



5 - Les trois temps de la production

ANNA SMYK ET TANGUY BARTHÉLÉMY
Division Recueil et Traitement de l'Information
Département des Méthodes Statistiques

Objectifs de cette séquence

Cette séquence a pour but de vous présenter l'organisation d'un processus de production

Après cette séquence vous connaîtrez :

- les principales questions à se poser au moment de la mise en place d'un processus de désaisonnalisation
- les différentes étapes du processus, leurs objectifs et leur déroulement

Sommaire

1. Installation du processus de désaisonnalisation

1.1 Les questions, les choix et les difficultés

1.2 Approche directe ou indirecte ?

2. Les campagnes annuelles CVS-CJO

Les principales questions (1/2)

- Les besoins des utilisateurs
 - Quelles sont les variables d'intérêt à désaisonnaliser ?
 - À quel niveau ?
- Le choix de la méthode
 - X-13-ARIMA
 - TRAMO-SEATS
 - ...
- Le choix de l'outil
 - JDemetra+
(<https://jdemetradocumentation.github.io/JDemetra-documentation/>)
 - ...
- Le choix de l'approche : directe ou indirecte ? Des contraintes de diffusion ?

Les principales questions (2/2)

- La correction des effets de calendrier
 - Autorisée ou non ?
 - Quels jeux de régresseurs ?
- Quelle longueur d'estimation et quelle période de révision ?
Attention aux séries « trop longues » (>20 ans)

Les « réponses » ? Plutôt des recommandations. . .

Au sein du SSP, il est vivement conseillé de suivre les recommandations d'Eurostat en matière d'ajustement saisonnier (ESS guidelines on seasonal adjustment)

<https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-15-001>

(Philosophie, objectifs et contenu des guidelines : un aperçu dans ESS_Guidelines_2015_overview.pdf)

Approche Directe

Cadre type : séries à différents niveaux d'une nomenclature (ex Naf)

Approche Directe : traitement de désaisonnalisation indépendant à chaque niveau de la nomenclature

Exemple : $X.cvs.NAF.0510 = CVS(X.brute.NAF.0510)$

$X.cvs.A88|05 = CVS(X.brute.A88|05)$

Mais a priori $X.cvs.A88|05 \neq X.cvs.NAF.0510 + X.cvs.NAF.0520$

(NB. En réalité cette moyenne est probablement pondérée.)

Il est en général très important pour le producteur de conserver cette additivité en diffusion, pour les séries brutes comme pour les séries CVS.

Approche Indirecte

Approche Indirecte : Traitement cvs au niveau le plus bas uniquement, séries CVS des niveaux agrégés obtenues par somme pondérée des séries CVS des niveaux inférieurs

On a toujours $X.cvs.NAF.0510 = CVS(X.brute.NAF.0510)$

Mais $X.cvs.A88|05 = X.cvs.NAF.0510 + X.cvs.NAF.0520$

(nb : en réalité cette moyenne est probablement pondérée)

Ainsi, a priori :

$X.cvs.A88|05 \neq CVS(X.brute.A88|05)$

Approche Directe / indirecte : avantages et inconvénients

	Approche directe	Approche indirecte
Avantages	<ul style="list-style-type: none">– plus « lisible »– on supprime la composante saisonnière directement sur toutes les séries (y c niveau agrégé)	<ul style="list-style-type: none">– moins coûteux– additivité préservée (diffusion)
Inconvénients	<ul style="list-style-type: none">– plus coûteux– pas d'additivité	<p>peu de diagnostics au niveau agrégé : risque de saisonnalité résiduelle</p>

Sommaire

1. Installation du processus de désaisonnalisation

2. Les campagnes annuelles CVS-CJO

2.1 Pourquoi faire une campagne annuelle ?

2.2 Exemple de bilan qualité

2.3 Les campagnes mensuelles/trimestrielles CVS-CJO

Faire une campagne annuelle pour s'adapter aux nouvelles données

Une campagne annuelle CVS-CJO est une révision des modèles de pré-ajustement et des paramètres de la décomposition.

Pourquoi faire une campagne annuelle ?

- S'assurer de l'adaptation du modèle aux données (on dispose d'une année de données en plus. . .)
 - Pas d'effets saisonniers résiduels sur la série CVS-CJO
 - Pas d'effets de calendrier résiduels sur la série CVS-CJO
 - Qualité des prévisions issues du modèle Reg-ARIMA (révisions faibles)

Dans la pratique :

- on compare le modèle "en vigueur" avec un modèle automatique entièrement réestimé
- on utilise un ensemble de critères statistiques pondérés (Score).

Le bilan qualité utilisé au DMS

Un score de qualité est construit à partir des diagnostics et pondéré par le poids de la série (en diffusion)

Cela permet de répondre à 2 besoins :

- Valider les modèles et les filtres utilisés pour le traitement CVS-CJO
- Prioriser les séries qui feront l'objet d'une expertise « manuelle » dans le cadre d'une production de masse ("selective editing")

Choix (possible) de pondérations pour le score

Diagnostics		Pondération (sur 100)
Pré-ajustement	Résidus du modèle ARIMA	30
	Effets de calendrier résiduels	20
Décomposition (X11)	Effets saisonniers résiduels	45
	Qualité de la décomposition (stats M)	5

(cf Cruncher et Bilan Qualité)

Des campagnes infra-annuelles pour prendre en compte l'information la plus fraîche

Lorsqu'un nouveau point est disponible pour la série brute, il faut l'intégrer dans la série CVS, dont on doit rapidement publier une version actualisée.

Il y a plusieurs manières d'estimer la nouvelle série cvs, en donnant plus ou moins d'importance à l'information apportée par le nouveau point.

Un continuum d'options est disponible allant de :

- on utilise des coefficients saisonniers calculés lors de la dernière campagne annuelle (information du nouveau point non prise en compte)

à

- la remise en cause complète de l'estimation (information totalement prise en compte)

La démarche recommandée se situe en général entre ces deux extrêmes. (voir séquence suivante)

Les essentiels

La production des CVS-CJO s'organise en 3 temps :

- installation du processus de traitement (choix methodologiques structurants : algorithme, longueur des séries..)
- campagnes annuelles de mise à jour des régresseurs, modèles et filtres : révision complète
- campagnes de production mensuelles/trimestrielles : révision partielle, focus sur les dernières valeurs diffusées