12 Désaisonnalisation et estimation de la tendance-cycle

En mars 2000, des lignes directrices concernant la désaisonnalisation ont été rédigées à l'intention du Comité des méthodes et des normes de Statistique Canada. Le présent texte propose leur mise à jour, à la suite de l'adoption de X-12-ARIMA (Findley et al., 1998) par Statistique Canada et leur normalisation en vue de les harmoniser avec celles du Census Bureau des États-Unis (McDonal-Johnson et al., 2006a, 2006b) et d'EUROSTAT (Mazzi, 2008).

12.1 Portée et objet

12.1.1 Désaisonnalisation

Une série chronologique est une séquence de mesures visant une variable observée au fil du temps. Dans la plupart des cas, ces mesures sont interdépendantes et c'est cette interdépendance qui présente un intérêt dans le cadre de la désaisonnalisation. Pour les besoins de la désaisonnalisation, nous supposons que la série chronologique est observée mensuellement ou trimestriellement et qu'elle est constituée de trois éléments distincts, à savoir la tendance-cycle, les effets saisonniers et de calendrier combinés et l'irrégulier. L'objectif de la désaisonnalisation est de déceler et d'estimer les effets saisonniers et de calendrier combinés pour les éliminer de la série chronologique. La série qui en résulte est alors dite désaisonnalisée et ne comprend plus que la tendance-cycle et l'irrégulier.

Les effets saisonniers sont les fluctuations infra-annuelles (mensuelles, trimestrielles) qui se répètent plus ou moins régulièrement d'année en année. Ils résultent des effets combinés des événements reliés au climat, des décisions institutionnelles ou des modes de fonctionnement qui se reproduisent avec une certaine régularité au cours de l'année. Les effets de calendrier sont liés à la composition du calendrier. Ils comprennent les effets de jours ouvrables associés aux nombres de jours de semaine contenus dans un mois, les effets des congés à occurrence variable associés aux congés à date non fixe tels que Pâques et d'autres événements prévisibles du calendrier. Les effets de jours ouvrables se produisent lorsque le niveau d'activité varie d'un jour à l'autre de la semaine. L'effet de Pâques s'assimile à la variation de niveau causée par le déplacement d'un volume d'activité d'avril à mars lorsque Pâques tombe en mars au lieu d'avril (comme c'est habituellement le cas).

La série désaisonnalisée permet d'évaluer la direction de la tendance-cycle, grâce à des comparaisons de mois en mois ou de trimestre en trimestre.

La tendance est le mouvement sous-jacent à long terme. Ce mouvement dure plusieurs années. Le cycle, qu'on appelle également cycle économique, est une oscillation quasi périodique, dont la durée excède un an, autour de la tendance à long terme. Il est caractérisé par l'alternance de périodes d'expansion et de contraction. Il est difficile d'estimer séparément la tendance et le cycle; on les étudie et on les analyse donc ensemble comme un tout appelé tendance-cycle.

L'irrégulier, ou la composante irrégulière, se manifeste dans des variations aléatoires, qui s'avèrent être des fluctuations imprévisibles causées par des événements (de toutes sortes) indépendants de la tendance-cycle, de la saisonnalité ou des effets de calendrier.

12.1.2 Estimation de la tendance-cycle

La désaisonnalisation d'une série très volatile ne suffit pas nécessairement pour pouvoir tirer des conclusions sur la direction de la tendance-cycle. Le cas échéant, il est bon de poursuivre le lissage de la série désaisonnalisée pour éliminer le plus possible la composante irrégulière. L'estimation de la tendance-cycle qui en résulte doit être considérée comme de l'information auxiliaire sur la série désaisonnalisée.

12.1.3 X-12-ARIMA

Le fondement du programme de désaisonnalisation de Statistique Canada est la variante X-11 de la Census Method II de 1967 (Shiskin et al., 1967; Ladiray et Quenneville, 2001). En 1980, Statistique Canada ajoutait les modèles autorégressifs à moyenne mobile intégrée (ARIMA pour autoregressive integrated moving average) (Box et Jenkins, 1976) pour pouvoir projeter et rétropoler les séries avant leur désaisonnalisation. Plusieurs autres changements ont alors été instaurés pour améliorer la variante X-11. Cette nouvelle version a été appelée X-11-ARIMA (Dagum, 1980). En général, l'extension des séries, grâce aux prévisions ARIMA, a permis de réduire la révision des séries désaisonnalisées. En 1988, Statistique Canada a diffusé une version améliorée de X-11-ARIMA (Dagum, 1988) utilisable sur micro ordinateurs.

