

Les principales questions autour de la qualité

L'ajustement saisonnier au spectre de la qualité

Hien PHAM

Division Recueil et Traitement de l'Information
Département des Méthodes Statistiques



Mesurer pour comprendre



Plan de la présentation

- ◆ Introduction
- ◆ Quelle approche de la qualité ?
- ◆ Rappel de quelques notions techniques
- ◆ CVS et l'aspect « lisse »
- ◆ CVS et les révisions

La Correction des Variations Saisonnières, une pratique répandue et « banalisée »

Ce qui est « banal » devient un état de fait, détaché de sa finalité

Pourquoi désaisonnalise-t-on ?

Les réponses spontanées :

- ça fait partie des bonnes pratiques

La technique se justifie par elle-même

- « Eurostat nous le demande »

Une exigence institutionnelle, un fait institutionnalisé

- il y a une saisonnalité dans la série

Condition nécessaire..... mais pas suffisante

Ex : saisonnalité touristique a une utilité : décider des mesures pour stabiliser l'activité

En précisant la question : à quoi sert une série CVS ?

Mettre en évidence ce qui est « nouveau »

Ex : Etude de la Conjoncture

La saisonnalité est structurelle, elle persiste et se répète

Comparaisons temporelles

Lire aisément les phases d'expansion, de récession et points de retournement

Comparaisons spatiales

Entre pays n'ayant pas les mêmes climats, traditions, institutions....

La Qualité est une notion qui dépend de la Finalité

« *Quality is defined as the totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs.* » Source : *SDMX*

Quelle approche de la qualité ? 1/3

Il n'y a pas de convergence sur les critères de qualité

Exemple de Nerlove (1964) : critères fondés sur la comparaison des spectres

Penser la qualité comme une démarche opérationnelle pour :

Valider la production

Besoin de critères pour décider Ok / Nok pour diffuser

Harmoniser les pratiques au sein du SSP

Besoin de critères qui s'appliquent aux différents systèmes de production, répondant à des besoins différents

Prioriser les traitements

Besoin de critères qui optimisent l'utilisation des ressources

Cf. présentation « Démarche du diagnostic de la qualité »

Quelle approche de la qualité ? 2/3

Penser la qualité comme un compromis entre la « performance technique » et la « finalité » :

Proximité des modèles aux données

Considérations techniques basées sur un ensemble d'indicateurs statistiques et des recommandations européennes (guidelines)

Proximité de la production aux besoins des utilisateurs

Considérations pragmatiques basées sur l'analyse des besoins

Cf. présentation de la Dares

Compromis car ces deux aspects peuvent être en concurrence :

Exemple : Approches Directe versus Indirecte

Quelle approche de la qualité ? 3/3

Penser la qualité comme un processus d'amélioration continu ...

... enchaînement ordonné et itératif de 3 étapes :

Installation du processus de désaisonnalisation

Quelle méthode, quel outil, quelle imbrication.... ??

Bien connaître les besoins, les recommandations,

Phase structurante et coûteuse, nombreux tests et simulations

Cf. présentation Dera & Aco (+processus harmonisé)

Révisions annuelles des modèles

Adapter les modèles aux données récentes

Outil « Bilan Qualité » du DMS

Mettre en concurrence « modèle optimisé » et « précédent modèle »

Productions mensuelles ou trimestrielles

Vérifier et valider les éléments nouveaux : outliers, révisions

Diagnostics simples produits par JDemetra+

Quelle approche de la qualité ?

Réponse :

La qualité est une **démarche opérationnelle**,
résultant d'un **compromis** entre la plus grande fiabilité
statistique et la plus grande satisfaction des utilisateurs,
et qui s'inscrit dans un **processus d'amélioration continu**.

Deux sujets plus précis de discussion

- ◆ Une bonne CVS est-elle lisse ?
- ◆ Une bonne CVS est-elle stable ?

Rappel de quelques notions techniques

Quelle démarche pour obtenir une CVS ?

