080	0.00688	-0.01156	-0.04072	0.00199	0.00149	0.01466	0.00513

Et 10 fois la moyenne des valeurs absolues de ces données conduit à la valeur de M9 :

$$\mathrm{M9} = \frac{0.0362 + 0.1525 + 0.1062 + 0.0598 + 0.0080 + 0.0688 + 0.1156 + 0.4072 + 0.0199 + 0.0149 + 0.1466 + 0.0513}{12}$$
$$\approx 0.099$$

Pour M11, on ferait la même chose mais sur les données de avril 1990 à mars 1993, ce qui conduit à la valeur 0.151.
Enfin, on déduit la valeur de la statistique Q finale :

$$Q = \frac{1.08 + 1.199 + 8.71 + 0.233 + 8.569 + 2.45 + 1.998 + 0.882 + 0.693 + 0.652 + 0.604}{100} \approx 0.27$$

Chapitre 5 MODÉLISATION DE L'EFFET DE PÂQUES

X11-ARIMA88 et X12-ARIMA proposent différents modèles de correction de l'effet de Pâques ***à partir d'une estimation de la composante irrégulière***. Les modèles proposés et la méthodologie utilisée sont parfois assez différents dans les deux logiciels et c'est pourquoi nous ne l'avons pas directement intégrée dans l'exemple ci-dessus.

5.1 La fête de Pâques

5.1.1 Un peu d'histoire

Selon les Évangiles, la résurrection du Christ a eu lieu pendant la *Pâque juive*, fête judaïque annuelle célébrée après la première pleine lune de Printemps et au cours de laquelle un agneau est sacrifié[33] pour commémorer l'exode d'Égypte. Les Chrétiens souhaitèrent conserver le lien symbolique entre ce sacrifice et le sacrifice de Jésus et c'est pourquoi il fut décidé, au concile de Nicée en 325, que la fête chrétienne de Pâques serait célébrée le premier dimanche après la première pleine lune suivant le solstice de Printemps. Malheureusement, l'ancien calendrier julien était basé sur une année un peu plus longue que celle en usage aujourd'hui et, peu à peu, le « solstice de Printemps » se rapprochait des mois d'hiver. Lorsqu'en 1582 le Pape Grégoire XIII introduisit le calendrier grégorien encore en usage de nos jours, le solstice de Printemps était presque en Février et l'un des principaux objectifs du changement de calendrier fut sans aucun doute de ramener la fête de Pâques au Printemps !

5.1.2 Le calcul des dates de Pâques

La détermination à l'avance des dates de Pâques a fait l'objet de travaux de mathématiciens célèbres et Gauss lui même est l'auteur d'algorithmes savants mais hélas complexes. Gardner (1981) cite un algorithme simple, dû à Thomas H. O'Beirne (1966), valable de 1900 à 2099 inclus :

- Soit Y l'année. Soustraire 1900 de Y et soit N la différence.
- Diviser N par 19 ; soit A le reste de cette division.
- Diviser (7A+1) par 19 et soit B le quotient de la division dont on ignore le reste.
- Diviser (11A+4-B) par 29 ; soit M le reste de cette division.
- Diviser N par 4 et soit Q le quotient de la division dont on ignore le reste.
- Diviser (N+Q+31-M) par 7 ; soit W le reste.
- La date de Pâques est alors 25-M-W. Si le résultat est positif, le mois est avril, et mars dans le cas contraire (en interprétant 0 comme le 31 mars, -1 comme le 30 mars et ainsi de suite jusqu'à –9 pour le 22 mars).

[33]« … Le Seigneur passera pendant la nuit et tuera tous les premiers nés égyptiens. Saura-t-il discerner les enfants d'Israël ? Oui, car chaque famille aura, la veille au soir, immolé un agneau; prête au départ, en tenue de voyage, ayant renoncé au pain levé des villes, elle attendra. Sur la porte elle aura inscrit un signe avec le sang de cette victime qui, semblable au bélier d'Abraham, rachète la vie des enfants de Dieu (...) La Pâque est née, la fête du «passage» ; Israël la commémorera d'année en année, au souvenir de la nuit où la puissance de mort «passa outre» et contraignit la force brutale à laisser agir Dieu. »
DANIEL-ROPS, le Peuple de la Bible, II, I.

Par exemple, pour les années 1985 à 1995 les dates de Pâques sont les suivantes :

ANNÉE	Date de Pâques
1985	7 avril
1986	30 mars
1987	19 avril
1988	3 avril
1989	26 mars
1990	15 avril
1991	31 mars
1992	19 avril
1993	11 avril
1994	3 avril
1995	16 avril

Par exemple, pour 1989, l'algorithme donne :

$$Y = 1989$$
$$N = 1989\text{-}1900 = 89$$
$$89 = 19*4 + 13 \quad \text{et donc} \quad A = 13$$
$$(7A+1)/19 = 92/19 = 19*4 + 16 \quad \text{et donc} \quad B = 4$$
$$(11A + 4\text{-}B)/29 = 143/29 = 29*4 + 27 \quad \text{et donc} \quad M = 27$$
$$N/4 = 89/4 = 22*4 + 1 \quad \text{et donc} \quad Q = 22$$
$$(N + Q + 31\text{-}M)/7 = 115/7 = 16*7 + 3 \quad \text{et donc} \quad W = 3$$
$$25\text{-}M\text{-}W = \text{-}5$$

Et donc Pâques est en mars (résultat négatif), le 26 (31-5)

La répartition des dates de Pâques sur la période 1900-2099 est illustrée à la figure 20. On note que le dimanche de Pâques tombe beaucoup plus souvent au mois d'avril qu'au mois de mars : 156 fois contre 44 fois sur cette période.

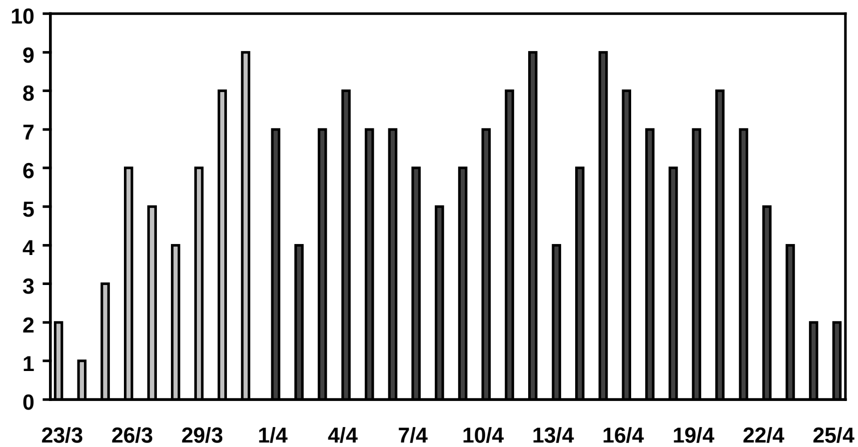
5.1.3 Pâques et la désaisonnalisation

Pourquoi se préoccupe-t-on de la fête de Pâques en analyse des séries temporelles et en désaisonnalisation ? Tout simplement parce que Pâques entraîne un changement du niveau de l'activité de nombreux secteurs : Vendredi Saint et Lundi de Pâques fériés, modification des habitudes alimentaires (chocolat, viande d'agneau …) etc. De plus, le dimanche de Pâques peut être au plus tôt le 22 mars (la dernière fois remonte à 1818 et la prochaine fois sera en 2285) et au plus tard le 25 avril (la dernière fois en 1943 et la prochaine fois en 2038) soit au premier ou au second trimestre : les effets potentiels de cette fête ne sont donc pas complètement pris en compte dans la saisonnalité de la série.

Les modèles élaborés pour la prise en compte de l'effet de Pâques dépendent de la nature même de la série.

- 1997 est la dernière année où le dimanche de Pâques est tombé en mars (le 30), ce qui ne s'était pas produit depuis 1991 où Pâques est tombé le 31 mars. Pour les ventes mensuelles d'automobiles par exemple, la variation observée entre mars et avril 1997 n'est pas comparable à celles observées entre les mêmes mois des années précédentes. La fermeture des magasins le vendredi saint a entraîné une baisse inhabituelle des ventes en mars 1997. Inversement, on a enregistré un relativement plus grand nombre de mariages en mars 1997. Pour ces séries, l'impact de Pâques, positif ou négatif, est *immédiat* dans le sens où il se concentre sur le mois dans lequel tombe Pâques.
- Pour d'autres séries, les variations liées à Pâques peuvent se faire sentir non seulement pendant la fête elle-même mais aussi pendant les jours où les semaines qui précèdent. Cet effet *graduel* s'observe par exemple dans les ventes de chocolat, de fleurs …. Dans ce cas, l'effet constaté dépend non seulement du fait que Pâques tombe en mars ou en avril, mais aussi de la date à laquelle il tombe en avril. Les chiffres de mars seront alors d'autant plus affectés que Pâques tombe tôt en avril.

Figure 20 : Répartition des dates de Pâques sur la période 1900-2099



- Enfin, la plupart des séries de stock, qui présentent l'évolution d'une grandeur mesurée à une date précise (nombre de chômeurs au 1^er^ de chaque mois par exemple) ne sont que très rarement affectées par la fête de Pâques. Il en est de même des données qui sont mesurées sur des périodes de référence, une semaine donnée en général. Ainsi, les enquêtes emploi, qui mesurent en particulier le nombre d'heures travaillées, portent elles généralement sur une semaine de référence et ne sont qu'exceptionnellement touchées par la fête de Pâques.

Les modèles que nous allons présenter par la suite s'appliquent essentiellement aux séries de flux, les séries de stock relevant plus de modèles d'intervention dans lesquels les événements particuliers comme Pâques sont modélisés sous forme de variables indicatrices particulières et leur effet ponctuel estimé par des méthodes de régression.

Les effets de Pâques peuvent enfin être estimés soit sur les données brutes, à l'aide de modèles de régression avec erreurs ARIMA, soit sur une estimation préalable de la composante irrégulière obtenu après élimination des autres effets présents dans la série (tendance, saisonnalité, effets de jours ouvrables …). La première approche n'est proposée que par X12-ARIMA, dans son module RegARIMA (Findley et *alii* 1998) et la seconde est disponible dans les deux logiciels X11-ARIMA88 et X12-ARIMA; c'est sur les modèles utilisés pour cette seconde approche que nous allons nous concentrer à présent. Nous en distinguerons six :

- Les 3 modèles proposés par X11-ARIMA88 et qui estiment l'effet de Pâques à partir de la composante irrégulière du tableau D13 : *le modèle à effet ponctuel, le modèle à effet ponctuel corrigé et le modèle à effet graduel*.
- Les 3 modèles proposés par X12-ARIMA : le *modèle de Bateman & Mayes* qui estime l'effet de Pâques à partir du tableau D13, et *deux modèles à effet graduel (« Sceaster » et « Easter »)* qui estiment l'effet de Pâques à partir des estimations de l'irrégulier des tableaux B13 et C13, le cas échéant en même temps que les effets de jours ouvrables.

