
La correction des ruptures de profil saisonnier

Avantages et Inconvénients

Plan

La présentation comporte 4 parties

1. Qu'est ce qu'un SO (Seasonal Outlier) ?
2. Quels effets l'introduction des SO a sur les modèles ?
 - 2.1 - dans la phase de pré-ajustement
 - 2.2 - dans la phase de décomposition
3. Quels sont les risques éventuels de leur emploi ?
4. Dans quelles situations les utiliser ?

Rappels

1/2

Désaisonnalisation en deux étapes avec X13-ARIMA

Phase de pré-ajustement : linéariser la série avec un modèle Reg-ARIMA : correction des effets de calendrier, correction des points atypiques, prévision pour prolonger la série linéarisée

Phase de décomposition : module X11

Problématique : Quels sont les effets des SO sur chacune de ces étapes ?

Les points atypiques (outliers)

Les **Additive Outliers** affectent l'Irrégulier,

Les **Level Shift** et **Transitory Change** affectent la Tendence

Les **Seasonal Outliers** affectent la composante saisonnière



3

Insee

Rappels

2/2

Un phénomène saisonnier

Fluctuations à l'intérieur d'une année, mensuelles ou trimestrielles, qui se répètent plus ou moins régulièrement d'une année sur l'autre.

Modèle déterministe : $\sum_{i=1}^s c_i = 0$ Saisonnalité constante

Modèle stochastique : $\sum_{i=0}^{s-1} c_{t-i} = \varepsilon_t$

→ Espérance du bruit nulle, variance petite

→ La saisonnalité évolue lentement, i.e. localement stable

Les méthodes de désaisonnalisation, en particulier X11, s'appuient sur l'hypothèse d'une saisonnalité localement stable

« Les filtres saisonniers utilisés par X11 fonctionnent de façon optimale pour les saisonnalités constantes. Si ce mouvement saisonnier évolue au cours des ans, les estimations des coefficients saisonniers peuvent être erronées. » Ladiray et Quenneville



4

Insee

1. Qu'est ce qu'un SO ?

1/4

Rupture soudaine et durable du profil saisonnier

Rupture du niveau d'un ou de plusieurs coefficients saisonniers

Déphasage : les amplitudes sont inchangées, mais le profil saisonnier est décalé d'une ou plusieurs périodes – type de rupture peu fréquent

C'est un changement qui perdure (sinon AO)

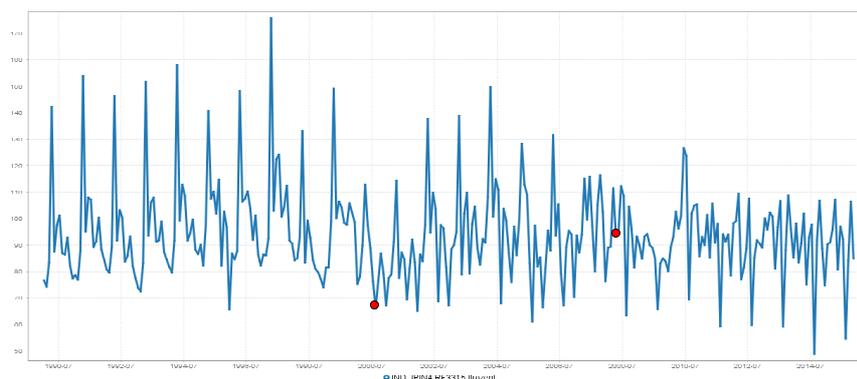
Exemples :

- cause économique : changement de période de fermeture annuelle qui entraîne une répartition différente des creux d'activité
- cause institutionnelle : suppression des primes versées aux fonctionnaires
- cause technique : changement de méthode de traitement des données

1. Qu'est ce qu'un SO ?

2/4

Série brute IPI 3315 « réparation et maintenance navale »



Disparition des pics d'activité en avril 2008
Apparition de creux d'activité en août, à partir de 2000

1. Qu'est ce qu'un SO ?

3/4

Comment le modélise-t-on ?