Modéliser, Décomposer et Estimer

X observée = Tendance + Cycle + Saisonnalité + Irrégulier (schéma additif)

Objectif de la technique : bien estimer **S** \rightarrow CVS = $X - S$

Par quelles méthodes ?

Deux méthodes recommandées : X12-ARIMA et TRAMO-SEATS

Point commun : même structure en 2 phases

1. Pré-ajustement : nettoie et prolonge la série initiale avant filtrage
2. Décomposition : applique des filtres pour extraire les composantes

Différences importantes dans la méthode de Décomposition

Décomposition X12 (ou méthode X11)

En entrée du module : série linéarisée et prolongée issue de la phase pré-ajustement

On applique à cette série, une succession de moyennes mobiles pour :

Extraire une estimation de la Tendance-Cycle de la série

Ce qui donne $X - TC = S + I$

Extraire la composante saisonnière de $S + I$

Ce qui donne $CVS = X - S$

Ce procédé est mis en œuvre en plusieurs itérations pour obtenir des estimations finales

On parle de méthode « non paramétrique », i.e. qu'il n'y a pas de modèle explicite dont les paramètres sont à estimer, même si l'application successive des MM revient à appliquer un modèle sous-jacent.

Décomposition SEATS

En entrée : la série linéarisée et prolongée et le modèle ARIMA de la série

A partir du modèle ARIMA de la série linéarisée, un modèle ARIMA pour chaque composante est déduit

En répartissant les racines du polynôme AR de la série linéarisée (selon sa fréquence dans le spectre)

En choisissant de maximiser la variance du bruit (décomposition canonique)

A partir du modèle ARIMA des composantes, un filtre de type Wiener-Kolmogorov est déduit pour chaque composante

Ces filtres appliqués à la série linéarisée donnent les composantes estimées

On parle de méthode « paramétrique », i.e. chaque composante peut être décrite par un modèle explicite dont on estime les paramètres.

Une bonne CVS est-elle lisse ?