Les composantes irrégulières qui serviront d’exemple pour illustrer les différents modèles sont les suivantes :

B13: SÉRIE DES FACTEURS IRRÉGULIERS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.535	100.471	97.916
1986	103.127	98.213	95.390	107.358	99.648	98.460	102.766	94.737	100.185	102.178	98.661	101.186
1987	95.874	101.153	101.588	102.296	98.892	100.097	102.280	96.086	100.774	98.541	99.781	101.664
1988	96.897	102.248	101.498	98.202	101.130	101.564	97.460	101.047	101.999	95.053	100.374	101.647
1989	101.028	99.859	98.753	99.913	99.930	102.961	97.968	100.893	98.592	99.146	102.274	98.212
1990	101.457	98.700	100.249	99.026	102.444	97.743	100.912	102.425	96.873	101.503	101.259	96.413
1991	103.060	99.489	97.167	101.999	100.597	97.834	103.528	99.549	98.111	101.645	98.541	98.994
1992	100.887	100.985	99.061	101.438	97.878	100.337	101.233	98.474	101.637	101.286	99.015	100.115
1993	96.558	101.568	101.173	99.389	98.776	101.209	100.148	101.117	100.329	97.915	100.604	101.319
1994	98.715	99.149	100.309	98.327	101.865	100.715	97.384	103.706	100.376	97.811	100.452	100.387
1995	100.292	98.592	100.227	.	.	.	.	.	.	.	.	.

D13: COMPOSANTE IRRÉGULIÈRE FINALE

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	99.954	100.970	99.281
1986	100.676	99.653	99.099	106.850	99.569	100.182	100.533	97.269	99.670	100.824	99.781	99.403
1987	93.862	100.497	100.837	99.721	100.466	99.319	100.495	99.027	100.544	99.579	100.536	100.410
1988	99.416	99.859	100.020	99.587	101.167	100.154	99.169	99.792	100.465	97.512	98.616	101.281
1989	100.192	100.244	97.009	102.498	98.779	101.571	99.973	99.260	99.980	99.593	100.711	100.402
1990	99.865	99.352	100.315	99.900	100.459	98.858	100.342	100.927	99.593	100.535	100.040	98.587
1991	100.834	100.041	99.502	100.659	99.046	100.752	102.212	100.138	99.441	99.789	100.525	99.617
1992	99.931	100.478	100.018	100.294	100.092	98.995	99.407	100.461	100.653	100.758	99.833	98.148
1993	98.999	102.515	100.691	97.966	100.338	100.045	100.112	100.420	99.512	99.565	99.603	100.134
1994	100.955	100.045	99.280	99.774	101.018	100.102	100.356	101.840	100.022	99.235	99.609	101.212
1995	100.153	99.499	100.038	.	.	.	.	.	.	.	.	.

5.2 Les modèles utilisés dans X11-ARIMA88

Comme nous l’avons vu, il est plus fréquent que Pâques tombe en avril et, dans les modèles qui vont suivre, c’est cette situation qui est considérée comme « normale » : si un « effet Pâques » est détecté, seules les données des années où Pâques tombe en mars seront corrigées. Cette correction affectera les mois de mars et avril des années concernées et, comme il est souhaitable que le niveau de la série ne soit pas modifié, on va faire en sorte que la somme des coefficients correcteurs soit égale à 0 pour un schéma additif, ou à 2 pour un schéma multiplicatif.
Enfin, notons qu’en travaillant sur les estimations de la composante irrégulière du tableau D13, c’est un effet de Pâques *résiduel* qui est estimé.

5.2.1 Le modèle à effet ponctuel

Le modèle à effet ponctuel est la version la plus simple des modèles actuellement disponibles dans le logiciel et qui seront explicités par la suite.

5.2.1.1 Modèle et estimation des effets

- Soit $I_{i,j}$ la valeur de la composante irrégulière du tableau D13 correspondant au mois *j* de l’année *i*.
- $I_{i,4}$ et $I_{i,3}$, $i = 1....n$, sont ainsi les valeurs de la composante irrégulière des mois d’avril et de mars, pour les *n* années disponibles. On note $Y_i = I_{i,4} - I_{i,3}$ leurs différences.
- Soit X_i la variable indicatrice valant 1 si Pâques tombe en Mars (ou au premier trimestre) et 0 s’il est en avril (ou au second trimestre).

- Soient enfin Z_i le nombre de jours qui séparent le dimanche de Pâques de l'année i et le 22 mars (date la plus précoce pour cette fête) et $f(Z_i)$ la fonction définie par :

  $$f(Z_i) = \begin{cases} 1 & \text{si } Z_i \leq 9 \quad \text{(Pâques tombe en mars)} \\ 0 & \text{si } Z_i > 9 \quad \text{(Pâques tombe en avril)} \end{cases}$$

  Bien que n'intervenant pas explicitement dans les calculs relatifs au modèle à effet ponctuel, la variable Z_i et la fonction $f(Z_i)$ permettent d'uniformiser la présentation des différents modèles et ainsi d'en faciliter leur compréhension. C'est pourquoi nous les introduisons ici où d'ailleurs on a $X_i = f(Z_i)$.

L'effet de Pâques peut s'obtenir en expliquant les valeurs de l'irrégulier par la variable X_i, c'est à dire en mesurant l'impact sur l'irrégulier du fait que Pâques tombe en mars. On a donc les modèles suivants :

$$\begin{cases} I_{i,3} = a_M + b_M X_i + \eta_i \\ I_{i,4} = a_A + b_A X_i + \xi_i \end{cases}$$

où b_M et b_A mesurent l'impact de Pâques sur les valeurs des mois de mars et avril et où η_i et ξ_i sont les termes d'erreurs des régressions. Dans ces modèles additifs, l'hypothèse de conservation du niveau de la série impose que $b_A = -b_M = b/2$ et on a, en prenant la différence de ces modèles :

$$I_{i,4} - I_{i,3} = Y_i = a + bX_i + \varepsilon_i.$$

- *a* s'interprète comme un effet moyen de la fête de Pâques sur la différence des irréguliers. Il traduit donc une différence structurelle entre les mois de mars et d'avril (ou les premier et second trimestres). Les données du tableau D13 ne présentant en principe ni tendance ni saisonnalité, le paramètre *a* doit être nul en théorie et son estimation proche de zéro.
- *b* est l'effet supplémentaire dû au fait que Pâques tombe en mars. C'est donc bien cette quantité qui mesure l'effet propre à la fête de Pâques.

Soit $\hat{b}$ l'estimateur du paramètre b ; alors, sous l'hypothèse de conservation du niveau de la série, les coefficients correcteurs appliqués sont :

	Schéma additif	Schéma multiplicatif
Mois de mars	$-\hat{b} * f(Z_i)/2$	$1 - \hat{b} * f(Z_i)/2$
Mois d'avril	$+\hat{b} * f(Z_i)/2$	$1 + \hat{b} * f(Z_i)/2$

Seules les données des années où Pâques tombe en mars, et pour lesquelles $f(Z_i)$ ne sera pas nul, seront donc corrigées.

L'estimateur $\hat{b}$ de b, obtenu par les moindres carrés ordinaires, peut se calculer explicitement.

- Soient n_M et n_A les nombres d'années où Pâques tombe en mars ou en avril. On a bien sur $n = n_M + n_A$.
- Soient YM et YA les sommes des valeurs Y_i pour les années où Pâques tombe respectivement en mars et en avril.

On a, pour une régression linéaire simple par les moindres carrés ordinaires, $\hat{b} = \frac{Cov(X,Y)}{Var(X)}$ or ici :

$$Cov(X,Y) = \frac{1}{n}\sum_i X_i Y_i - \overline{X}\overline{Y} = \frac{1}{n}YM - \left(\frac{n_M}{n}\right)\left(\frac{YM+YA}{n}\right) = \frac{1}{n}\left(\frac{(n-n_M)*YM}{n} - \frac{n_M * YA}{n}\right)$$
$$= \frac{1}{n}\left(\frac{n_A * YM}{n} - \frac{n_M * YA}{n}\right)$$

$$Var(X) = \frac{1}{n}\sum_i X_i^2 - \left(\overline{X}\right)^2 = \frac{n_M}{n}\frac{n_A}{n}$$

On en déduit directement $\hat{b} = \frac{YM}{n_M} - \frac{YA}{n_A} = \overline{Y}_M - \overline{Y}_A$.

Et, de même, l'estimateur $\hat{a}$ de a sera donné par $\hat{a} = \overline{Y} - \hat{b}\overline{X} = \frac{YM+YA}{n} - \left(\frac{YM}{n_M} - \frac{YA}{n_A}\right)\frac{n_M}{n} = \frac{YA}{n_A} = \overline{Y}_A$, ce qui montre que la valeur de l'estimateur devrait être proche zéro. En effet, le cas où Pâques tombe en avril correspondant à la situation « normale », l'effet de Pâques est pour ces années compris dans la saisonnalité et n'affecte donc pas les valeurs de l'irrégulier des mois de mars et avril pour ces années. Celles ci sont donc en théorie de moyenne égale à 0 ou à 1 selon le schéma de composition et leur différence est ainsi de moyenne nulle. En conséquence, on a aussi $\hat{b} = \overline{Y}_M - \overline{Y}_A \cong \overline{Y}_M$.

Enfin, il est possible de tester la présence d'un effet de Pâques en utilisant le test global de Fisher de la régression.

5.2.1.2 Exemple

Seules les données de 1986 à 1994 permettent de calculer la différence des irréguliers de mars et avril. On a donc :

Année i	$f(Z_i) = X_i$	Z_i	$I_{i,3}$	$I_{i,4}$	Y_i
1986	1	8	0.99099	1.06850	0.07751
1987	0	28	1.00837	0.99721	-0.01116
1988	0	12	1.00020	0.99587	-0.00433
1989	1	4	0.97009	1.02498	0.05489
1990	0	24	1.00315	0.99900	-0.00415
1991	1	9	0.99502	1.00659	0.01156
1992	0	28	1.00018	1.00294	0.00276
1993	0	20	1.00691	0.97966	-0.02726
1994	0	12	0.99280	0.99774	0.00494

On en déduit facilement

$$\hat{b} = \frac{0.07751+0.05489+0.01156}{3} - \frac{-0.01116-0.00433-0.00415+0.00276-0.02726+0.00494}{6}$$
$$= 0.04799 + 0.00653 = 0.05452$$

Les résultats de l'analyse de la variance et le test de Fisher associés à cette régression sont

	SOMME des CARRÉS	D.D.L	MOYENNE DES CARRÉS	F-VALUE	PROB > F
EFFET PÂQUES	0.0059	1	0.0059	14.2262	0.0070
ERREUR	0.0029	7	0.0004		
TOTAL	0.0089				

L'effet de Pâques est donc jugé significatif au niveau de 1%.