Rupture du profil saisonnier en t_0 , $SO_{t_0}(t) = \begin{cases} 0 & \text{si } t \geq t_0 \\ 1 & \text{si } t < t_0 \text{ et } \text{mois}(t) = \text{mois}(t_0) \\ \frac{-1}{s-1} & \text{sinon} \end{cases}$
« s » la périodicité

Régresseur introduit dans le modèle Reg-ARIMA, le coefficient de régression estimé = ampleur du changement de niveau du coefficient saisonnier en t_0

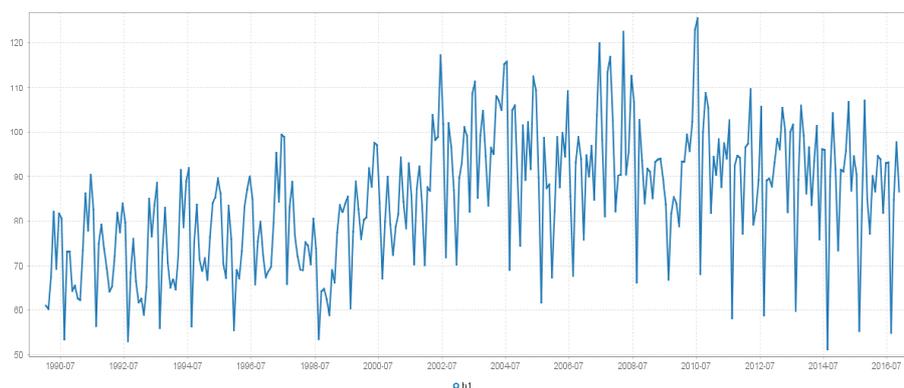
Les points antérieurs à t_0 sont corrigés, les points postérieurs à t_0 ne sont pas modifiés – idée « on corrige le profil du passé pour le faire ressembler à celui du présent »

Somme des valeurs du régresseur sur une année est nulle – sauf quand l'intervalle contient à la fois des points antérieurs et postérieurs à t_0 . On préserve une saisonnalité constante, sauf autour du point t_0

1. Qu'est ce qu'un SO ?

4/4

Série linéarisée (CJO et corrigée des outliers) avec la correction des SO



Profil saisonnier devient stable : les pics d'avril disparaissent et les creux d'août apparaissent sur le passé.

2.1 Effets sur le pré-ajustement ? 1/6

La modélisation des SO améliore le pré-ajustement

Ces résultats ont été démontrés par Kaiser et Maravall (2001).

Sans correction des SO :

- les résidus des modèles présentent plus d'autocorrélation
- l'estimation des paramètres MA saisonniers d'un modèle Airline est biaisée

Les tests empiriques réalisés sur 200 séries IPI (dont 93 présentent des SO) donnent des résultats qui vont dans le même sens. Plusieurs indicateurs ont été comparés pour évaluer la validité du modèle :

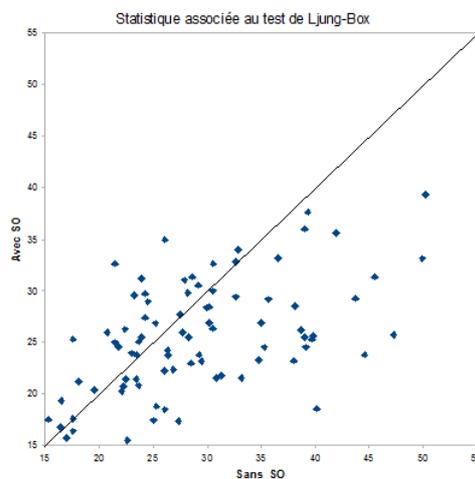
- Autocorrélation des résidus
- Homoscédasticité
- Normalité

Et pour évaluer la performance du modèle :

- AIC
- EQM des révisions

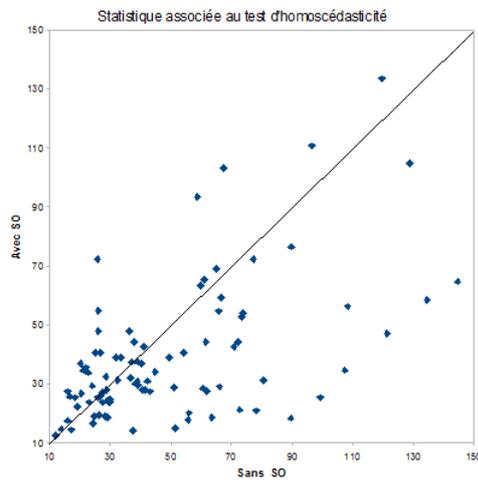
2.1 Effets sur le pré-ajustement ? 2/6

Test de Ljung-Box (autocorrélation des résidus)