1/5

Série Brute et CVS de l'IPI, Fabrication de meubles de cuisine

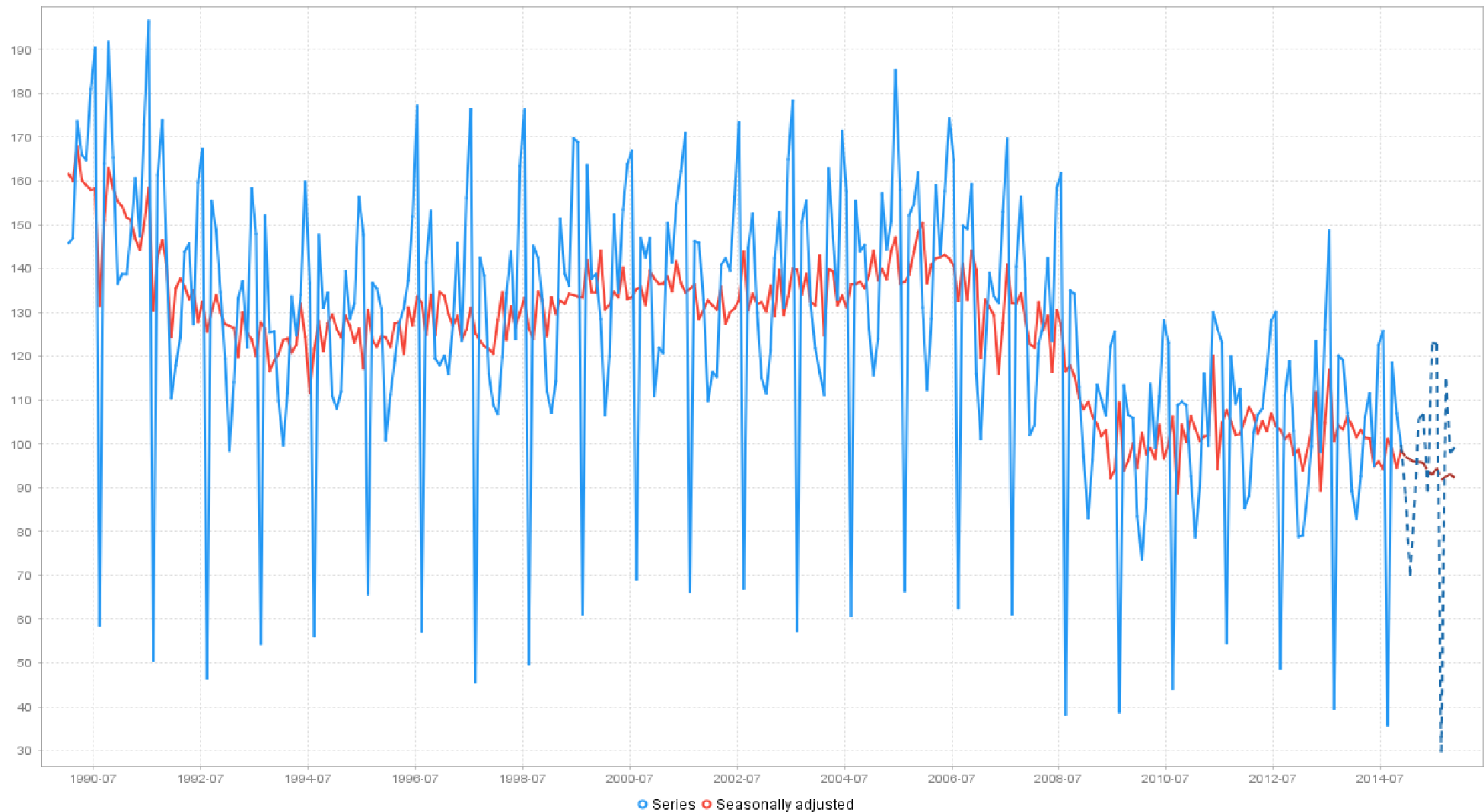


Schéma de décomposition

$$X = TC + S + I \quad \text{et} \quad CVS = TC + I$$

La part relative de I est plus importante dans la CVS que dans X

Qu'est-ce qui caractérise S ?

Régularité des coefficients saisonniers mensuels
et Fréquence annuelle

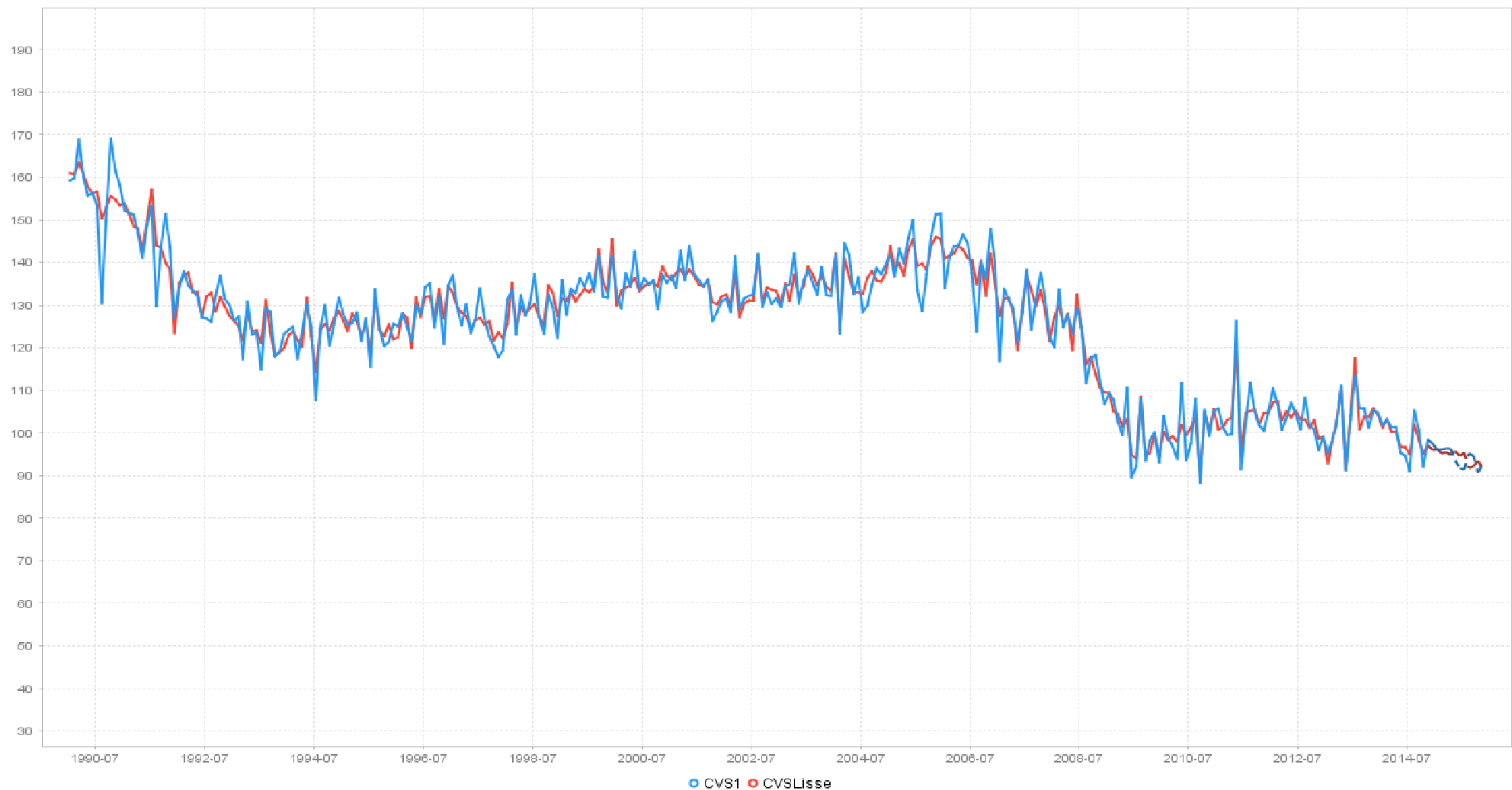
Estimer S, avec des coefficients mensuels instables va à l'encontre de sa « nature »