Les coefficients de correction seront nuls pour les années où Pâques tombe en avril et, le modèle étant multiplicatif, ils seront, pour les années où il tombe en mars, de :

- $1+\frac{0.05452}{2}=1.02726$ pour les données du mois d'avril
- $1-\frac{0.05452}{2}=0.97274$ pour les données du mois de mars

Ce qui conduit dans X11-ARIMA88 au tableau suivant :

A11: EFFET DE PÂQUES, MODÈLE À EFFET PONCTUEL DE X11-ARIMA88

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.000	100.000	100.000
1986	100.000	100.000	97.274	102.726	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1987	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1989	100.000	100.000	97.274	102.726	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1991	100.000	100.000	97.274	102.726	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1993	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1994	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1995	100.000	100.000	100.000	.	.	.	.	.	.	.	.	.

5.2.2 Le modèle à effet ponctuel corrigé

X11-ARIMA88 incorpore un modèle à effet ponctuel, mis au point au Bureau Australien de Statistique (Laker, 1976), légèrement différent du modèle précédent et qui prend en compte les cas où le week-end pascal est à cheval sur mars et avril, c'est à dire lorsque le dimanche de Pâques tombe le 31 mars, le 1 ou le 2 avril. Dans ce cas, on suppose que les deux mois sont affectés par la fête de Pâques et on considère que cet effet est alors la moitié de l'effet observé lorsque le week-end pascal tombe entièrement en mars.

5.2.2.1 Modèle

Le modèle s'écrit alors $Y_i = a + bX_i + \varepsilon_i$ avec :

- $I_{i,j}$ la valeur de la composante irrégulière du tableau D13 correspondant au mois *j* de l'année *i*. $I_{i,4}$ et $I_{i,3}$, les valeurs de la composante irrégulière des mois d'avril et de mars, pour les *n* années disponibles et $Y_i = I_{i,4} - I_{i,3}$ leurs différences.
- X_i la variable indicatrice valant 1 si Pâques tombe en Mars, 0.5 si Pâques tombe à cheval sur mars et avril, et 0 s'il est en avril.

- Z_i le nombre de jours qui séparent le dimanche de Pâques de l'année *i* et le 22 mars (date la plus précoce pour cette fête) et $f(Z_i)$ la fonction définie par :

$$f(Z_i) = \begin{cases} 1 & \text{si } Z_i \le 8 \quad \text{(Pâques tombe en mars)} \\ 0.5 & \text{si } 9 \le Z_i \le 11 \quad \text{(Pâques tombe en mars - avril)} \\ 0 & \text{si } Z_i \ge 12 \quad \text{(Pâques tombe en avril)} \end{cases}$$

La valeur de $\hat{b}$ peut encore se calculer explicitement.

- Soient n_M, n_{MA} et n_A les nombres d'années où Pâques tombe en mars, en « mars-avril » et en avril. On a bien sur $n = n_M + n_{MA} + n_A$.
- Soient YM, YMA et YA les sommes des valeurs Y_i pour les années où Pâques tombe respectivement en mars, en « mars-avril » et en avril ; et soit $YT = YM + YMA + YA$

On a alors :

$$Cov(X,Y) = \frac{1}{n}\sum_i X_i Y_i - \overline{X}\overline{Y} = \frac{1}{n}(YM + 0.5 * YMA) - \left(\frac{n_M + 0.5 * n_{MA}}{n}\right)\left(\frac{YT}{n}\right)$$

$$= \frac{1}{n}(YM + 0.5 * YMA - r * YT)$$

avec $r = \dfrac{n_M + 0.5 * n_{MA}}{n}$

$$Var(X) = \frac{1}{n}\sum_i X_i^2 - \left(\overline{X}\right)^2 = \frac{1}{n}(n_M + 0.25 * n_{MA}) - r^2 = r - 0.25 * \frac{n_{MA}}{n} - r^2$$

$$= r * (1 - r) - 0.25 * \frac{n_{MA}}{n}$$

et donc $\hat{b} = 2\dfrac{YM + 0.5 * YMA - r * YT}{2 * n * r * (1 - r) - 0.5 * n_{MA}}$

Ici encore, et comme le fait le programme X11-ARIMA88, la présence d'un effet de Pâques peut être testée par le test de Fisher global de la régression.

5.2.2.2 Estimation des effets

L'estimation « naturelle » des corrections est la même que précédemment, et est appliquée telle quelle pour un schéma additif :

	Schéma additif
Mois de mars	$-\hat{b} * f(Z_i)/2$
Mois d'avril	$+\hat{b} * f(Z_i)/2$

Ainsi, pour une année où Pâques tombe à cheval sur mars et avril, la correction sera moitié moindre que celle apportée aux années où Pâques tombe en mars.

Dans le cas d'un schéma de composition multiplicatif, le programme impose une correction supplémentaire.
Pour les années où Pâques tombe en avril, et uniquement pour elles, notons AIA la moyenne des irréguliers d'avril $I_{i,4}$ et AIM la moyenne des irréguliers de mars $I_{i,3}$. Alors les effets correcteurs appliqués sont définis par :

	Schéma multiplicatif
Mois de mars	$1-\hat{b}*f(Z_i)/(2*AIM)$
Mois d'avril	$1+\hat{b}*f(Z_i)/(2*AIA)$

5.2.2.3 Justification de la correction apportée pour un schéma multiplicatif

Dans la présentation théorique du modèle (Laker, 1976), l'auteur considère l'estimation de l'effet de Pâques à partir d'une estimation de la composante Saisonnier-Irrégulier. Il pose donc, de façon similaire à ce que nous avons vu jusqu'à présent :
$\begin{cases} SI_{i,3} = \overline{SI}_{.3} - bX_i + \eta_i \\ SI_{i,4} = \overline{SI}_{.4} + bX_i + \xi_i \end{cases}$ soit, par différence, $\left(SI_{i,4} - SI_{i,3}\right) = \left(\overline{SI}_{.4} - \overline{SI}_{.3}\right) + 2bX_i + \varepsilon_i$

Où :

- $SI_{i,3}$ et $SI_{i,4}$ désignent les valeurs de la composante Saisonnier-Irrégulier pour les mois de mars et avril,
- b est l'effet de Pâques,
- $\overline{SI}_{.,3}$ est la moyenne théorique de la composante Saisonnier-Irrégulier pour un mois de mars « normal », c'est à dire non affecté d'un effet de Pâques,
- et $\overline{SI}_{.,4}$ est la moyenne théorique de la composante Saisonnier-Irrégulier pour un mois d'avril « normal », c'est à dire affecté d'un effet de Pâques.

Dans ces conditions, si $2\hat{b}$ est l'estimation issue du modèle aux différences, on aura dans le cas d'un schéma multiplicatif et par exemple pour le mois de mars :
$SI_{i,3} \approx \overline{SI}_{.3} - \hat{b}X_i = \overline{SI}_{.3}\left[1-\frac{\hat{b}X_i}{\overline{SI}_{.3}}\right]$ soit $\overline{SI}_{.3} \approx SI_{i,3} \Big/ \left(1-\frac{\hat{b}X_i}{\overline{SI}_{.3}}\right)$ et donc le ratio $\left(1-\frac{\hat{b}X_i}{\overline{SI}_{.3}}\right)$ permet de passer d'une valeur de la composante Saisonnier-Irrégulier affectée par l'effet de Pâques à une valeur corrigée de cet effet. Malheureusement, la quantité $\overline{SI}_{.3}$ est inconnue et Laker propose de l'estimer par la moyenne des valeurs de la composante Saisonnier-Irrégulier des mois de mars des années où Pâques tombe en avril (donc des mois de mars « normaux »).
Une correction de même philosophie est proposée pour les valeurs de la composante Saisonnier-Irrégulier du mois d'avril.

Dans notre cas, l'estimation de l'effet de Pâques se faisant sur les estimations de la composante irrégulière, Laker propose d'adopter le même principe de correction, ce qui conduit aux formules du paragraphe précédent.

En fait, cette correction s'avère ici superflue dans la mesure où nous connaissons la moyenne théorique de l'irrégulier, ici égale à 1, et qu'il est inutile de l'estimer.

5.2.2.4 Exemple

Les données sont les mêmes et seule la fonction $F(Z_i)$ change pour l'année 1991 où le dimanche de Pâques tombant le 31 mars, le week-end pascal est à cheval sur mars et avril.

Année i	$f(Z_i)$	Z_i	$I_{i,3}$	$I_{i,4}$	Y_i
1986	1	8	0.99099	1.06850	0.07751
1987	0	28	1.00837	0.99721	-0.01116
1988	0	12	1.00020	0.99587	-0.00433
1989	1	4	0.97009	1.02498	0.05489
1990	0	24	1.00315	0.99900	-0.00415
1991	0.5	9	0.99502	1.00659	0.01156
1992	0	28	1.00018	1.00294	0.00276
1993	0	20	1.00691	0.97966	-0.02726
1994	0	12	0.99280	0.99774	0.00494

On a donc :

$$r = \frac{n_M + 0.5 * n_{MA}}{n} = \frac{2 + 0.5 * 1}{9} = 0.27778$$

$$\hat{b} = 2\frac{YM + 0.5 * YMA - r * YT}{2 * n * r * (1-r) - 0.5 * n_{MA}} = 2 * \frac{0.13240 + 0.5 * 0.01156 - 0.27778 * 0.104762}{2 * 9 * 0.27778 * 0.72222 \text{ - } 0.5 * 1} = 0.07012$$

Les résultats de l'analyse de la variance et le test de Fisher associés à cette régression sont

	SOMME des CARRÉS	D.D.L	MOYENNE DES CARRÉS	F-VALUE	PROB > F
EFFET PÂQUES	0.0076	1	0.0076	43.8393	0.0003
ERREUR	0.0012	7	0.0002		
TOTAL	0.0089				

L'effet de Pâques est donc jugé très significatif.
Comme notre schéma est multiplicatif, le calcul des coefficients correcteurs nécessite de connaître les moyennes des irréguliers pour les années où Pâques tombe en avril. On a :

$$AIM = \frac{1.00837 + 1.00020 + 1.00315 + 1.00018 + 1.00691 + 0.99280}{6} = 1.001935$$

$$AIA = \frac{0.99721 + 0.99587 + 0.99900 + 1.00294 + 0.97966 + 0.99774}{6} = 0.995403$$

Et les effets correcteurs appliqués sont définis par :

	Pâques en mars	Pâques en mars-avril
Mois de mars	$1 - \hat{b}/(2 * AIM) = 0.9650$	$1 - 0.5 * \hat{b}/(2 * AIM) = 0.9825$
Mois d'avril	$1 + \hat{b}/(2 * AIA) = 1.0352$	$1 - 0.5 * \hat{b}/(2 * AIA) = 1.0176$

Ce qui conduit au tableau suivant de X11-ARIMA88 :

A11: EFFET DE PÂQUES, MODÈLE À EFFET PONCTUEL CORRIGÉ DE X11-ARIMA88

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.000	100.000	100.000
1986	100.000	100.000	96.501	103.522	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1987	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1989	100.000	100.000	96.501	103.522	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1991	100.000	100.000	98.250	101.761	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1993	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1994	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1995	100.000	100.000	100.000	.	.	.	.	.	.	.	.	.