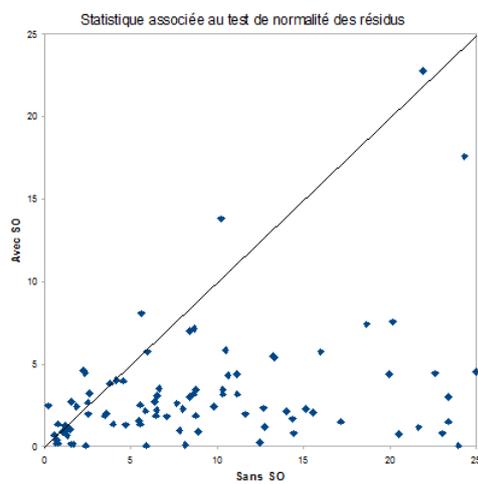
2.1 Effets sur le pré-ajustement ? 3/6

Test d'homoscédasticité



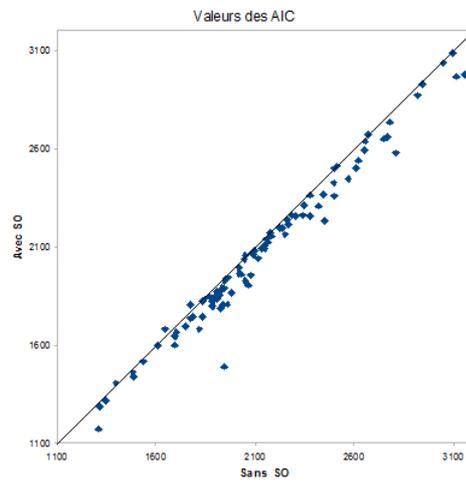
2.1 Effets sur le pré-ajustement ? 4/6

Test de normalité



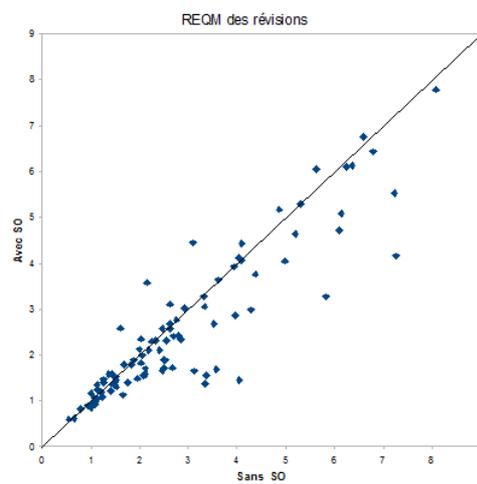
2.1 Effets sur le pré-ajustement ? 5/6

Critère AIC (sur les séries ayant les mêmes ordres de différenciation)



2.1 Effets sur le pré-ajustement ? 6/6

Racine carrée de EQM des révisions



2.2 Effets sur la décomposition ? 1/6

La modélisation des SO améliore la décomposition

Méthode paramétrique

Pour la méthode Tramo-Seats la qualité de la décomposition dépend de la qualité du modèle ARIMA.

Méthode non paramétrique

X11 utilise une succession de moyennes mobiles pour extraire les composantes.

Les moyennes mobiles produisent des coefficients saisonniers lissés (cohérent avec l'hypothèse de stabilité locale du profil saisonnier).

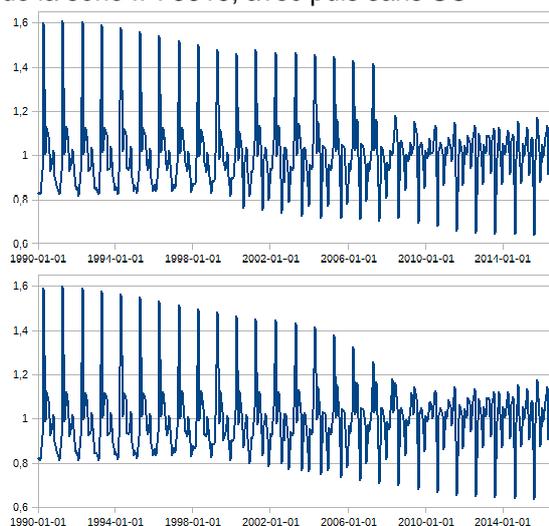
Par conséquent une rupture de profil saisonnier ne sera pas bien restituée, mais sera lissée.