Par construction, la méthode X12-ARIMA lisse les coefficients de S

Une bonne CVS est-elle lisse ?

3/5

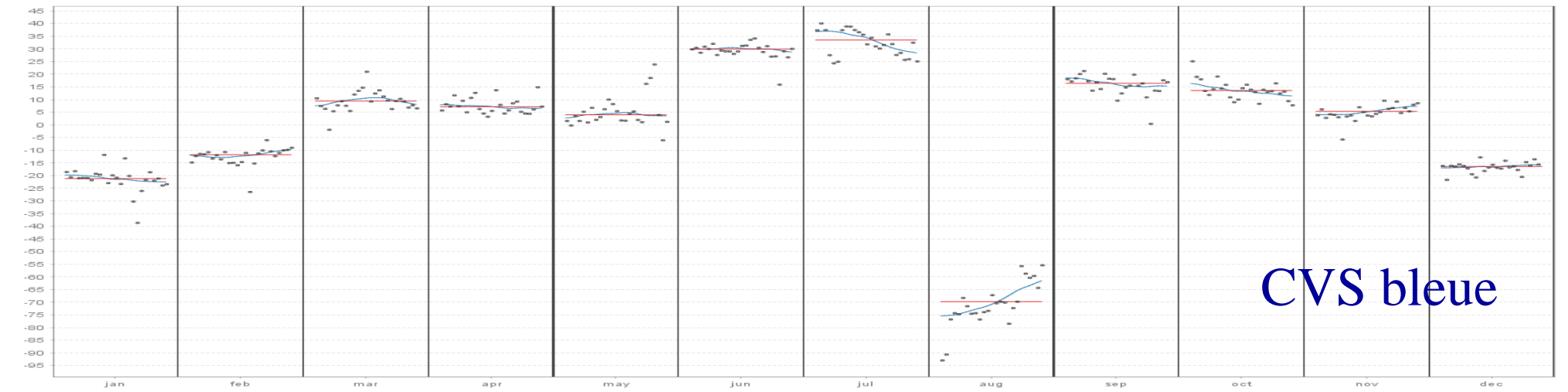
Deux versions de CVS, la rouge plus lisse



Une bonne CVS est-elle lisse ?