En 1998, le US Census Bureau a lancé la version X-12-ARIMA (Findley et al., 1998). Cette version s'appuie sur la régression linéaire avec erreurs ARIMA (modélisation regARIMA) pour estimer les effets de calendrier, des valeurs aberrantes additives, des changements de niveau et d'autres variables de régression prédéfinies. En outre, X-12-ARIMA permet à ses utilisateurs de définir une régression pour les effets de calendrier inhabituels ou non standards, et comprend une variante de l'algorithme TRAMO (Gomez et Maravall, 1996), pour la modélisation regARIMA automatique. D'autres fonctionnalités du programme sont décrites dans le guide de l'utilisateur (US Census Bureau, 2008). X-12-ARIMA s'adapte à diverses plateformes informatiques, ce qui inclut l'exécutable FORTRAN brut, une version C pour UNIX, et une pour FAME. Enfin, la procédure X12 (SAS Institute Inc, 2007) en SAS® assure la mise en œuvre des options les plus importantes. Sur ce plan, Statistique Canada recommande la méthode X-12-ARIMA, car X-11-ARIMA est progressivement mise hors service et ne bénéficiera plus d'aucun soutien.

12.2 Principes de la désaisonnalisation

La désaisonnalisation vise à éliminer les effets saisonniers et de calendrier combinés. Avant d'appliquer cette méthode, il faut donc s'assurer que de tels effets sont présents et qu'il est possible de les estimer correctement. Lorsqu'il est impossible de repérer des effets saisonniers et/ou de calendrier dans une série chronologique, cette série est réputée désaisonnalisée de facto. Les séries désaisonnalisées ne doivent pas présenter de saisonnalité résiduelle et sont généralement plus lisses que les séries brutes correspondantes.

À mesure que des données nouvelles sont accessibles, il est plus facile de bien estimer les composantes d'une série chronologique : on obtient alors des estimations révisées des valeurs désaisonnalisées antérieures qui sont plus exactes et dont il faut tenir compte. Cependant, il est préférable d'éviter les révisions trop fréquentes, car elles peuvent réduire l'utilité des données désaisonnalisées. Il faut privilégier des options de désaisonnalisation qui réduisent les révisions au minimum sans affecter pour autant la qualité globale de l'ajustement. Il faut également implanter une stratégie permettant de réduire au minimum la fréquence à laquelle les données publiées seront révisées.

Lorsqu'elles sont inappropriées, les options de désaisonnalisation peuvent fausser les résultats. Par conséquent, il faut consacrer suffisamment de temps et d'efforts à l'analyse des séries, au choix des options et à la maintenance de ces dernières. Comme chaque combinaison d'options peut produire des résultats différents d'autres combinaisons, il faut traiter les séries chronologiques comme un tout lorsque ces séries mesurent la même activité économique. Cette approche implique habituellement l'utilisation d'options d'ajustement similaires par les divers secteurs de programme concernés, pour garantir la cohérence des résultats.

12.3 Principes de l'estimation de la tendance-cycle

En tant que complément des séries désaisonnalisées, les estimations de la tendance-cycle peuvent révéler la direction de la tendance à court terme (durant l'année courante). Lorsque de nouveaux points de donnée sont ajoutés aux séries, on peut calculer plus précisément les estimations antérieures de la tendance-cycle; ces estimations font donc l'objet de révisions. Les estimations de la tendance-cycle sont sensibles à la phase courante du cycle économique (point de retournement, récession, reprise ou expansion); la fiabilité des estimations courantes de la tendance-cycle dépend donc de la proximité d'un point de retournement,

ainsi que de l'amplitude du cycle.

Les estimations de la tendance-cycle doivent correspondre en tous points à la série désaisonnalisée diffusée. Si les valeurs désaisonnalisées sont gelées dans une base de données, la tendance cycle doit être estimée d'après la série désaisonnalisée telle que présentée dans cette base. De même, si la série désaisonnalisée a subi d'autres ajustements, telles qu'une agrégation ou bien une réconciliation, la tendance-cycle doit être estimée d'après la série agrégée ou réconciliée.