5.2.3 Le modèle à effet graduel

La fête de Pâques peut aussi avoir un effet sur les jours qui précèdent sa célébration : ainsi comme il est habituel d'offrir et de consommer du chocolat ou d'offrir des fleurs à cette période, les industries concernées vont adapter leur production pour pouvoir, le moment venu, répondre à la demande.

5.2.3.1 Modèle et estimation

Dans ce cas, on suppose que l'effet varie linéairement pendant les *k* jours (*k* pouvant prendre les valeurs 1 à 9) qui précèdent le dimanche de Pâques. Soient alors :

- $I_{i,3}$ et $I_{i,4}$, $i = 1....n$, les valeurs de la composante irrégulière des mois de mars et avril du tableau D13 pour les *n* années disponibles et $Y_i = I_{i,4} - I_{i,3}$ leurs différences.
- X_i la variable indicatrice valant 1 si Pâques tombe en Mars et 0 s'il tombe le ou après le *k* avril.
- Soient enfin Z_i le nombre de jours qui séparent le dimanche de Pâques de l'année *i* et le 22 mars (date la plus précoce pour cette fête) et $f(Z_i)$ la fonction définie par :

$$f(Z_i) = \begin{cases} 1 & \text{si } Z_i \leq 9 & \text{(Pâques tombe en mars)} \\ \dfrac{k+9-Z_i}{k} & \text{si } 9 < Z_i < k+9 & \text{(Pâques tombe en avril, avant le } k \text{ avril)} \\ 0 & \text{si } Z_i \geq 9+k & \text{(Pâques tombe en avril, le ou après le } k \text{ avril)} \end{cases}$$

L'effet de Pâques est alors estimé, par les moindres carrés ordinaires, à partir du modèle $Y_i = a + bX_i + \varepsilon_i$; dans ce modèle les données des années où Pâques tombe entre le 1 et le *k* avril ne sont donc pas utilisées.

La valeur de *b* peut encore se calculer explicitement, en utilisant les résultats du modèle de base à effet ponctuel estimé sur les années pour lesquelles $f(Z_i) = 0$ ou bien $f(Z_i) = 1$. En notant :

- n_M et n_{LA} les nombres d'années où Pâques tombe en mars, et « fin avril »
- YM et YLA les sommes des valeurs Y_i pour les années où Pâques tombe respectivement en mars et « fin avril »,

on a alors immédiatement $\hat{b} = \frac{YM}{n_M} - \frac{YLA}{n_{LA}} = \overline{Y}_M - \overline{Y}_{LA}$ et $\hat{a} = \overline{Y}_{LA}$.

Les coefficients correcteurs appliqués sont alors les suivants :

	Schéma additif	Schéma multiplicatif
Mois de mars	$-\hat{b} * f(Z_i)/2$	$1 - \hat{b} * f(Z_i)/2$
Mois d'avril	$+\hat{b} * f(Z_i)/2$	$1 + \hat{b} * f(Z_i)/2$

5.2.3.2 Remarques

- L'utilisateur peut demander au programme X11-ARIMA88 de choisir automatiquement la valeur « optimale » de *k* parmi les valeurs possibles de 1 à 9. L'algorithme choisira alors la valeur de *k* qui conduit à la plus petite erreur quadratique moyenne.
- L'utilisateur peut demander au programme d'exclure les valeurs aberrantes avant d'estimer le modèle. Seules les valeurs correspondant aux années où Pâques est en avril sont concernées : il y a en général trop peu d'années où Pâques tombe en mars pour identifier correctement les valeurs extrêmes dans ce cas.

  1- On calcule l'écart-type des valeurs Y_i pour les années où Pâques tombe le ou après le *k* avril :

  $$\sigma_{LA} = \sqrt{\frac{1}{n_{LA}-1}\sum_{i\in LA}\left(Y_i - \overline{Y}_{LA}\right)^2} = \sqrt{\frac{1}{n_{LA}-1}\left[\sum_{i\in LA} Y_i^{\,2} - n_{LA}\overline{Y}_{LA}^{\,2}\right]}$$

  2- Et on exclut toutes les valeurs de ces années telles que : $\left|Y_i - \overline{Y}_{LA}\right| > 2 * \sigma_{LA}$
- Ce modèle est très proche du modèle « *sceaster* » utilisé dans X12-ARIMA (voir 5.3.2) à cela près que l'estimation est faite de façon différente.

5.2.3.3 Exemple

A titre d'exemple, nous allons supposer *k*=5. Les données sont :

Année i	Date de Pâques	$f(Z_i)$	Z_i	IM_i	IA_i	Y_i
1986	30 mars	1	8	0.99099	1.06850	0.07751
1987	19 avril	0	28	1.00837	0.99721	-0.01116
1988	3 avril	0.4	12	1.00020	0.99587	-0.00433
1989	26 mars	1	4	0.97009	1.02498	0.05489
1990	15 avril	0	24	1.00315	0.99900	-0.00415
1991	31 mars	1	9	0.99502	1.00659	0.01156
1992	19 avril	0	28	1.00018	1.00294	0.00276
1993	11 avril	0	20	1.00691	0.97966	-0.02726
1994	3 avril	0.4	12	0.99280	0.99774	0.00494

Les années où Pâques tombe le 1, 2, 3 ou 4 avril sont donc affectées d'un poids différent de 0 où 1. C'est le cas des années 1988 et 1994 où Pâques était le 3 avril ($Z_i = 12$). La valeur de $f(Z_i)$ est pour ces années :

$$f(Z_i) = \frac{k+9-Z_i}{k} = \frac{5+9-12}{5} = \frac{2}{5} = 0.4$$

La régression est faite sur les données pour lesquelles $f(Z_i) = 1$ (Pâques en mars) ou $f(Z_i) = 0$ (Pâques après le 4 avril), soit 7 années. On a :

$$b = \frac{0.07751 + 0.05489 + 0.01156}{3} - \frac{-0.01116 - 0.00415 + 0.00276 - 0.02726}{4} = 0.05794$$

Les résultats de l'analyse de la variance et le test de Fisher associés à cette régression sont

	SOMME des CARRÉS	D.D.L	MOYENNE DES CARRÉS	F-VALUE	PROB > F
EFFET PÂQUES	0.0058	1	0.0058	10.4940	0.0230
ERREUR	0.0027	5	0.0005		
TOTAL	0.0085				

L'effet de Pâques est donc jugé significatif.

Notre schéma étant multiplicatif, les coefficients correcteurs seront :

	Pâques en mars (86, 89, 91)	Pâques le 3 avril (88, 94)
Mois de mars	$1 - b/2 = 0.9710$	$1 - 0.4 * b/2 = 0.9884$
Mois d'avril	$1 + b/2 = 1.0289$	$1 + 0.4 * b/2 = 1.0116$

Pour une valeur de *k* égale à 5, on obtient donc le tableau suivant :

A11: EFFET DE PÂQUES, MODÈLE À EFFET GRADUEL DE X11-ARIMA88

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.000	100.000	100.000
1986	100.000	100.000	97.103	102.897	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1987	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	98.841	101.159	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1989	100.000	100.000	97.103	102.897	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1991	100.000	100.000	97.103	102.897	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1993	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1994	100.000	100.000	98.841	101.159	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1995	100.000	100.000	100.000	.	.	.	.	.	.	.	.	.

Les valeurs estimées des différences $\hat{Y}_i$ sont alors données par $\hat{Y}_i = \hat{a} + \hat{b} * f(Z_i)$.

Or $\hat{a}$ est égal à la moyenne des différences pour les années où Pâques tombe après le 4 avril et donc

$$\hat{a} = \frac{-0.011159 - 0.004152 + 0.0027566 - 0.027256}{4} = -0.009953$$

On peut alors évaluer les erreurs de prévision pour les années où Pâques tombe en avril :

Année i	Y_i	$f(Z_i)$	$\hat{b} * f(Z_i)$	$\hat{Y}_i = \hat{a} + \hat{b} * f(Z_i)$	$Y_i - \hat{Y}_i$
1987	-0.01116	0	0	-0.00995	-0.001206
1988	-0.00433	0.4	0.02318	0.01322	-0.017553
1990	-0.00415	0	0	-0.00995	0.005800
1992	0.00276	0	0	-0.00995	0.012709
1993	-0.02726	0	0	-0.00995	-0.017303
1994	0.00494	0.4	0.02318	0.01322	-0.008284

La moyenne des carrés de ces erreurs est alors égale à 0.0001455.

Choix optimal de la durée *k*

Si on demandait au programme de choisir la valeur de *k* optimale, c'est cette erreur quadratique qui servirait de critère et il prendrait la valeur de *k* minimisant cette quantité. Sur notre exemple, on trouverait les valeurs suivantes :

k	1	2	3	4	5	6	7	8	9
Erreur	0.0001132	0.0001132	0.0001132	0.0000958	0.0001455	0.0002065	0.0002639	0.0003144	0.0003581

Pour *k* égal à 1, 2 ou 3, on peut remarquer que pour les années qui nous intéressent (1986 à 1994), les fonctions $f(Z_i)$ seront égales et les résultats de la régression identiques, ce qui explique l'égalité des 3 premières erreurs quadratiques.

Prise en compte des valeurs extrêmes

Comme on l'a vu précédemment, $\hat{a} = \overline{Y}_{LA} = -0.009953$ et l'écart-type des valeurs pour les années où Pâques tombe le où après le 5 avril se calcule par exemple par $\sigma_{LA} = \sqrt{\frac{1}{n_{LA}-1}\left[\sum_{i\in LA} Y_i^2 - n_{LA}\overline{Y}_{LA}^2\right]}$, or :

Y_i	$f(Z_i)$	Y_i^2	$\lvert Y_i - \overline{Y}_{LA}\rvert$
-0.011159	0	0.0001245	0.0012063
-0.004152	0	0.0000172	0.0058004
0.002757	0	0.0000076	0.0127093
-0.027256	0	0.0007429	0.0173034

Et donc

$$\sigma_{LA} = \sqrt{\frac{1}{4-1}\left[(0.0001245 + 0.0000172 + 0.0000076 + 0.0007429) - 4*(-0.009953)^2\right]} \approx 0.01286.$$

Aucune valeur Y_i ne s'éloigne donc, en valeur absolue, de plus de 2 écart-types (soit 0.0257) de la moyenne et aucun point n'est donc exclu de la régression.

5.3 Les modèles de X12-ARIMA

X12-ARIMA propose des modèles de correction de l'effet de Pâques différents de ceux de X11-ARIMA88. Les modèles *Sceaster* et *Easter* évaluent l'effet de Pâques à partir des estimations de la composante irrégulière des tables B13 et C13, seul le modèle de Bateman et Mayes (1970) utilise les données du tableau D13.
Le modèle *Sceaster* est très proche du modèle à effet graduel proposé par X11-ARIMA88 et discuté précédemment (voir 5.2.3). Ce modèle et ceux de X11-ARIMA88 considèrent que la situation normale est celle où Pâques tombe en avril et ils ne corrigent donc que les données des années où Pâques tombe en mars. Ce n'est pas le cas du modèle *Easter* et du modèle de Bateman et Mayes qui vont, comme on le verra dans les paragraphes suivants, corriger les données des mois de mars, avril et éventuellement février.