2.2 Effets sur la décomposition ? 2/6

Exemple : profil saisonnier de la série IPI 3315, avec puis sans SO

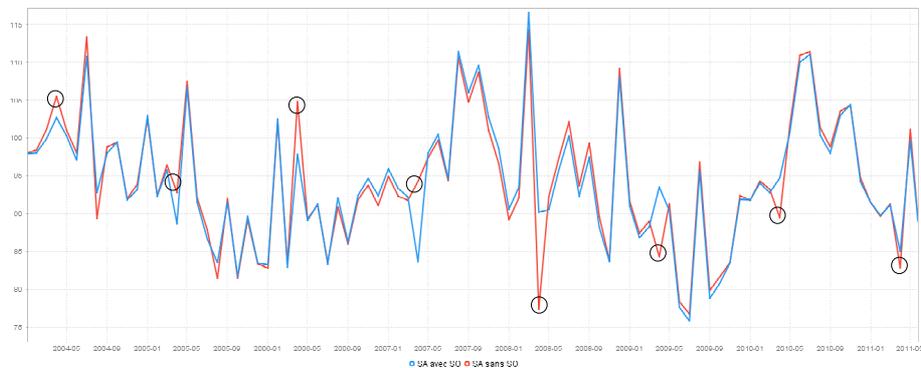
Avec SO,
coeffs avril sup à 1,4
jusqu'en 2007,
Puis chutent à 1 en 2008

Sans SO,
coeffs avril diminuent
progressivement
en 2007 = 1,25
en 2008 = 1,18
Coeffs sous-estimés
2003 → 2007,
puis ensuite surestimés



2.2 Effets sur la décomposition ? 3/6

CVS-CJO de la série IPI 3315, avec et sans SO



Sans correction du SO, sur la CVS, les points d'avril sont surestimés jusqu'en 2007. Puis, à partir de 2008 ils sont sous-estimés.

Saisonnalité résiduelle

2.2 Effets sur la décomposition ? 4/6

Biais sur les coeffs saisonniers, quand SO non détectés

Si la rupture de niveau d'un coeff saisonnier est Δ , l'utilisation d'une moyenne mobile M3x3 (5 termes, courte et s'adapte vite) entraîne un biais de :

- 1/9 de Δ , 2 ans avant le choc (apparition progressive avant la date du choc)
- 1/3 de Δ , 1 an avant le choc
- à la date du choc, ce dernier n'est restitué qu'à 2/3
- 1 an après, le choc est restitué à 8/9
- c'est seulement 2 an après qu'on retrouve le niveau du choc

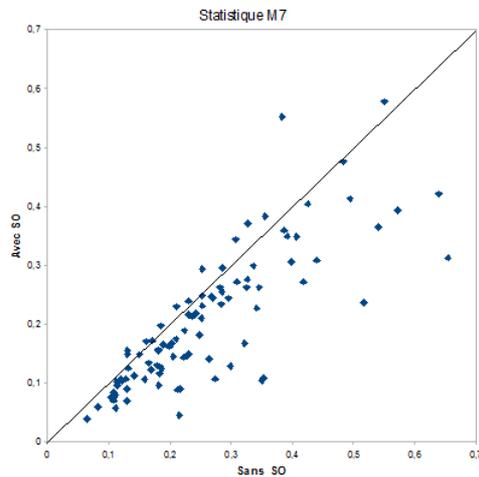
Les moyennes mobiles plus longues M3x5 et M3x9 (resp. 7 et 11 termes) produisent des biais sur de plus longues périodes (resp. 3 ans et 5 ans avant et après le choc)

2.2 Effets sur la décomposition ? 5/6

La modélisation des SO améliore la statistique M7

M7 mesure la part relative
des saisonnalités
stable et mobile

Les petites valeurs sont
préférables



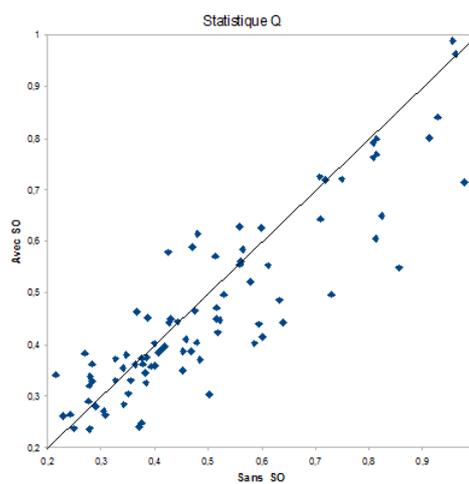
2.2 Effets sur la décomposition ? 6/6

La stat Q

Q est une moyenne pondérée
des statistiques M

C'est un indicateur synthétique
de la qualité de décomposition
réalisée par X11

Les petites valeurs sont
préférables



3. Quels risques à modéliser les SO ? 1/5

Décalage de niveau des CVS dans le cas de décomposition X11 des modèles multiplicatifs