12.4 Lignes directrices concernant la désaisonnalisation

- Avant de désaisonnaliser une série pour la première fois, s'assurer que la composante saisonnière est identifiable et qu'on peut l'estimer correctement.
- Si une série est exempte de saisonnalité ou d'effet de calendrier, n'appliquer aucun traitement; la série est alors réputée désaisonnalisée de facto.
- Une série désaisonnalisée ne doit contenir ni saisonnalité résiduelle, ni effet de calendrier résiduel.
- Afin d'identifier et de bien estimer les effets saisonniers et les effets de calendrier, il est recommandé d'utiliser des données couvrant une période de 10 à 15 ans. La période minimale est de cinq (5) ans pour estimer correctement un profil saisonnier et de sept (7) ans pour les effets de calendrier tels que les effets de jours ouvrables et de jours fériés mobiles (les congés à occurrence variable).
- Il est recommandé d'utiliser la modélisation regARIMA pour calculer les facteurs de correction des effets de calendrier et les corrections temporaires comme celles des valeurs aberrantes additives connues ou des changements de niveau. En général, il faut user du même modèle regARIMA pour extrapoler les séries, afin de réduire les révisions au sein de la série désaisonnalisée.
- Les options de désaisonnalisation sont multiples; les plus importantes concernent la sélection du modèle de décomposition, ainsi que la spécification d'un modèle regARIMA et de la longueur des filtres saisonniers et de la tendance-cycle. Les processus de sélection automatique de X-12-ARIMA peuvent être utilisés pour une initialisation de ces options. Si l'on dispose de suffisamment de temps ou que les attentes en matière de qualité sont grandes, la sélection automatique doit être examinée en s'appuyant sur d'autres statistiques, sur les connaissances préalables des experts et sur l'analyse graphique.
- Pour chaque série, il faut examiner périodiquement les options de désaisonnalisation, afin de s'assurer qu'elles demeurent applicables et appropriées et d'accroître la précision. À moins de motifs valables, il faut éviter de modifier les options de désaisonnalisation retenues entre deux révisions planifiées.
- Si les principales options de désaisonnalisation doivent généralement demeurer fixes, entre les révisions, les facteurs d'ajustement et les paramètres du modèle regARIMA doivent quant à eux être actualisés; autrement dit, il faut les recalculer en se servant de tous les points de données disponibles. Il peut y avoir exception à la règle, lorsqu'on sait que les observations les plus récentes ont déjà été soumises à des révisions majeures. Le cas échéant, les facteurs (prévus) pour l'année à venir peuvent s'avérer plus appropriés.
- La désaisonnalisation des séries agrégées (ou composées) comprenant plusieurs séries-composantes peut être indirecte – les composantes désaisonnalisées sont agrégées pour former la série agrégée désaisonnalisée – ou directe – la série agrégée est désaisonnalisée indépendamment. La méthode directe peut engendrer des écarts entre la série agrégée et les composantes agrégées après la désaisonnalisation. Au besoin, appliquer une méthode de réconciliation afin de faire concorder la série agrégée désaisonnalisée directement avec ses composantes désaisonnalisées, sans modifier les composantes non désaisonnalisées dans la mesure du possible.
- Qu'il résulte de l'approche indirecte ou de l'approche directe, l'agrégat ne devrait contenir aucune saisonnalité résiduelle et devrait être relativement lisse. L'approche directe avec réconciliation est la plus indiquée lorsqu'on accorde plus d'importance à la série agrégée qu'à ses composantes ou bien que les composantes présentent des composantes saisonnières fort similaires. En général, il est approprié de recourir à l'ajustement indirect quand les séries-composantes présentent des profils saisonniers très différents et que plusieurs d'entre elles peuvent être désaisonnalisées individuellement.

- Forcer les totaux annuels des données désaisonnalisées à être égaux à ceux de la série originale (ou de la série originale corrigée des effets de calendrier) est une approche rarement justifiée d'un point de vue théorique, mais qui pourrait être adoptée quand il est nécessaire d'assurer la cohérence avec des étalons externes, telles qu'au Système de comptabilité nationale, ou de réconcilier un agrégat et ses composantes.

12.4.1 Révisions des données désaisonnalisées

- La publication des données désaisonnalisées révisées doit suivre une politique de révision officielle et s'aligner sur le calendrier de diffusion des données non désaisonnalisées.
- Lorsqu'on emploie un facteur saisonnier actualisé, il n'est pas nécessaire de réviser les estimations désaisonnalisées en remontant en arrière de plus d'une période. Il peut y avoir des exceptions à cette règle, lorsqu'on utilise des observations préliminaires : il est alors recommandé de réviser les facteurs saisonniers chaque fois qu'on révise les données originales. Chaque année, il faut réviser les valeurs désaisonnalisées des trois dernières années dès que les données du premier mois (trimestre) de l'année suivante sont disponibles. Si les valeurs désaisonnalisées sont générées grâce aux facteurs de désaisonnalisation (prévus) de l'année à venir, cette révision annuelle doit s'appliquer aux quatre dernières années.