4/5

Les coefficients saisonniers de la CVS lisse plus instables



CVS bleue



CVS rouge

Conclusion

Une CVS plus lisse n'est pas nécessairement meilleure
On ne peut pas non plus conclure l'inverse

Le caractère « lisse » d'une série CVS n'est tout
simplement pas un critère d'évaluation de la qualité...
Si la finalité est de corriger la saisonnalité

Révisions totales :

- révisions liées aux révisions des données brutes
- révisions liées aux calculs des CVS (Révisions CVS)

On suppose ici que la série brute n'est pas révisée

Révisions CVS :

- dues au changement de modèle (affecte toute la série, rév. annuelles)
- dues aux ajouts de nvx points (affecte fin de série, rév. infra-annuelle)

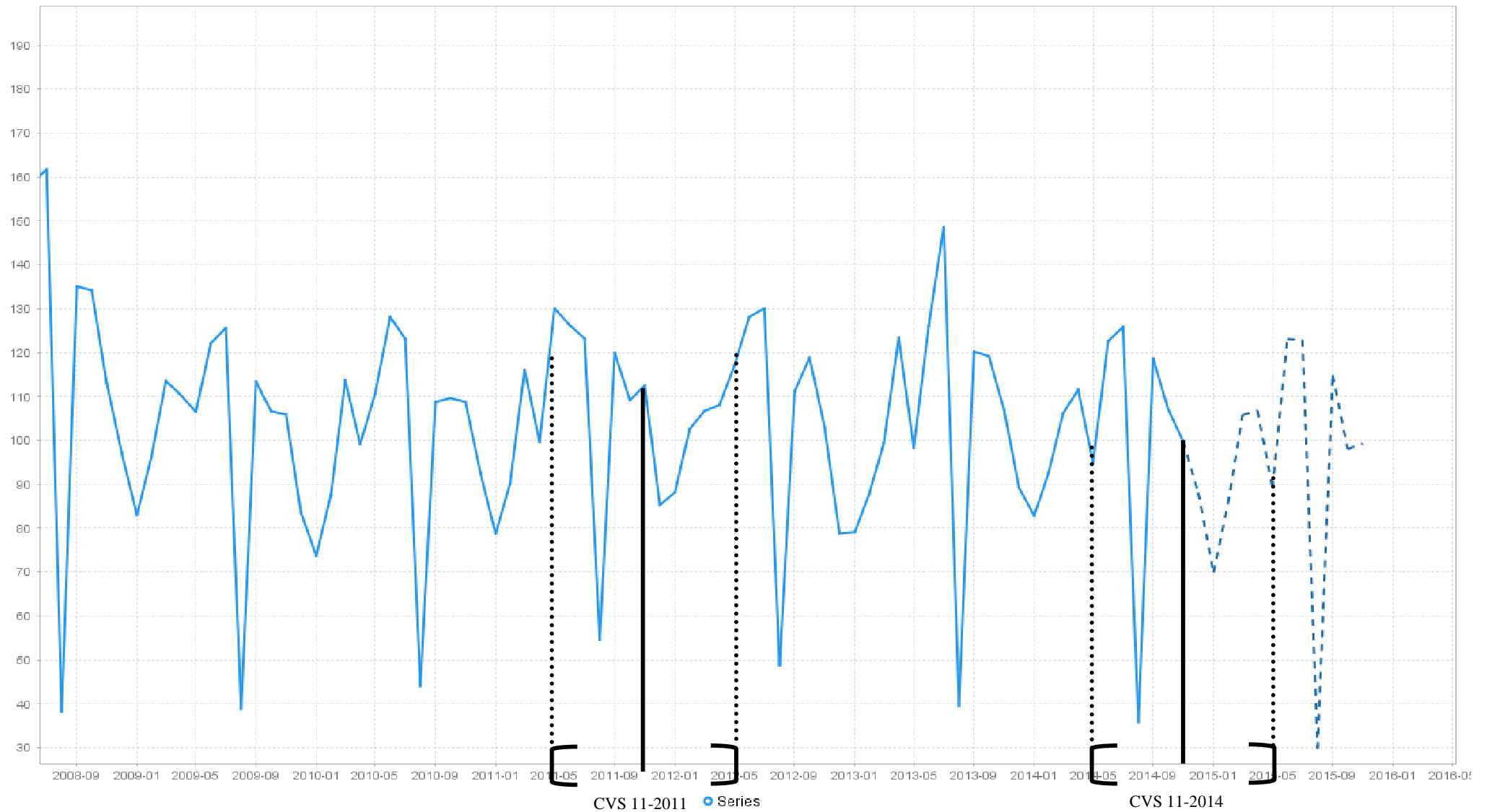
On s'intéresse ici plus particulièrement au deuxième cas