La modélisation SO devrait « corriger » la rupture du profil saisonnier sans affecter le niveau de la série, i.e. la tendance.

Quand le modèle est additif : $Z_t = Y_t + \alpha SO_{t_0}(t)$

La première estimation de la tendance par M2x12 donne le même résultat pour Y ou Z. Car la M2x12 supprime le terme en SO (sauf autour du t_0).

La tendance de la série corrigée Z est la même que celle de la série initiale Y

3. Quels risques à modéliser les SO ? 2/5

Quand le modèle est multiplicatif la régression est faite sur le ln :

$$\ln(Z_t) = \ln(Y_t) + \alpha SO_{t_0}(t)$$

X11 va ensuite décomposer la série : $Z_t = Y_t * \exp(\alpha SO_{t_0}(t))$

Pour les dates postérieure à t_0 , pas de problème, les valeurs du SO sont nulles. Donc les niveaux de Y et Z coïncident.

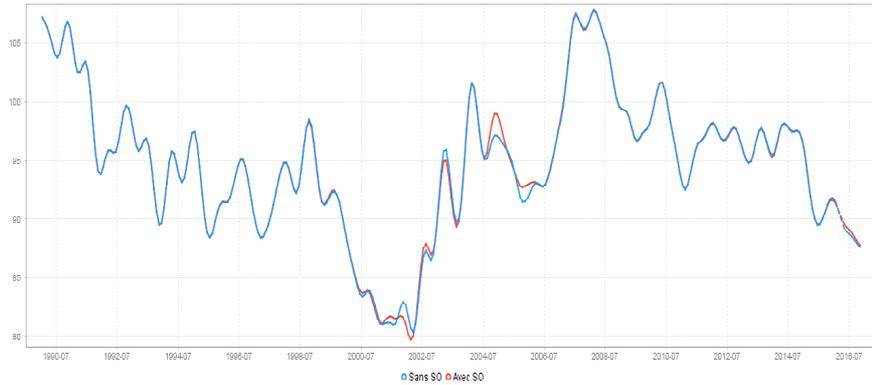
Pour les dates antérieures à t_0 , contrairement au cas additif, la moyenne M2x12 n'élimine pas le terme en SO. Donc le niveau de Z n'est plus égal à celui de Y sur les périodes antérieures au choc.