5.3.1 Le modèle de Bateman & Mayes

Ce modèle n'est utilisé que dans le cas d'un schéma de composition multiplicatif.

5.3.1.1 Modèle et estimation

L'effet de Pâques est estimé, à partir des valeurs de la composante irrégulière du tableau D13 pour les mois de mars et avril, en plusieurs étapes :

- Soient $I_{i,3}$ et $I_{i,4}$, $i = 1....n$, les valeurs de la composante irrégulière du tableau D13 pour les n années disponibles. Les valeurs des mois d'avril sont transformées et ce sont les variables $2 - I_{i,4}$ et $I_{i,3}$ qui seront utilisées dans la suite.
- Soit k_i le nombre de jours qui séparent le dimanche de Pâques de l'année *i* et le 22 mars.
- Les années sont réparties en 4 groupes définis en fonction de k :
  $G_1 : 1 \leq k \leq 10$ (du 23 mars au 1 avril)
  $G_2 : 11 \leq k \leq 17$ (du 2 avril au 8 avril)
  $G_3 : 18 \leq k \leq 24$ (du 9 avril au 15 avril)
  $G_4 : 25 \leq k \leq 34$ (du 16 avril au 25 avril)
  Dans chaque groupe, on va calculer maintenant des moyennes « tronquées » de l'ensemble des valeurs $2 - I_{i,4}$ et $I_{i,3}$,
- La moyenne tronquée m_1 est la moyenne des valeurs du groupe G_1 calculée en éliminant les valeurs qui s'éloignent de plus de deux écart-types de la moyenne simple du groupe G_1. La moyenne tronquée m_4 pour le groupe G_4 est calculée de la même façon.
- Le calcul des moyennes m_2 et m_3 comporte plusieurs étapes :
  a) Soient $\overline{m}_2$ et $\overline{m}_3$ les moyennes simples des groupes G_2 et G_3.
  b) Des facteurs préliminaires de correction de l'effet de Pâques pour le mois de mars, correspondant à l'observation *t*, $E_t(k)$ sont calculés ainsi :

$$E_t(k) = \begin{cases} m_1 + \dfrac{(\overline{m}_2 - m_1)(k-10)}{4} & \text{si } 11 \leq \text{k} \leq 14 \\ \overline{m}_2 + \dfrac{(\overline{m}_3 - \overline{m}_2)(k-14)}{7} & \text{si } 15 \leq \text{k} \leq 20 \\ \overline{m}_3 + \dfrac{(m_4 - \overline{m}_3)(k-21)}{4} & \text{si } 21 \leq \text{k} \leq 24 \end{cases}$$

  c) Le facteur préliminaire de correction pour le mois d'avril *t+1* s'en déduit par $E_{t+1}(k) = 2 - E_t(k)$.
  d) Calculer les erreurs quadratiques moyennes des coefficients correcteurs préliminaires pour $11 \leq k \leq 17$ et $18 \leq k \leq 24$.
  e) Les moyennes tronquées m_2 et m_3 sont les moyennes des valeurs des groupes G_2 et G_3 qui ne s'éloignent pas de plus de deux écart-types des valeurs des facteurs préliminaires qui leur sont affectées.

- L’effet de Pâques pour un mois de mars est alors calculé ainsi :

$$E_t(k)=\begin{cases} m_1 & \text{si } 1\le \text{k} \le 10 \\ m_1+\dfrac{(m_2-m_1)(k-10)}{4} & \text{si } 11\le \text{k} \le 14 \\ m_2+\dfrac{(m_3-m_2)(k-14)}{7} & \text{si } 15\le \text{k} \le 20 \\ m_3+\dfrac{(m_4-m_3)(k-21)}{4} & \text{si } 21\le \text{k} \le 24 \\ m_4 & \text{si } 25\le \text{k} \le 34 \end{cases}$$

  Pour le mois d’avril *t+1*, on aura : $E_{t+1}(k)=2-E_t(k)$ et, pour tous les autres mois, le coefficient correcteur sera égal à 1.
  La fonction qui à *k* associe l’effet de Pâques pour un mois de mars est donc une fonction linéaire par morceaux.
- Ces coefficients correcteurs doivent être alors ajustés pour tenir compte de la saisonnalité de Pâques dans la mesure où certaines dates sont plus fréquentes que d’autres (voir la figure 20). Pour cela, on calcule la quantité : $\overline{E}=\sum_{k=0}^{34} w(k)E_t(k)$ où $w(k)$ est la proportion d’années où Pâques tombe le $k-ième$ jour après le 22 mars, proportion calculée sur un calendrier de 38 000 années.
  Les valeurs de $w(k)$ utilisées par X12-ARIMA sont les suivantes :

k	0	1	2	3	4	5	6
w(k)	0.0290	0.0285	0.0285	0.0290	0.0280	0.0290	0.0280

k	7	8	9	10	11	12	13
w(k)	0.0290	0.0285	0.0285	0.0290	0.0280	0.0290	0.0280

k	14	15	16	17	18	19	20
w(k)	0.0290	0.0285	0.0285	0.0290	0.0280	0.0290	0.0280

k	21	22	23	24	25	26	27
w(k)	0.0290	0.0285	0.0285	0.0290	0.0280	0.0290	0.0280

k	28	29	30	31	32	33	34
w(k)	0.0290	0.0285	0.0285	0.0290	0.0280	0.0290	0.0280

Les estimations finales des facteurs correctifs de l’effet de Pâques sont alors :

$$\tilde{E}_t=\begin{cases} E_t/\overline{E} & \text{si } t \text{ est un mois de mars} \\ E_t/(2-\overline{E}) & \text{si } t \text{ est un mois d'avril} \\ 1 & \text{sinon} \end{cases}$$

5.3.1.2 Exemple

Les données brutes (tableau D13) et les données transformées sont les suivantes :

Données brutes

ANNÉE	MARS	AVRIL
1986	99.099	106.850
1987	100.837	99.721
1988	100.020	99.587
1989	97.009	102.498
1990	100.315	99.900
1991	99.502	100.659
1992	100.018	100.294
1993	100.691	97.966
1994	99.280	99.774
1995	100.038	.

Données transformées

k	MARS	2-AVRIL
8	0.99099	0.93150
28	1.00837	1.00279
12	1.00020	1.00413
4	0.97009	0.97502
24	1.00315	1.00100
9	0.99502	0.99341
28	1.00018	0.99706
20	1.00691	1.02034
12	0.99280	1.00226
25	1.00038	.

Ainsi par exemple, la donnée transformée pour avril 1987 est $APR87 = 2 - 0.99721 = 1.00279$.
Des valeurs de *k*, on déduit les sous-ensembles G_1, G_2, G_3 et G_4 :

$G_1 = \{1986, 1989, 1991\}$

$G_2 = \{1988, 1994\}$

$G_3 = \{1990, 1993\}$

$G_4 = \{1987, 1992, 1995\}$

- **Calcul de m_1 :**

Les données de G_1 sont :

k	MARS	2-AVRIL
8	0.99099	0.93150
4	0.97009	0.97502
9	0.99502	0.99341

La moyenne $\overline{m}_1$ et l'écart-type σ_1 de ces six valeurs sont : $\overline{m}_1 = 0.97601$ et $\sigma_1 = 0.02201$.
La matrice des différences absolues à $\overline{m}_1$ est :

MARS	2-AVRIL
0.01498	0.04451
0.00591	0.00098
0.01902	0.01741

Et seule la première valeur d'avril (0.93150) est à plus de deux écart-types de la moyenne $\overline{m}_1$ (écart de 0.04451) et doit donc être éliminée du calcul final.
On a donc :

$$m_1 = \frac{0.99099 + 0.97009 + 0.99502 + 0.97502 + 0.99341}{5} = 0.98491$$

- **Calcul de m_4 :**

Les données de G_4 sont :

k	MARS	2-AVRIL
28	1.00837	1.00279
28	1.00018	0.99706
25	1.00038	.

La moyenne $\overline{m}_4$ et l'écart-type σ_4 de ces six valeurs sont : $\overline{m}_4 = 1.00176$ et $\sigma_4 = 0.003774$.
La matrice des différences absolues à $\overline{m}_4$ est :

MARS	2-AVRIL
0.00661	0.00103
0.00158	0.00469
0.00136	.

Et ici, aucun point n'étant jugé extrême, la valeur de m_4 est 1.00176.

- **Calcul de m_2 et m_3 :**

Données de G_2

k	MARS	2-AVRIL
12	1.00020	1.00413
12	0.99280	1.00226

Données de G_3

k	MARS	2-AVRIL
24	1.00315	1.00100
20	1.00691	1.02034

Les moyennes simples de ces deux groupes de 4 valeurs sont $\overline{m}_2 = 0.99985$ et $\overline{m}_3 = 1.00785$ et les estimations préliminaires des effets pour mars et avril sont :

Données de G_2

k	MARS	AVRIL
12	0.99238	1.00762
12	0.99238	1.00762

Données de G_3

k	MARS	AVRIL
24	1.00328	0.99672
20	1.00671	0.99329

Ainsi, pour *k*=12 (données de G_2),

$$E_t(12) = m_1 + \frac{(\overline{m}_2 - m_1)(k-10)}{4} = 0.98491 + \frac{(0.99985 - 0.98491)(12-10)}{4} = 0.99238$$

$$E_{t+1}(12) = 2 - 0.99238 = 1.00762$$

Et, de même, pour les données de G_3 , on a :

$$E_t(20) = \overline{m}_2 + \frac{(\overline{m}_3 - \overline{m}_2)(k-14)}{7} = 0.99985 + \frac{(1.00785 - 0.99985)(20-14)}{7} = 1.00671$$

$$E_{t+1}(20) = 2 - 1.00328 = 0.99329$$

et

$$E_t(24) = \overline{m}_3 + \frac{(m_4 - \overline{m}_3)(k-21)}{4} = 1.00785 + \frac{(1.00176 - 1.00785)(24-21)}{4} = 1.00328$$

$$E_{t+1}(24) = 2 - 1.00328 = 0.99329$$

Les matrices des écarts absolus aux estimations préliminaires des effets sont :

Données de G_2

K	MARS	AVRIL
12	0.00782	0.00349
12	0.00043	0.00537

Données de G_3

k	MARS	AVRIL
24	0.00013	0.00428
20	0.00020	0.02705

Les erreurs quadratiques moyennes sont alors :
Pour I_2 ($11 \leq k \leq 17$) :

$$EQM_2 = \sqrt{\frac{1}{4}\left[(0.00782)^2 + (0.00349)^2 + (0.00043)^2 + (0.00537)^2\right]} = 0.00506$$

Pour I_3 ($18 \leq k \leq 24$) :

$$EQM_3 = \sqrt{\frac{1}{4}\left[(0.00013)^2 + (0.00428)^2 + (0.00020)^2 + (0.02705)^2\right]} = 0.01369$$

Aucun point n'est donc jugé atypique et les moyennes pour chaque sous-groupe sont donc les moyennes simples : $m_2 = 0.99985$ et $\mathrm{m}_3 = 1.00785$

- **Calcul des effets de Pâques**

Ces effets se calculent ainsi :

1. Pour les données de G_1, les effets de mars sont égaux à m_1 (0.98491) et ceux d'avril à $2 - m_1$
2. Pour les données de G_4, les effets de mars sont égaux à m_4 (1.00176) et ceux d'avril à $2 - m_4$
3. Pour les données de G_2 et de G_3, les effets sont égaux aux effets préliminaires calculés précédemment puisqu'aucun point extrême n'a été détecté.

