

EFFETS SAISONNIERS DANS LES SÉRIES DÉMOGRAPHIQUES

J.-C. Labat

Les séries démographiques (nombres de naissances, de mariages...) présentent des mouvements saisonniers, plus ou moins marqués, qui évoluent avec le temps. Ces phénomènes sont intéressants en eux-mêmes d'un point de vue sociologique.

En outre, comme pour les séries économiques, une analyse aussi précise que possible des variations saisonnières est nécessaire pour apprécier convenablement l'évolution réelle des phénomènes de fond (natalité, nuptialité...) notamment d'un point de vue conjoncturel. Des travaux concrets dans ce domaine sont faits à l'INSEE, au département de la Démographie, depuis de nombreuses années. La méthode utilisée et les résultats obtenus pour les mariages sont présentés ici.

Les nombres mensuels de mariages sont soumis à des variations saisonnières **de très forte amplitude** : ainsi en 1991, il y a eu six fois plus de mariages en juin qu'en janvier. En outre, le profil saisonnier **évolue rapidement** avec le temps : par exemple le rapport entre les nombres de mariages de juin et de janvier, six en 1990, n'était que de deux en 1966. Enfin, la série présente un trait qui vient compliquer l'analyse ; elle est affectée de fortes variations, en apparence irrégulières, en tout cas autres que saisonnières au sens classique : ainsi le nombre de mariages de juin a augmenté de 35 % de 1989 à 1990, mais baissé de 8 % de 1990 à 1991, alors que le nombre annuel a varié respectivement de + 3 % et - 2 %.

Ces irrégularités, par leur importance, rendent l'évaluation des coefficients saisonniers difficile, notamment pour les derniers mois, et leur extrapolation pour les mois à venir très incertaine.

La méthode utilisée pour réduire progressivement les irrégularités et améliorer simultanément l'estimation des coefficients saisonniers et de la tendance est décrite ci-après. Elle comporte plusieurs phases (A à F).

Les désaisonnalisations sont faites avec le logiciel X-11 ARMMI confectionné par Statistique Canada. Elles portent toutes sur la période allant de janvier 1966 à avril 1992. On utilise évidemment un modèle multiplicatif :

$$X = T \times FA \times CS \times FI$$

X: Série à désaisonnaliser

T: Tendance

FA: Facteur d'ajustement préalable (le cas échéant)

CS: Coefficient saisonnier

FI: Facteur irrégulier

A) La désaisonnalisation de la série brute des nombres mensuels de mariages, sans ajustement en fonction des jours ouvrables, fournit des facteurs irréguliers dont l'écart-type est 8,2 (les facteurs sont exprimés en % ; leur moyenne est voisine de 100).

B) Ces irrégularités apparentes résultent en grande partie de l'inégale répartition des mariages par jour de la semaine.

La régression "en fonction des jours ouvrables" réalisée au cours de la désaisonnalisation (mais non appliquée) explique environ les 2/3 de la variance. Les coefficients de la régression sont respectivement (du lundi au dimanche) : 0,8 ; 0,2 ; 0,5 ; 0,3 ; 1,1 ; 3,3 ; 0,8.

En fait, la statistique des mariages par jour est disponible. Sur la période étudiée les poids de chaque jour sont en moyenne : 0,3 ; 0,1 ; 0,1 ; 0,3 ; 0,8 ; 5,3 ; 0. Le poids du samedi est ainsi bien supérieur à l'estimation réalisée par X11.

En introduisant ces poids *a priori* dans la désaisonnalisation, on aboutit à des facteurs irréguliers dont l'écart-type est 4,9.

C) Au cours de la période les poids des différents jours ont évolué. Ainsi le poids du samedi a augmenté de 4,8 à 5,7. On doit donc réduire quelque peu la variance résiduelle en divisant la série par un coefficient CJOURM tenant compte de la composition journalière de chaque mois et calculé *a priori* à partir des poids journaliers du moment fournis par la statistique. En fait l'amélioration est très faible globalement, l'écart-type des facteurs irréguliers étant encore de 4,9. Elle devrait néanmoins ne pas être insignifiante pour les dernières années de la période (de même que pour les premières).