→ **Biais sur le niveau**, qui dépend du paramètre alpha

3. Quels risques à modéliser les SO ? 3/5

Tendances de la série IPI – 1395 : schéma additif

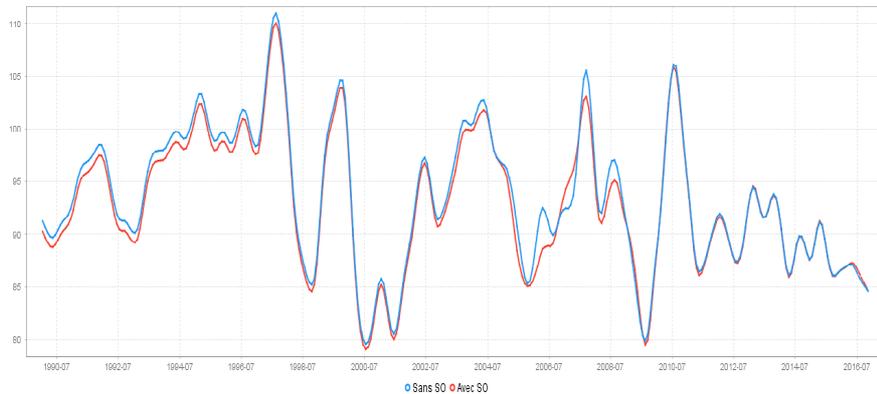
Un SO en décembre 2002



3. Quels risques à modéliser les SO ? 4/5

Tendances de la série IPI – 3315 : schéma multiplicatif

SO en août 2000 et avril 2008



3. Quels risques à modéliser les SO ? 5/5

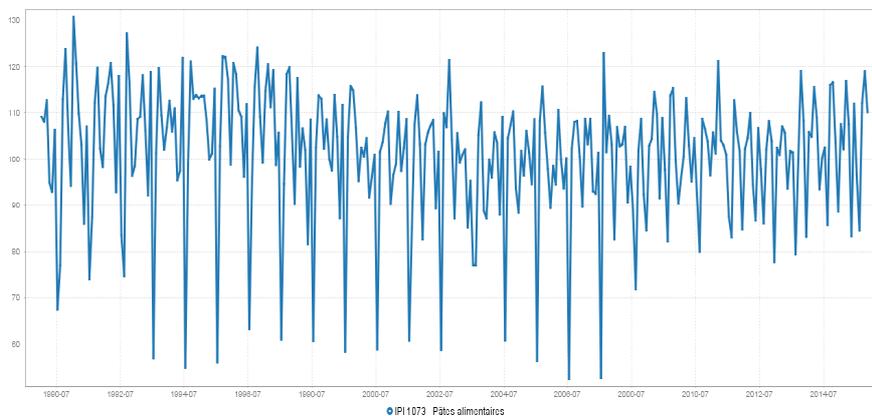
La gestion de ce type de ruptures est plus compliquée que pour les autres types d'outliers : pertinence ou non de leur intégration, lecture plus difficile du niveau de correction, etc.

→ Nous verrons cela pendant l'atelier de cet AM

4. Quand faire la détection des SO ? 1/4

Ruptures visibles du profil saisonnier sur les graphiques

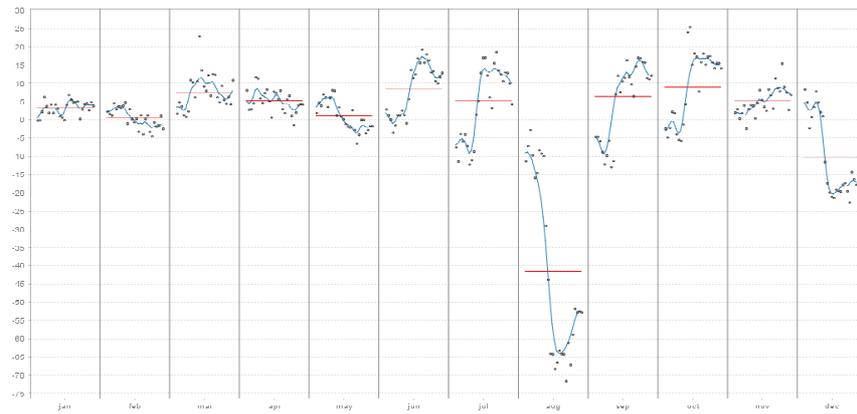
Sur la série brute : juillet et août 1993, puis juillet et août 2008



4. Quand faire la détection des SO ? 2/4

Ruptures visibles du profil saisonnier sur les graphiques

Graphique des SI-ratios – IPI 2312, façonnage et transformation de verre
Ex : août 99



4. Quand faire la détection des SO ? 3/4

Saisonnalité résiduelle en fin de série (due au modèle Reg-ARIMA)

CVS-CJO – IPI 1724, fabrication de papier peint (SO 08/2013)



4. Quand faire la détection des SO ? 4/4

Des statistiques M de mauvaise qualité

Quand la statistique M7 (part relative des saisonnalités stable et mobile) est supérieure à 1, on peut essayer de chercher la présence éventuelle de SO.

Les statistiques M8 à M11, qui mesurent l'ampleur des variations de la composante saisonnière, sont assez piégeuses à interpréter dans le cas où l'on corrige des ruptures de profil saisonnier.

→ Nous verrons cela pendant l'atelier de cet AM

Conclusion

La détection des SO peut se faire lors des campagnes annuelles

Une rupture est un SO que si le nouveau profil est confirmé l'année d'après → il y aura donc forcément des révisions

La détection des SO est ciblée sur certaines séries (cf. partie précédente)

Ne pas faire de détection automatique systématique

En présence de rupture du profil saisonnier, sans modélisation des SO, risque de saisonnalité résiduelle

Soit du fait de la décomposition réalisée par X11

Soit du fait de la modélisation Reg-ARIMA