Les effets de Pâques peuvent se calculer pour toute valeur de *k* à partir des moyennes m_1, m_2, m_3 et m_4 : la courbe de ces effets est présentée dans la figure 21.

On peut alors calculer la composante saisonnière de ces effets : $\overline{E} = \sum_{k=0}^{34} w(k) E_t(k) = 0.99639$

et les effets corrigés :

Effets de Pâques « bruts »

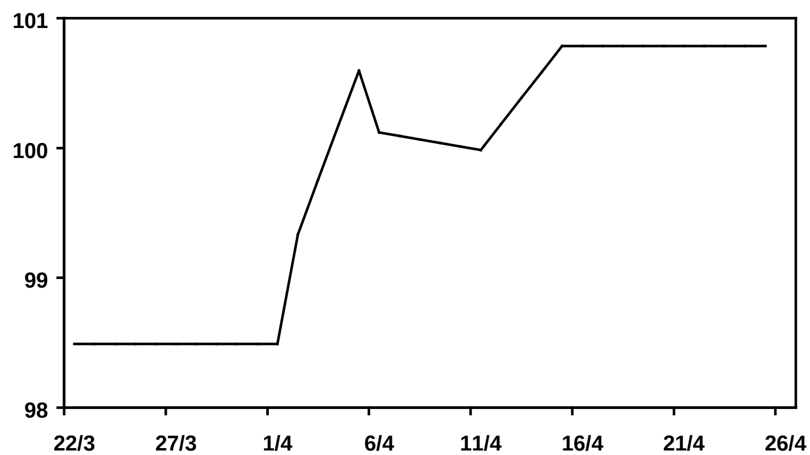
ANNÉE	MARS	AVRIL
1986	98.491	101.509
1987	100.176	99.824
1988	99.238	100.762
1989	98.491	101.509
1990	100.328	99.672
1991	98.491	101.509
1992	100.176	99.824
1993	100.671	99.329
1994	99.238	100.762
1995	100.176	99.824

Effets de Pâques « corrigés »

MARS	AVRIL
98.847	101.144
100.538	99.465
99.597	100.400
98.847	101.144
100.691	99.314
98.847	101.144
100.538	99.465
101.035	98.972
99.597	100.400
100.538	99.465

Avec par exemple :
$MAR86 = 98.491 / 0.99639 = 98.847$ et $APR86 = 101.509 / (2 - 0.99639) = 101.144$

Figure 21 : Effet de Pâques selon la date de Pâques (du 22 mars au 25 avril)



Ce qui conduit au résultat final :

CHL: EFFET DE PÂQUES, MODÈLE DE X12

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.000	100.000	100.000
1986	100.000	100.000	98.847	101.144	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1987	100.000	100.000	100.538	99.465	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	99.597	100.400	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1989	100.000	100.000	98.847	101.144	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	100.691	99.314	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1991	100.000	100.000	98.847	101.144	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	100.538	99.465	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1993	100.000	100.000	101.035	98.972	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1994	100.000	100.000	99.597	100.400	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1995	100.000	100.000	100.538	.	.	.	.	.	.	.	.	.

5.3.2 Le modèle « Sceaster »

Ce modèle entre dans le cadre plus général de l'estimation des effets de calendrier proposée par la méthode X11 à la fin des étapes B et C et plus précisément à partir des estimations de la composante irrégulière des tableaux B13 et C13.

5.3.2.1 Modèle et estimation

La fête de Pâques est supposée, comme dans le modèle à effet graduel proposé dans X11-ARIMA88, avoir un impact sur les *w* jours ($1 \leq w \leq 24$) qui précèdent le dimanche pascal et on pose le modèle $Y_{ij} = a + bX_{ij}(w) + \varepsilon_{ij}$ où :

- Y_{ij} est la valeur de l'irrégulier (du tableau B13 ou C13) correspondant à l'année *i* et à la période (mois ou trimestre) *j*.
- Soit, pour une année *i* donnée, n_i le nombre de jours, parmi les *w* jours avant Pâques (Pâques compris) qui tombent en mars (ou au premier trimestre). Alors :

$$X_{ij}(w) = \begin{cases} n_i / w & \text{pour un mois de mars ou un premier trimestre } (j = 3 \text{ ou } j = 1) \\ -n_i / w & \text{pour un mois d'avril ou un second trimestre } (j = 4 \text{ ou } j = 2) \\ 0 & \text{sinon} \end{cases}$$

Les contraintes imposées à w ($1 \leq w \leq 24$) entraînent que seules les valeurs du régresseur pour mars et avril (ou du premier et second trimestre) ne sont pas nulles[34]. Les valeurs de a et b sont estimées par moindres carrés ordinaires.
À ce stade, il est possible de faire plusieurs remarques.

- Tout d'abord, si Pâques tombe en mars, la valeur associée $X_{i3}(w)$ est égale à 1. De même, si Pâques tombe après le w avril, cette valeur $X_{i3}(w)$ est égale à 0. De façon plus générale, si on considère les valeurs de la variable X pour les mois de mars, on a, en reprenant les notations associées aux modèles de X11-ARIMA88 : $X_{i3}(w) = f(Z_i)$
- Compte tenu de la forme simple de la variable explicative, on peut donner une forme plus explicite de l'estimateur $\hat{b}$. On a, en supposant que la série comporte T observations :
  $\hat{b} = \dfrac{Cov(X,Y)}{Var(X)}$ avec $Cov(X,Y) = \dfrac{1}{T}\sum_i \sum_j X_{ij}Y_{ij} - \overline{X}\overline{Y}$ et
  $Var(X) = \dfrac{1}{T}\sum_i \sum_j X_{ij}^2 - \overline{X}^2$
  Les seules valeurs non nulles de la variable X sont celles des mois de mars et avril qui sont par ailleurs opposées. Si l'on suppose que la série ne commence pas un mois d'avril ou ne se termine pas en mars, à tout mois de mars correspondra un mois d'avril et la somme des deux valeurs de la variable X de ces mois sera nulle. La moyenne $\overline{X}$ est donc elle aussi nulle dans ce cas, et dans le cas général proche de 0.

  On aura donc :
  $$Var(X) = \frac{1}{T}\sum_i \sum_j X_{ij}^2 - \overline{X}^2 = \frac{1}{T}\sum_i \sum_j X_{ij}^2 = \frac{2}{T}\sum_i \left(\frac{n_i}{w}\right)^2$$
  $$Cov(X,Y) = \frac{1}{T}\sum_i \sum_j X_{ij}Y_{ij} - \overline{X}\overline{Y} = \frac{1}{T}\sum_i \sum_j X_{ij}Y_{ij} = \frac{1}{T}\sum_i \frac{n_i}{w}\left[Y_{i3} - Y_{i4}\right]$$
  Et on voit donc apparaître les valeurs des différences, pour chaque année, des irréguliers de mars et avril.
- Enfin, comme $\hat{a} = \overline{Y} - \hat{b}\overline{X}$ et que la moyenne $\overline{X}$ est proche de 0, $\hat{a}$ est très proche de $\overline{Y}$, moyenne de la composante irrégulière (soit proche de la moyenne théorique a égale à 0 pour un schéma additif et à 1 pour un schéma multiplicatif).

L'effet de Pâques s'en déduit par $\hat{E}_{ij} = \hat{Y}_{ij} = \hat{a} + \hat{b}X_{ij}(w)$ ou, plus exactement par $\hat{E}_{ij} = a + \hat{b}X_{ij}(w)$, ce qui permet d'assurer que Pâques n'a pas d'effet en dehors des mois de mars et avril.

[34] En fait, ce n'est pas tout à fait vrai puisque par exemple, en 2008 Pâques tombera un 23 mars et si w=24, on aura alors un jour en février.

Le modèle proposé par X12-ARIMA ressemble donc beaucoup au modèle à effet graduel de X11-ARIMA88[35], et au modèle à effet ponctuel pour $w=1$, mais ici l'estimation du modèle de régression se fait en utilisant toutes les années disponibles et non pas seulement celles des années où Pâques tombe en mars ou après le *w* avril. Par ailleurs, l'estimation est ici faite une première fois sur les données du tableau B13 et une seconde fois sur les données du tableau C13.

Rappelons enfin qu'à cette étape du traitement, X12-ARIMA permet aussi de rechercher dans la composante irrégulière d'autres effets de calendrier.

5.3.2.2 Exemple

Si on utilise ce modèle dans X12-ARIMA, l'estimation est faite une première fois sur les données du tableau B13 puis, une seconde fois, sur les données du tableau C13.
Les valeurs de la variable à expliquer Y sont donc ici les 114 valeurs de B13 (divisées par 100) :

B13: SÉRIE DES FACTEURS IRRÉGULIERS

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.535	100.471	97.916
1986	103.127	98.213	95.390	107.358	99.648	98.460	102.766	94.737	100.185	102.178	98.661	101.186
1987	95.874	101.153	101.588	102.296	98.892	100.097	102.280	96.086	100.774	98.541	99.781	101.664
1988	96.897	102.248	101.498	98.202	101.130	101.564	97.460	101.047	101.999	95.053	100.374	101.647
1989	101.028	99.859	98.753	99.913	99.930	102.961	97.968	100.893	98.592	99.146	102.274	98.212
1990	101.457	98.700	100.249	99.026	102.444	97.743	100.912	102.425	96.873	101.503	101.259	96.413
1991	103.060	99.489	97.167	101.999	100.597	97.834	103.528	99.549	98.111	101.645	98.541	98.994
1992	100.887	100.985	99.061	101.438	97.878	100.337	101.233	98.474	101.637	101.286	99.015	100.115
1993	96.558	101.568	101.173	99.389	98.776	101.209	100.148	101.117	100.329	97.915	100.604	101.319
1994	98.715	99.149	100.309	98.327	101.865	100.715	97.384	103.706	100.376	97.811	100.452	100.387
1995	100.292	98.592	100.227	.	.	.	.	.	.	.	.	.

Si nous supposons par exemple *w*=5, on peut calculer la variable X (nulle partout sauf pour les mois de mars et avril) :

ANNÉE	PÂQUES	5 jours avant	Nombre de jours en mars	$X_{i,mars}$	$X_{i,avril}$
1985	7 avril	3 avril	0	0	0
1986	30 mars	26 mars	5	1	-1
1987	19 avril	15 avril	0	0	0
1988	3 avril	30 mars	2	0.4	-0.4
1989	26 mars	22 mars	5	1	-1
1990	15 avril	11 avril	0	0	0
1991	31 mars	27 mars	5	1	-1
1992	19 avril	15 avril	0	0	0
1993	11 avril	7 avril	0	0	0
1994	3 avril	30 mars	2	0.4	-0.4
1995	16 avril	12 avril	0	0	0

Ainsi, en 1988, dans les 5 jours qui précédaient Pâques (Pâques compris), 2 étaient en mars (les 30 et 31).