D) On observe que l'écart-type des facteurs irréguliers est particulièrement élevé pour mars (8,0) et avril (6,8). En général, lorsque mars est fort, avril est faible et vice versa. Ce phénomène serait lié aux vacances scolaires de printemps, dont les dates ont varié très sensiblement suivant les années.

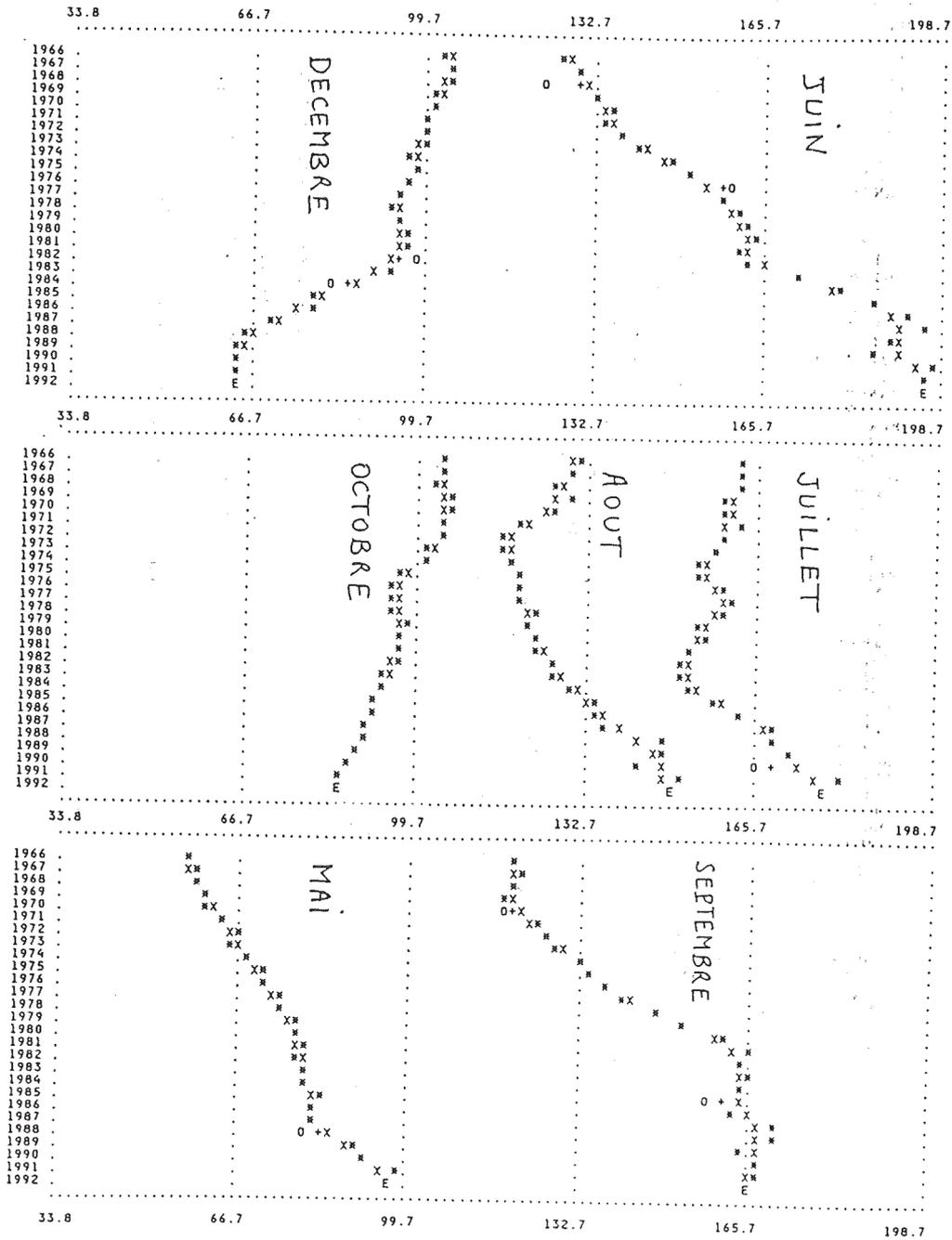
Les nombres de mariages de mars et d'avril de chaque année sont remplacés par une répartition, au prorata de CJOURM, de leur somme.