12.4.2 Estimation de la tendance-cycle

- Il faut appliquer la méthode d'estimation de la tendance-cycle à la série désaisonnalisée publiée pour s'assurer que la ligne de tendance soit centrée sur la série désaisonnalisée. La méthode de Dagum (1996) ou une adaptation/variante appropriée de cette méthode est recommandée pour l'estimation de la tendance-cycle.
- Informer les utilisateurs du fait que les dernières estimations de la tendance-cycle (et surtout de la toute dernière) peuvent subir des révisions lorsque s'ajoute un point de donnée supplémentaire. Il est possible de signaler cette variabilité (celle des estimations qui se trouvent vers l'extrémité de la série) en se servant d'une ligne pointillée sur le graphique de tendance, par exemple, ou en publiant une note d'information en même temps que les données.
- Il faut réviser les estimations de la tendance en remontant aussi loin que pour la révision des estimations désaisonnalisées. Par ailleurs, lorsqu'on révise une série mensuelle type, il faut ajouter trois mois (deux pour une série trimestrielle) durant l'année et six mois (deux pour une série trimestrielle) au moment de la révision annuelle.

12.4.3 Présentation des données et accès aux données

- Il faut calculer les taux de croissance et les variations – d'un mois à l'autre (ou d'un trimestre à l'autre) – en se servant des données désaisonnalisées et les utiliser avec prudence lorsque les séries chronologiques sont hautement instables. Lorsqu'on compare le même mois, d'une année à l'autre, il faut se servir des données corrigées des effets de calendrier ou, en l'absence d'effets de calendrier, des données brutes.
- Les utilisateurs doivent avoir accès à toute la série historique brute, à la série désaisonnalisée et, sur demande, aux options de désaisonnalisation.

12.4.4 Mise en œuvre

Pour obtenir de l'aide concernant l'interprétation et la mise en œuvre de ces lignes directrices, s'adresser au Centre de recherche et d'analyse en séries chronologiques (CRASC), Division des méthodes d'enquête auprès des entreprises.

12.5 Indicateurs de qualité

Les indicateurs suivants peuvent servir à déterminer si une série contient une composante saisonnière :

- Un simple graphique temporel et un graphique année-sur-année. Ces outils permettent d'inspecter la série en vue de déceler des profils saisonniers et de repérer – visuellement – d'autres perturbations;
- Diverses statistiques, telles que les deux tests de Fisher pour la saisonnalité stable et la saisonnalité évolutive, ainsi que les graphiques de spectre décrits dans (Ladiray et Quenneville, 2001; Findley et al., 1998).

La détection de la saisonnalité résiduelle peut se faire grâce aux tests ci-dessus appliqués aux données désaisonnalisées ou au moyen d'autres tests (Ladiray et Quenneville, 2001).

Les statistiques permettant d'évaluer la signification des composantes regARIMA estimées et la qualité globale du modèle ajusté sont décrites dans plusieurs manuels traitant du sujet.

Pour quantifier les révisions, on peut utiliser des statistiques sommaires sur les révisions historiques des niveaux et des changements dans la série désaisonnalisée. Dans le cadre de la désaisonnalisation, les révisions augmentent généralement la précision, car elles reposent sur des observations qui n'étaient pas disponible initialement.

Bibliographie

BOX, G.E.P. et G.M. JENKINS. 1976. *Time Series Analysis, Forecasting and Control*, San Francisco, Holden Day.

DAGUM, E.B. 1980. *La méthode de désaisonnalisation X-11-ARIMA,* publication nº 12-564F au catalogue de Statistique Canada, Ottawa.

DAGUM, E.B. 1988. *The X11ARIMA/88 Seasonal Adjustment Method - Foundations and User's Manual, Time Series Research and Analysis Division*. Rapport technique de Statistique Canada.

DAGUM, E.B. 1996. « A New Method to Reduce Unwanted Ripples and Revisions in Trend-Cycle Estimates from X-11-ARIMA », *Survey Methodology*, nº 22, p. 77 à 83.

FINDLEY, D.F., B.C. MONSELL, W.R. BELL, M.C. OTTO et B.C. CHEN. 1998. « New Capabilities of the X-12-ARIMA Seasonal Adjustment Program (with Discussion) », *Journal of Business and Economic Statistics*, nº 16, p. 127 à 177.

GOMEZ, V. et A. MARAVALL. 1996. *Programs TRAMO (Time Series Regression with Arima Noise, Missing Observations, and Outliers) and SEATS (Signal Extraction in Arima Time Series). Instructions for the User.* Document de travail 9628 du Département de recherche de la Banque d'Espagne. Voir aussi le document en ligne : http://www.bde.es/servicio/software/tramo/summprogs.pdf.

LADIRAY, D. et B. QUENNEVILLE. 2001. « Seasonal Adjustment with the X-11 Method », *Lecture Notes in Statistics*, nº 158, New York, Springer-Verlag.