[35] C'est d'ailleurs pour cela que cette modélisation est référencée dans X12-ARIMA par le mot clef Sceaster, le SC signifiant Statistique Canada

D'où la variable explicative *X* :

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985										0	0	0
1986	0	0	1	-1	0	0	0	0	0	0	0	0
1987	0	0	0	0	0	0	0	0	0	0	0	0
1988	0	0	0.4	-0.4	0	0	0	0	0	0	0	0
1989	0	0	1	-1	0	0	0	0	0	0	0	0
1990	0	0	0	0	0	0	0	0	0	0	0	0
1991	0	0	1	-1	0	0	0	0	0	0	0	0
1992	0	0	0	0	0	0	0	0	0	0	0	0
1993	0	0	0	0	0	0	0	0	0	0	0	0
1994	0	0	0.4	-0.4	0	0	0	0	0	0	0	0
1995	0	0	0									

On a $\overline{Y} = 0.9999$ et $\overline{X} = 0$, ce qui entraîne :

$$\hat{a} = \overline{Y} - \hat{b}\overline{X} = 0.9999$$

$$Var(X) = \frac{2}{T}\sum_i \left(\frac{n_i}{w}\right)^2 = \frac{2}{114}\left[1 + (0.4)^2 + 1 + 1 + (0.4)^2\right] = \frac{2 * 3.32}{114}$$

$$Cov(X,Y) = \frac{1}{T}\sum_i \frac{n_i}{w}\left[Y_{i3} - Y_{i4}\right]$$

$$= \frac{1}{114}\left[\frac{(95.390-107.358)+0.4*(101.498-98.202)+(98.753-99.913)+(97.167-101.999)+0.4*(100.309-98.327)}{100}\right]$$

$$= \frac{-0.15849}{114}$$

et donc $\hat{b} = \dfrac{-0.15849}{6.64} = -0.02387$

L'effet de Pâques est donc estimé par $\hat{E}_{ij} = 1 + \hat{b}X_{ij}(w) = 1 - 0.02387 * X_{ij}(w)$, ce qui conduit au tableau :

B16H : ESTIMATION PRÉLIMINAIRE DE L'EFFET DE PAQUES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.000	100.000	100.000
1986	100.000	100.000	97.613	102.387	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1987	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	99.045	100.955	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1989	100.000	100.000	97.613	102.387	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1991	100.000	100.000	97.613	102.387	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1993	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1994	100.000	100.000	99.045	100.955	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1995	100.000	100.000	100.000	.	.	.	.	.	.	.	.	.

Avec, par exemple pour mars 1988 : $\hat{E}_{1988,3} = 1 - 0.02387 * 0.4 = 0.99045$

5.3.3 Le modèle « Easter »

Ce modèle, comme le précédent, entre dans le cadre plus général de l'estimation des effets de calendrier proposée par la méthode X11 à la fin des étapes B et C et plus précisément à partir des estimations de la composante irrégulière des tableaux B13 et C13.

5.3.3.1 Modèle et estimation

La fête de Pâques est supposée ici encore avoir un impact sur les *w* jours ($1 \leq w \leq 25$) qui précèdent le dimanche pascal et on pose le modèle $Y_{ij} = a + bX_{ij}(w) + \varepsilon_{ij}$ où :

- Y_{ij} est la valeur de l'irrégulier (du tableau B13 ou C13) correspondant à l'année *i* et à la période (mois ou trimestre) *j*.
- Soit, pour une année *i* donnée, n_{ij} le nombre de jours, parmi les *w* jours avant Pâques (Pâques exclus) qui tombent dans le mois (ou trimestre) *j*. Alors on défini d'abord la variable $Z_{ij}(w) = \frac{n_{ij}}{w}$. Compte tenu des contraintes sur *w*, cette variable est nulle sauf pour les mois de février, mars et avril. Par contre elle possède une certaine saisonnalité : les valeurs concernant le mois de février seront par exemple structurellement plus faibles. La variable $X_{ij}(w)$ s'obtient en enlevant à $Z_{ij}(w)$ la moyenne $\overline{Z}_{.j}(w)$ correspondant au mois *j* et calculée sur les années disponibles. On a donc : $X_{ij} = \frac{n_{ij}}{w} - \overline{Z}_{.j}(w)$. Ainsi, cela permet de conserver le niveau de la série en annulant l'effet de Pâques sur l'ensemble des mois concernés. La variable explicative est alors de moyenne nulle et sans saisonnalité.

Les contraintes sur la valeur de *w* entraînent que seuls les mois de février, mars et avril peuvent faire l'objet de corrections. Comme précédemment, les valeurs de *a* et *b* sont estimées par moindres carrés ordinaires et l'effet de Pâques s'en déduit par $\hat{E}_{ij} = a + \hat{b}X_{ij}(w)$.

5.3.3.2 Exemple

Si on utilise ce modèle dans X12-ARIMA, l'estimation est faite une première fois sur les données du tableau B13 puis, une seconde fois, sur les données du tableau C13.
Les valeurs de la variable à expliquer *Y* sont donc ici les 114 valeurs de B13 (divisées par 100).

Si nous supposons par exemple *w*=5, on peut calculer la variable *Z* (nulle ici partout sauf pour les mois de mars et avril) :

ANNÉE	PÂQUES	5 jours avant	Nombre de jours en mars	Nombre de jours en avril	$Z_{i,mars}$	$Z_{i,avril}$
1985	7 avril	2 avril	0	5	0	1
1986	30 mars	25 mars	5	0	1	0
1987	19 avril	14 avril	0	5	0	1
1988	3 avril	29 mars	3	2	0.6	0.4
1989	26 mars	21 mars	5	0	1	0
1990	15 avril	10 avril	0	5	0	1
1991	31 mars	26 mars	5	0	1	0
1992	19 avril	14 avril	0	5	0	1
1993	11 avril	6 avril	0	5	0	1
1994	3 avril	29 mars	3	2	0.6	0.4
1995	16 avril	11 avril	0	0	0	1

Ainsi, en 1988, dans la période de 5 jours avant Pâques, on en avait 3 en mars (les 29, 30 et 31) et 2 en avril (les 1 et 2).

Les moyennes de $Z_{i,mars}$ et $Z_{i,avril}$ sont calculées sur les années complètes donc en enlevant les années 1985 et 1995. Dans ces conditions, $\overline{Z}_{mars} = (1+0.6+1+1+0.6)/9 = 0.46667$ et $\overline{Z}_{avril} = 0.53333$.

En corrigeant la variable *Z* de ces moyennes mensuelles, on obtient la variable explicative *X* :

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	0	0	0
1986	0	0	0.5333	-0.5333	0	0	0	0	0	0	0	0
1987	0	0	-0.4667	0.4667	0	0	0	0	0	0	0	0
1988	0	0	0.1333	-0.1333	0	0	0	0	0	0	0	0
1989	0	0	0.5333	-0.5333	0	0	0	0	0	0	0	0
1990	0	0	-0.4667	0.4667	0	0	0	0	0	0	0	0
1991	0	0	0.5333	-0.5333	0	0	0	0	0	0	0	0
1992	0	0	-0.4667	0.4667	0	0	0	0	0	0	0	0
1993	0	0	-0.4667	0.4667	0	0	0	0	0	0	0	0
1994	0	0	0.1333	-0.1333	0	0	0	0	0	0	0	0
1995	0	0	-0.4667	.	.	.	.	.	.	.	.	.

Dans le cas le plus général, les valeurs du régresseur pour les mois de février, mars et avril peuvent ne pas être nulles et le calcul de l'estimation de *b* relativement complexe. Ici, comme dans le cas précédent, on devine que ce calcul ne fera intervenir que les différences entre les valeurs de l'irrégulier des mois de mars et avril. Quoiqu'il en soit, on obtient $\hat{a} = 0.999646$ et $\hat{b} = -0.02604$.

L'effet de Pâques est donc estimé par $\hat{E}_{ij} = 1 + \hat{b}X_{ij}(w) = 1 - 0.02604 * X_{ij}(w)$, ce qui conduit au tableau :

B16H : ESTIMATION PRÉLIMINAIRE DE L'EFFET DE PAQUES

ANNÉE	JANV	FÉVR	MARS	AVRIL	MAI	JUIN	JUIL	AOÛT	SEPT	OCT	NOV	DÉC
1985	.	.	.	.	.	.	.	.	.	100.000	100.000	100.000
1986	100.000	100.000	98.6114	101.3886	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1987	100.000	100.000	101.2150	98.7850	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	99.6529	100.3471	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1989	100.000	100.000	98.6114	101.3886	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	101.2150	98.7850	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1991	100.000	100.000	98.6114	101.3886	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	101.2150	98.7850	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1993	100.000	100.000	101.2150	98.7850	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1994	100.000	100.000	99.6529	100.3471	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1995	100.000	100.000	101.2150	.	.	.	.	.	.	.	.	.

Avec, par exemple pour mars 1988 : $\hat{E}_{1988,3} = 1 - 0.02604 * 0.1333 = 0.996529$.

BIBLIOGRAPHIE

AKAIKE H.(1980), *Seasonal Adjustment by a Bayesian Modeling,* Journal of Time Series Analysis, n°1, pp 1-13

ARMATTE M.(1992), *Conjonctions, conjoncture et conjecture. Les baromètres économiques (1885-1930),* Histoire et Mesure, Vol 7, n°1/2, pp 99-149

BARTLETT. (1950) : *Periodogram Analysis and Continuous Spectra*, Biometrika, Vol n°35, pp 1-16

BATEMAN D.V., MAYES F. (1970) *Holiday Adjustment of Retail Sales,* U.S. Bureau of the Census, unpublished memorandum

BELL W. R., CHEN B., FINDLEY D. F., MONSELL B. C., OTTO M. C. (1997), *New capabilities and methods of the X-12-ARIMA seasonal adjustment program,* publication of the Time Series Staff, Statistical Research Division, US Bureau of the Census.

BELL W. R., CHEN B., FINDLEY D. F., MONSELL B. C., OTTO M. C. (1997), *X12-ARIMA Reference Manual (Final version 0.1),* publication of the Time Series Staff, Statistical Research Division, US Bureau of the Census.

BELL W.R., HILLMER S.C. (1984), *Issues Involved With the Seasonal Adjustment of Economic Time Series,* Journal of Business and Economic Statistics, Vol 2, n°4

BONGARD, J. (1962), *Quelques remarques sur les moyennes mobiles,* L'élimination des variations saisonnières à l'aide de calculatrices électroniques, OCDE.