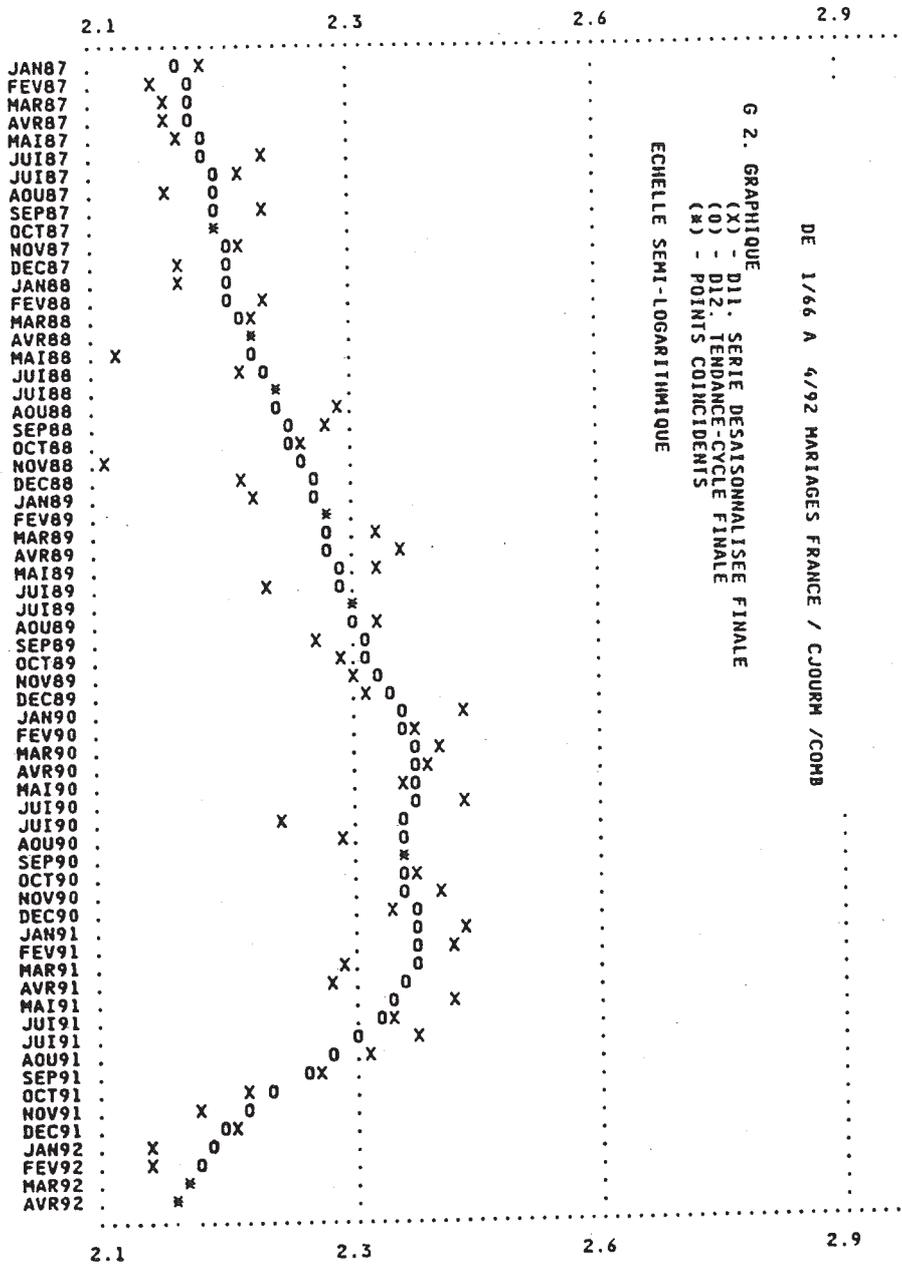
L'écart-type des facteurs irréguliers est ainsi ramené à 4,0 (1,8 pour mars et avril). La série analysée (nombre de mariages, corrigé pour mars et avril, divisé par CJOURM) est appelée FRAMJJ.