Exemple avec la méthode TRAMO-SEATS :

- filtre symétrique (de type Wiener-Kolmogorov)
- filtre fini (longueur 13 par exemple)
- on peut dériver un filtre de type WK pour la CVS

Une bonne CVS est-elle stable ?

2/4



Au point 11-2011 estimation finale :

$$CVS_{11-2011/11-2014} = v_6 x_{05-2011} + v_5 x_{06-2011} + \dots + v_1 x_{10-2011} + v_0 x_{11-2011} + v_1 x_{12-2011} \dots + v_5 x_{04-2012} + v_6 x_{05-2012}$$

Au point 11-2014 estimation préliminaire :

$$CVS_{11-2014/11-2014} = v_6 x_{05-2014} + v_5 x_{06-2014} + \dots + v_1 x_{10-2014} + v_0 x_{11-2014} + v_1 \hat{x}_{12-2014/11-2014} \dots + v_5 \hat{x}_{04-2015/11-2014} + v_6 \hat{x}_{05-2015/11-2014}$$

$$CVS_{11-2014/12-2014} = \dots + v_1 x_{12-2014} \dots + v_5 \hat{x}_{04-2015/12-2014} + v_6 \hat{x}_{05-2015/12-2014}$$

$$CVS_{11-2014/04-2015} = \dots + v_1 x_{12-2014} \dots + v_5 x_{04-2015} + v_6 \hat{x}_{05-2015/04-2015}$$

$$CVS_{11-2014/05-2015} = \dots + v_1 x_{12-2014} \dots + v_5 x_{04-2015} + v_6 x_{05-2015}$$

Si on note l'erreur de prévision à une date j $e_j = x_j - \hat{x}_{j/11-2014}$

Révision finale du point 11-2014 dépend de 2 paramètres :

- nature stochastique de la série (modèle)
- erreurs de prévisions (donc à minimiser)

$$R_{11-2014} = v_1 e_{12-2014} \dots + v_5 e_{04-2015} + v_6 e_{05-2015}$$

Révisions successives dues aux nouvelles obs + maj prévisions,
ne pas les prendre en compte c'est ignorer de l'information récente

Conclusion :

Par construction, la fin de série CVS est révisée avec un nv. point.

Ce mécanisme traduit la prise en compte de nouvelles info.

Juger la qualité d'une CVS uniquement par la « quantité » de révisions, c'est comme juger la qualité d'un estimateur par sa vitesse de convergence, quand bien même il serait biaisé.

Les différentes options de rafraichissement :

Méthode Current : coeff S projetés et figés, aucune révision CVS en infra-annuelle

Méthode Partial-Concurrent : coeff S ré-estimés, avec légère maj du modèle (paramètres + outliers)

Méthode Concurrent : coeff S ré-estimés avec modèle ré-identifié

Fin

Merci de votre attention !

Contact

M. Hien PHAM

Tél. : 01 41 17 63 67

Courriel : hien.pham@insee.fr

Insee

18 bd Adolphe-Pinard
75675 Paris Cedex 14

www.insee.fr  

Informations statistiques :

www.insee.fr / Contacter l'Insee

09 72 72 4000

(coût d'un appel local)

du lundi au vendredi de 9h00 à 17h00