BOURNAY J., LAROQUE G. (1979) *Réflexions sur la méthode d'élaboration des comptes trimestriels,* Annales de l'INSEE, n°36

BOX G.E.P., JENKINS G.M. (1970), *Time Series Analysis : Forecasting and Control,* San Francisco : Holden Day.

BOX G.E.P., TIAO G.C. (1975), *Intervention Analysis with Applications to Economic and Environmental Problems*, Journal of The American Statistical Association, Vol n°70, pp 70-79

BUREAU OF THE CENSUS, Time Series Staff (1998) *X-12-ARIMA Reference Manual Final Version 0.1*

BURMAN J.P. (1980), *Seasonal Adjustment by Signal Extraction,* Journal of the Royal Statistical Society, Serie A, n°143, pp321-337

CHEN B., FINDLEY D.F. (1998) *Comparison of X11 and RegARIMA Easter Holiday Adjustments,* Statistical Research Division, U.S. Bureau of the Census, manuscript in preparation.

CHEN C., LIU L.M. (1993), *Joint Estimation of Model Parameters and Outlier Effects in Time Series*, Journal of The American Statistical Association, Vol n°88, pp 284-297

CHOLETTE P.A. (1978) : *A Comparison and Assessment of Various Adjustment Methods of Sub-Annual Series to Yearly Benchmarks,* Seasonal Adjustment and Time Series Analysis Staff, Statistics Canada, Ottawa.

CHOLETTE P.A. (1981) : *A Comparison of Various Trend-Cycle Estimators,* Time Series Analysis, O.D. Anderson & M.R. Perryman Editors, North-Holland.

CHOLETTE P.A., DAGUM E.B. (1994) : *Benchmarking Time Series with Autocorrelated Survey Errors,* International Statistical Review, Vol. 62, n°3.

CLEVELAND W.S. (1979), *Robust Locally Weighted Regression and Smoothing Scatterplots*, Journal of The American Statistical Association, Vol n°74, pp 829-836

CLEVELAND R.B., CLEVELAND W.M., McRAE J.E., TERPENNING I. (1990) : *STL A Seasonal-Trend Decomposition Procedure Based on Loess,* Journal of Official Statistics, Vol 6, n°1, pp 3-73

COOLEY J.W., TUKEY J.W. (1965), *An Algorithm for the Machine Calculation of Complex Fourier Series*, Mathematics of Computation, Vol n°19, pp 297-301

DAGUM E. B. (1980), *La méthode de désaisonnalisation X-11-ARMMI*, Statistique Canada, n°12-564F.
DAGUM E. B. (1988), *The X11-ARIMA/88* S*easonal Adjustment Method, Foundations and User's Manual*, Time Series Research and Analysis Division, Statistics Canada.
DAGUM E.B., HUOT G., MORRY M. (1988) *Seasonal adjustment in the eighties: Some problems and solutions*, The Canadian Journal of Statistics, Vol. 16, Supplement
DOHERTY, M. (1992), *The surrogate Henderson filters in X11*, Technical report, Department of statistics, Wellington, New Zealand.
DURBIN J., KENNY P.B. (1982), *Local trend estimation and seasonal adjustment of economic and social time series*, Journal of the Royal Statistical Society A, n°145
FINDLEY D.F., MONSELL B.C., BELL W.R., OTTO M.C., CHEN B. (1998) : *New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program*, Journal of Business and Economic Statistics, Vol. 16, n°2
FISHER, A. (1937), *A brief note on seasonal variation*, Journal of Accountancy, n°64, pp 54-59
FISCHER B. (1995), *Decomposition of times series : comparing different methods in theory and in practice*, EUROSTAT.
GARDNER M. (1981) *Mathematical Games*, Scientific American, February 1981, 17-20
GOMEZ V., MARAVALL A. (1997), *Programs TRAMO and SEATS Instructions for the User (Beta Version)*.
GOURIÉROUX C., LEGALLO F. (1981), *Construction de moyennes mobiles par minimisation sous contraintes d'une forme quadratique des coefficients*, Annales de l'INSEE, n°42.
GOURIEROUX, C. , MONFORT, A. (1990), *Séries temporelles et modèles dynamiques*, Économica.
GRAY A., THOMSON P. (1996), *On a family of moving-average trend filters for the ends of series*, ASA proceedings of the survey methods section.
GRAY A., THOMSON P. (1996), *Design of moving-average trend filters using Fidelity and Smoothness Criteria*, Athens conference in applied probability and time series, Volume II: Time series analysis in memory of E.J. HANNAN, Ed. P.M. ROBINSON and M. ROSENBLATT.
GRUN-REHOMME M., LADIRAY D. (1994*), Moyennes mobiles centrées et non centrées : construction et comparaison*, Revue de Statistique Appliquée, Volume XLII, n°3.
HARVEY A.C., KOOPMAN S.J., DOORNIK J.A., SHEPARD N. (1995), *Structural Time Series Modelling Program Manual*, Chapman and Hall.
HENDERSON R. (1916), *Note on graduation by adjusted average*, Transactions of the Actuarial Society of America, Volume 17, pages 43-48.
HENDERSON R. (1924), *A new method of graduation*, Transactions of the Actuarial Society of America, Volume 25, pages 29-40.
HIGGINSON J. (1975) : *Test F de la présence d'une saisonnalité mobile lors de l'utilisation de la variante X-11 de la Census Method II*, document de recherche, Division de la désaisonnalisation et de l'analyse des séries chronologiques, Statistique Canada, Ottawa
HILLMER S.C., TIAO G.C. (1982), *An ARIMA Model Based Approach to Seasonal Adjustment*, Journal of the American Statistical Association, n°77 pp 63-70
HYLLEBERG S. (1992), *The historical perspective* in *Modelling seasonality*, Oxford University Press, pp 15-25.
JEVONS W.S (1862), *On the study of periodic commercial fluctuations* in *Investigations in currency and finance*, London : Macmillan, 1884
KENDALL M. (1973), *Time Series*, Griffin.
KITAGAWA G., GERSCH W.(1984), *A Smoothness Priors State Space Modelling of Time Series with Trend and Seasonality*, Journal of the American Statistical Association, n°79 pp 378-389
KOOPMANS L.H. (1974), *The Spectral Analysis of Time Series*, Academic Press.

LAKER L.G. (1976) *Mathematical Note on Easter Correction*, working paper, Australian Bureau of Statistics

LAKER L.G. (1976) *Slightly Less Mathematical Note on the Mathematical Note on Easter Correction*, working paper, Australian Bureau of Statistics

LANIEL N. (1985) : *Design Criteria for the 13-term Henderson End-Weights*, working paper, Methodology Branch, Statistics Canada, Ottawa

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, pages 105-127.

LOTHIAN J. (1978) : *The Identification and Treatment of Moving Seasonality in the X-11 Seasonal Adjustment Method,* working paper 78-10-004, Methodology Branch, Seasonal Adjustment and Time Series Analysis Staff, Statistics Canada, Ottawa

LOTHIAN J., MORRY M. (1978) : *A Set of Quality Control Statistics for the X11-ARIMA Seasonal Adjustment Method*, working paper 78-10-005, Methodology Branch, Seasonal Adjustment and Time Series Analysis Staff, Statistics Canada, Ottawa

LOTHIAN J., MORRY M. (1978) : *Test de la présence d'une saisonnalité identifiable dans le programme X-11-ARMMI*, document de recherche 78-10-002F, Division de la désaisonnalisation et de l'analyse des séries chronologiques, Statistique Canada, Ottawa

MABALLÉE, Colette et Berthe (1906), "Peaks and Peaks and Correlograms", *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 3, 139-167.

MACAULAY, F.R. (1931), *The Smoothing of Time Series,* National Bureau of Economic Research

MARAVALL A. (1995), *Unobserved Components in Economic Time Series,* dans Handbook of Applied Econometrics, Blackwell Handbooks in Economics, M. Hashem Pesaran et Mike Wickens (Editeurs)

MARCH, L. (1905), *Comparaison Numérique de Courbes Statistiques*, Journal de la Société Statistique de Paris

MENDERHAUSEN, H. (1937), *Annual Survey of Statistical Technique : Methods of Computing and Eliminating Changing Seasonal Fluctuations,* Econometrica, Vol n°5 pp 234-262

MUSGRAVE J. (1964), *A Set of End Weights to End all End Weights*, working paper, US Bureau of Census, Washington.

MUSGRAVE J. (1964), *Alternative Sets of Weights for Proposed X-11 Seasonal Factor Curve Moving Averages*, working paper, US Bureau of Census, Washington.

NERLOVE M., GRETHER D.M., CARVALHO J.L. (1979), *Analysis of Economic Time Series : a Synthesis,* New York : Academic Press

O'BEIRNE T. (1966) *The Regularity of Easter*, Bulletin of the Institute of Mathematics and Its applications, Vol. 2, n°2, pages 46-49

PANNEKOEK A. (1961), *A history of Astronomy*, New York : Wiley

PERSONS W.M. (1919), *Indices of Business Conditions*, Review of Economic Statistics n°1 pp 5-107

PRIESTLEY M.B. (1965), *Evolutionary Spectra and Nonstationary Processes,* Journal of the Royal Statistical Society, Serie B, Vol n°27 pp 204-237

SHISKIN J., YOUNG A., MUSGRAVE J.C. (1967), *The X11 variant of the Census Method II Seasonal Adjustment Program*, Washington DC, Technical Paper n°15, Bureau of Census, US Department of Commerce

SHISKIN J., MUSGRAVE J., SOMER M., BAKKA N., ROSENBLATT H.M. (1969) : *X11Information for the user,* Papers prepared for the Seminar on seasonal adjustments of the National Association of Business Economists, Philadelphia, Pennsylvania, March 10, 1969, US Department of Commerce, Bureau of the Census, Washington.

SLUTSKY E (1927), *The summation of random causes as the source of cyclical processes*, Econometrica, vol n°84 pp 105-146 (1937), traduction d'un article paru en russe (Conjoncture Institute, Moscou)
SPENCER J. (1904), *On the graduation of rates of sickness and mortality*, Journal of the institute of actuaries, volume 38.
TUKEY J. W. (1950), *The Sampling Theory of Power Spectrum Estimates* in Proceedings of the Symposium on Applications of Autocorrelation Analysis to Physical Problems, Washington DC, Department of the Navy, Office of Naval Research
YOUNG A.H. (1965), *Estimating Trading-Day Variations in Monthly Economic Series*, Washington DC, Technical Paper n°12, Bureau of Census, US Department of Commerce
YULE, G.U. (1921), *On the Time Correlation Problem, With Especial reference to the Variate-Difference Correlation Method*, Journal of The Royal Statistical Society, n°84 pp 497-526
YULE, G.U. (1927), *On a method of investigating periodicities in disturbed series with special reference to Wolfer's sunspot numbers*, Philosophical Transactions of The Royal, serie A, n°226 pp 267-298
WALLIS K. (1982), *Seasonal adjustment and revision of current data: linear filters for the X11 Method*, Journal of the Royal Statistical Society A, 145, pages 74-85 .
WHITTAKER E. T. (1923), *On a new method of graduation*, Proceedings of the Edinburgh mathematical Society, n°41.