G 3. GRAPHIQUE

- (X) - D10. FACTEURS SAISONNIERS FINALS
- (O) - D 8. RAPPORTS S-I FINAL(E)S NON MODIFIE(E)S
- (+) - D 9. RAPPORTS FINAL(E)S SAHS LES VALEURS EXTREMES
- (*) - POINTS COINCIDENTS
- (E) - FACTEURS SAISONNIERS EXTRAPLES

ECHELLE ARITHMETIQUE





E) Un examen attentif des irréguliers montre que la place des samedis joue un rôle important. Par exemple, sur les quatre mois de décembre de la période qui se terminent par un samedi, trois ont un facteur irrégulier inférieur à 90 (c'est-à-dire présentant une valeur inférieure de plus de 10 % à la "normale"). Quant au quatrième, en 1983, son facteur irrégulier (inférieur à 100 mais nettement supérieur à 90) est vraisemblablement mal estimé, en raison d'une cassure dans l'évolution du coefficient saisonnier de décembre.

Une série de coefficients mensuels COMA, où la valeur d'un mois donné est fonction du nom du mois d'une part et du dernier jour du mois d'autre part (par exemple 87 pour les mois de décembre qui se terminent par un samedi), est donc définie.

Toutefois, cette série n'est pas utilisée directement comme facteur d'ajustement préalable. En effet, la moyenne mobile sur plusieurs mois du produit CJOURM x COMA, pondérée par les coefficients saisonniers correspondants, présente une évolution sinueuse dont les bosses et les creux peuvent atteindre des valeurs sensiblement éloignées de l'unité (à une puissance de 10 près). L'estimation de la tendance risquerait donc d'être faussée. C'est la série COMB, obtenue en modulant la série COMA pour atténuer ce phénomène, qui est utilisée comme facteur d'ajustement préalable de la série FRAMJJ.

On obtient ainsi des facteurs irréguliers dont l'écart-type vaut 2,2 (en imposant la même moyenne mobile pour l'estimation de la tendance que celle retenue pour les traitements antérieurs, c'est-à-dire une moyenne de Henderson à 23 termes).

F) La composante irrégulière ayant été très sensiblement réduite, il peut sembler judicieux, afin d'améliorer l'évaluation des coefficients saisonniers :

- d'opérer un lissage moins puissant pour les obtenir (moyenne mobile 3x3, au lieu de 3 x 5 antérieurement) ;
- de corriger moins les valeurs extrêmes (limites 2-3,5 au lieu de 1,5-2,5 antérieurement);
- de conserver en revanche une moyenne de Henderson à 23 termes pour estimer la tendance (alors que l'option choisie par X11 serait une moyenne à 13 termes).

Ces changements d'option abaissent l'écart-type des facteurs irréguliers à 1,6, valeur qui reste néanmoins assez élevée. C'est de cet ajustement saisonnier que sont tirés les résultats présentés en annexe. Même si l'incertitude portant sur les résultats du traitement a été réduite, elle reste toutefois non négligeable, notamment en ce qui concerne l'estimation des coefficients saisonniers et de la tendance des derniers mois, surtout si ce sont, comme c'est le cas ici, des mois à faible nuptialité.

Ainsi la forte baisse de la nuptialité qui apparaît pour les derniers mois est vraisemblablement exagérée.